MITei
MIT Energy Initiative

UTILITY OF THE FUTURE

An MIT Energy Initiative response to an industry in transition

In collaboration with IIT-Comillas

Universidad Pontificia Comillas ICAI ICADE Madrid
Instituto de Investigación Tecnológica

MIT

Full report can be found at: energy.mit.edu/uof

Incorporated in the cover art is an
image of a voltage tower. © iStock
and an aerial view of buildings © Shutterstock

ISBN (978-0-692-80824-5)

UTILITY OF THE FUTURE

An MIT Energy Initiative response to an industry in transition

Study Participants

Principal Investigators

IGNACIO PÉREZ-ARRIAGA
Professor, Electrical Engineering, Institute for Research in Technology, Comillas Pontifical University

Visiting Professor, MIT Energy Initiative

CHRISTOPHER KNITTEL
George P. Shultz Professor of Applied Economics, Sloan School of Management, MIT

Director, Center for Energy and Environmental Policy Research, MIT

Project Directors

RAANAN MILLER
Executive Director, Utility of the Future Study, MIT Energy Initiative

RICHARD TABORS
Visiting Scholar, MIT Energy Initiative

Research Team

ASHWINI BHARATKUMAR
PhD Student, Institute for Data, Systems, and Society, MIT

MICHAEL BIRK
SM, Technology and Policy Program ('16), MIT

SCOTT BURGER
PhD Student, Institute for Data, Systems, and Society, MIT

JOSÉ PABLO CHAVES
Associate Research Professor, Institute for Research in Technology, Comillas Pontifical University

PABLO DUENAS-MARTINEZ
Postdoctoral Associate, MIT Energy Initiative

IGNACIO HERRERO
Research Assistant, Institute for Research in Technology, Comillas Pontifical University

SAM HUNTINGTON
SM, Technology and Policy Program ('16), MIT

JESSE JENKINS
PhD Candidate, Institute for Data, Systems and Society, MIT

MAX LUKE
SM, Technology and Policy Program ('16), MIT

RAANAN MILLER
Executive Director, Utility of the Future Study, MIT Energy Initiative

PABLO RODILLA
Research Scientist, Institute for Research in Technology, Comillas Pontifical University

RICHARD TABORS
Visiting Scholar, MIT Energy Initiative

KAREN TAPIA-AHUMADA
Research Scientist, MIT Energy Initiative

CLAUDIO VERGARA
Postdoctoral Associate, MIT Energy Initiative

NORA XU
SM, Technology and Policy Program ('16), MIT

Faculty Committee

ROBERT ARMSTRONG

Director, MIT Energy Initiative

CARLOS BATLLE

Research Scholar, MIT Energy Initiative

Professor, Institute for Research in Technology, Comillas Pontifical University

MICHAEL CARAMANIS

Professor of Mechanical Engineering and Systems Engineering, College of Engineering, Boston University

JOHN DEUTCH

Institute Professor, Department of Chemistry, MIT

TOMÁS GÓMEZ

Professor, Institute for Research in Technology, Comillas Pontifical University

WILLIAM HOGAN

Raymond Plank Professor of Global Energy Policy, John F. Kennedy School of Government, Harvard University

STEVEN LEEB

Professor, Electrical Engineering & Computer Science and Mechanical Engineering, MIT

RICHARD LESTER

Associate Provost and Japan Steel Industry Professor of Nuclear Science and Engineering, Office of the Provost, MIT

LESLIE NORFORD

Professor, Department of Architecture, MIT

JOHN PARSONS

Senior Lecturer, Sloan School of Management, MIT

RICHARD SCHMALENSEE

Howard W. Johnson Professor of Economics and Management, Emeritus

Dean, Emeritus, Sloan School of Management, MIT

Research and Project Advisors

LOUIS CARRANZA

Associate Director, MIT Energy Initiative

STEPHEN CONNORS

Director, Analysis Group for Regional Energy Alternatives, MIT Energy Initiative

CYRIL DRAFFIN

Project Advisor, MIT Energy Initiative

PAUL MCMANUS

Master Lecturer, Questrom School of Business, Boston University

ÁLVARO SÁNCHEZ MIRALLES

Senior Associate Professor, Institute for Research in Technology, Comillas Pontifical University

FRANCIS O'SULLIVAN

Research Director, MIT Energy Initiative

ROBERT STONER

Deputy Director for Science and Technology, MIT Energy Initiative

Advisory Committee

PHILIP SHARP
Utility of the Future Advisory Committee Chair and Retired President, Resources for the Future

RICHARD O'NEILL
Utility of the Future Advisory Committee Vice Chair
Chief Economic Advisor, Federal Energy Regulatory Commission

M. MASOOD AHMAD
Division Head, Power Systems Planning Department, Saudi Aramco

JANET GAIL BESSER
Executive Vice President, Northeast Clean Energy Council

ALAIN BURTIN
Director, Energy Management, EDF R&D*

PAUL CENTOLELLA
President, Paul Centolella & Associates LLC
Senior Consultant, Tabors Caramanis Rudkevich

MARTIN CROUCH
Head of Profession for Economists and Senior Partner, Improving Regulation, Ofgem

ELIZABETH ENDLER
Research Program Manager, Shell International Exploration & Production (US) Inc.

PHIL GIUDICE
Chief Executive Officer, President and Board Member, Ambri Inc.

TIMOTHY HEALY
Chief Executive Officer, Chairman and Co-Founder, EnerNOC

MARIANA HEINRICH
Manager, Climate & Energy, World Business Council for Sustainable Development

PAUL JOSKOW
President, Alfred P. Sloan Foundation
Elizabeth and James Killian Professor of Economics, Emeritus, MIT

MELANIE KENDERDINE
Director of the Office of Energy Policy and Systems Analysis and Energy Counselor to the Secretary, U.S. Department of Energy

CHRISTIANA LA MARCA
Head of Innovation, Global Thermal Generation, Enel

ALEX LASKEY
President & Founder, Opower*

ANDREW LEVITT
Sr. Market Strategist, PJM Interconnection*

WILLIAM LILLEY
Renewable and Distributed Energy Developer, Saudi Aramco

BLANCA LOSADA MARTIN
Chief Executive Officer, Gas Natural Fenosa Engineering.
Chief Technology Officer, Gas Natural Fenosa

LUCA LO SCHIAVO
Deputy Director, Infrastructure Regulation, AEEGSI — Italian Regulatory Authority for Electricity, Gas, and Water

GARY RAHL
Executive Vice President, Booz Allen Hamilton

MARK RUTH
Principal Project Lead, Strategic Energy Analysis Center, National Renewable Energy Laboratory

MIGUEL SÁNCHEZ-FORNIE
Director, Global Smart Grids, Iberdrola

MANUEL SÁNCHEZ-JIMÉNEZ, PHD
Team Leader Smart Grids, European Commission

LAURENT YANA
Director Advisor of Global BUs, Group Strategy Division, Engie

AUDREY ZIBELMAN
Chair, New York State Public Service Commission

*** The MIT Energy Initiative also wishes to thank the following former Advisory Committee members:**

REGINE BELHOMME (EDF)

SUSAN COVINO (PJM)

JIM KAPSIS (Opower)

Contents

VI Foreword and Acknowledgments

VIII Executive Summary

1 **Part 1: Understanding Electricity Services and How Distributed Energy Resources Affect the Design and Operation of Power Systems**

1 A Power Sector in Transition

19 New Options for the Provision and Consumption of Electricity Services

35 Envisioning a Future with Distributed Energy Resources

75 **PART 2: A Framework for an Efficient and Evolving Power System**

75 A Comprehensive and Efficient System of Prices and Regulated Charges for Electricity Services

137 The Future of the Regulated Network Utility Business Model

185 Restructuring Revisited: Electricity Industry Structure in a More Distributed Future

227 The Re-Evolution of Short- and Long-Term Electricity Market Design

265 **Part 3: Insights on the Economics of Distributed Energy Resources and the Competition between Centralized and Distributed Resources**

265 Understanding the Value of Distributed Energy Resources

307 **Part 4: A Policy and Regulatory Toolkit for the Future Power System**

307 A Toolkit for Regulators and Policy Makers

325 **Appendix A**

325 A Description of the Computational Models Used in the Utility of the Future Study

339 **Appendix B**

339 A Review of Business Models for Distributed Energy Resources

Full report and related working papers are available at: ***energy.mit.edu/uof***

Foreword and Acknowledgments

An important evolution in the provision and consumption of electricity services is now under way, driven to a significant degree by a confluence of factors affecting the distribution side of the power system. A range of more distributed technologies — including flexible demand, distributed generation, energy storage, and advanced power electronics and control devices — is creating new options for the provision and consumption of electricity services. In many cases, these novel resources are enabled by increasingly affordable and ubiquitous information and communication technologies and by the growing digitalization of power systems. In light of these developments, the MIT Energy Initiative's *Utility of the Future* study examines how the provision and consumption of electricity services is likely to evolve over the next 10 to 15 years in different parts of the world and under diverse regulatory regimes, with a focus on the United States and Europe.

The Utility of the Future study is the first of a new series of reports that is being produced by the MIT Energy Initiative (MITEI) to serve as balanced, fact-based, and analysis-driven guides to key topic areas in energy for a wide range of decision makers in government and industry. This study specifically aims to serve as a guide for policy makers, regulators, utilities, existing and startup energy companies, and other power-sector stakeholders to better understand the factors that are currently driving change in power systems worldwide. The report distills results and findings from more than two years of primary research, a review of the state of the art, and quantitative modeling and analysis.

This study does not attempt to predict the future. We follow the dictum of poet and author Antoine de Saint Éxupéry: "As for the future, your task is not to foresee, but to enable it." We identify key barriers and skewed incentives that presently impede the efficient evolution of the power sector and offer a framework for regulatory and market reform, based on a comprehensive system of efficient economic signals, that will enable an efficient outcome, regardless of how technologies or policy objectives develop in the future.

MITEI's Utility of the Future study was supported by a consortium of 23 diverse organizations from across the energy sector, and it is complemented by distinguished Advisory Committee and Faculty Committee. We gratefully acknowledge the support of the following consortium members at the Sponsor level: Booz Allen Hamilton, EDF, Enel, Engie, Gas Natural Fenosa, General Electric Corporation, Iberdrola, National Renewable Energy Laboratory, PJM, Saudi Aramco, Shell, US Department of Energy, and World Business Council for Sustainable Development. At the Participant level we wish to thank: The Charles Stark Draper Laboratory, Duke Energy, Enzen, Eversource, Lockheed Martin, NEC Corporation, PSE&G, Siemens, and Statoil. At the Observer level we wish to thank Paul and Matthew Mashikian. In addition to providing financial support, a number of our sponsors provided data that were helpful for our modeling activities. We are very grateful for this assistance.

Our Advisory Committee members dedicated a significant amount of their time to participate in meetings and to comment on our preliminary analysis, findings, and recommendations. We would especially like to acknowledge the efficient conduct of Advisory Committee meetings under the able and experienced direction of Chairman Philip R. Sharp and Vice Chairman Richard O'Neill.

This study was initiated and performed within MITEI. Professor Robert Armstrong has supported this study in his role as director of MITEI and as an active participant in the faculty committee. Louis Carranza, associate director of MITEI, structured the commercial model and worked closely with study executive director Raanan Miller in assembling the consortium members. MITEI staff provided administrative and financial management assistance to this project; we would particularly like to thank Emily Dahl, Jennifer Schlick, and Chelsey Meyer for communications support, and Debra Kedian and Jessica Smith for event support. Finally, we would like to thank Kathryn O'Neill and Marika Tatsutani for editing this document with great skill and patience, and Opus Design for layout and figure design.

This report represents the opinions and views of the researchers who are solely responsible for its content, including any errors. The Advisory Committee and the Study Consortium Members are not responsible for, and do not necessarily endorse, the findings and recommendations it contains.

This report is dedicated to the memory of our friend and colleague Stephen Connors.

Executive Summary

Important changes in the provision and consumption of electricity services are now underway, driven to a significant degree by a confluence of factors affecting the distribution side of power systems. A variety of emerging distributed technologies — including flexible demand, distributed generation, energy storage, and advanced power electronics and control devices — are creating new options for the provision and consumption of electricity services. At the same time, information and communications technologies are rapidly decreasing in cost and becoming ubiquitous, enabling more flexible and efficient consumption of electricity, improved visibility of network use, and enhanced control of power systems.

These technologies are being deployed amidst several broad drivers of change in power systems, including growth in the use of variable renewable energy sources such as wind and solar energy; efforts to decarbonize the energy system as part of global climate change mitigation efforts; and the increasing interconnectedness of electricity grids and other critical infrastructure, such as communications, transportation, and natural gas networks.

The MIT Energy Initiative's *Utility of the Future* study presents a framework for proactive regulatory, policy, and market reforms designed to enable the efficient evolution of power systems over the next decade and beyond. The goal is to facilitate the integration of all resources, be they distributed or centralized, that contribute to the efficient provision of electricity services and other public objectives. This framework includes a comprehensive and efficient system of market-determined prices and regulated charges for electricity services that reflect, as accurately as possible, the marginal or incremental cost of providing these services; improved incentives for distribution utilities that reward cost savings, performance improvements, and long-term innovation; reevaluation of the power sector's structure to minimize conflicts of interest; and recommendations for the improvement of wholesale electricity markets. This study also offers a set of insights about the roles of distributed energy resources, the value of the services these resources deliver, and the factors most likely to determine the portfolio of cost-effective resources, both centralized and distributed, in different power systems. We consider a diverse set of contexts and regulatory regimes, but focus mainly on North America and Europe.

This study does not try to forecast the future or predict which technologies will prevail. Instead, it identifies unnecessary barriers and distortionary incentives that presently impede the efficient evolution of the power sector and provides a framework that will enable an efficient outcome regardless of how technologies or policy objectives develop in the future. In addition, we recognize that regulatory and policy reform often proceeds incrementally and that each jurisdiction faces unique challenges and contexts. As such, we offer this framework along with guidance on the key trade-offs regulators and policy makers confront as they pursue opportunities for progressive improvements.

The measures identified in this study could produce significant cost savings. Low-cost information and communications technologies and advanced metering enable more cost-reflective prices and charges for electricity services that can finally animate the "demand side" of the power system and align myriad decisions with the optimization of net social welfare. Efficient prices and charges will unlock flexibility in electricity consumption and appropriately value the services that distributed energy resources provide. To date, power systems have been designed to meet infrequent peaks in demand and to comply with engineering safety margins established in an era when electricity customers were largely inflexible and blind to the true costs and potential benefits of their electricity consumption or production decisions. In many cases, this has resulted in costly and significantly underutilized infrastructure. Smarter consumption of electricity and, where cost-effective, the deployment of distributed energy resources, could deliver billions of dollars in savings by improving the utilization of electricity infrastructure.

At the same time, the need for proactive reform is clear. Customers now face unprecedented choice regarding how they get their power and how they manage their electricity consumption — regardless of whether they are aware of those choices or are acting on them today. New opportunities include the ability to invest in distributed generation, smart appliances, and energy efficiency improvements. At present, the vast majority of power systems lack a comprehensive system of efficient prices and regulated charges for electricity services. As a result, some customers are making inefficient investments and are overcompensated for the services that they provide to the power system. At the same time, many more opportunities that could deliver greater value are being left untapped because of inadequate compensation. For example, the combination of simple volumetric tariffs and net metering policies has contributed to the rapid adoption of rooftop solar photovoltaics (PV) in several jurisdictions, while exposing several flaws in current ratemaking. The rapid uptake of solar PV also demonstrates how quickly customers can react to economic signals — whether well or poorly designed — and the importance of proactive, rather than reactive, policy-making and regulation. In multiple jurisdictions, challenges that once seemed insignificant have quickly become overwhelming, and failure to act can catch policy makers and regulators flat-footed.

The framework proposed in this study is designed to establish a level playing field for the provision and consumption of electricity services, whether via centralized or distributed resources. The goal is to remove inefficient barriers to the integration of cost-effective new sources of electricity services, rethink ill-designed incentives for certain resources, and present a system of prices and charges that can animate efficient decisions. With this framework in place, all customers and producers of electricity services can make efficient choices based on accurate incentives that reflect the economic value of these services and their own diverse personal preferences.

This study highlights several core findings:

The only way to put all resources on a level playing field and achieve efficient operation and planning in the power system is to dramatically improve prices and regulated charges (i.e., tariffs or rates) for electricity services.

- To establish a level playing field for all resources, cost-reflective electricity prices and regulated charges should be based only on what is metered at the point of connection to the power system — that is, the profile of injections and withdrawals of electric power at a given time and place, rather than the specific devices behind the meter. In addition, cost-reflective prices and regulated charges should be symmetrical, with injection at a given time and place compensated at the same rate that is charged for withdrawal at the same time and place.
- Increasingly affordable information and communications technologies (e.g., advanced meters or interval meters) enable detailed monitoring of electricity withdrawals and injections and therefore facilitate more efficient prices and charges. Without more accurate consumption and injection data from all customers, it is impossible to capture the full value of electricity services.

- Flat, volumetric tariffs are no longer adequate for today's power systems and are already responsible for inefficient investment, consumption, and operational decisions.

- Peak-coincident capacity charges that reflect users' contributions to incremental network costs incurred to meet peak demand and injection, as well as scarcity-coincident generating capacity charges, can unlock flexible demand and distributed resources and enable significant cost savings.

- Granularity matters. The value or cost of electricity services can vary significantly at different times and at different locations in electricity networks. Progressively improving the temporal and locational granularity of prices and charges for these services can deliver increased social welfare. However, these benefits must be balanced against the costs, complexity, and potential equity concerns of implementation.

- Care must be taken to minimize distortions from charges that are designed to collect taxes, recover the costs of public policies (such as efficiency programs, heating assistance, subsidies for renewable energy, cross-subsidies between different categories of customers, etc.), and recover residual network costs (i.e., those network costs that are not recovered via cost-reflective charges).

- Policy makers and regulators must be wary of the possibility of societally inefficient "grid defection" if residual network costs and policy charges become too high. This may suggest an upper limit on the portion of these costs that can be collected in electricity tariffs rather than through broader taxes or other means.

The regulation of distribution utilities must be improved to enable the development of more efficient distribution utility business models.

- Forward-looking, multi-year revenue trajectories with profit-sharing mechanisms can reward distribution utilities for cost-saving investments and operations, aligning utilities' business incentives with the continual pursuit of novel solutions.

- Several "state of the art" regulatory tools, including an incentive-compatible menu of contracts, an engineering-based reference network model, and automatic adjustment factors to account for forecast errors, can better equip regulators for an evolving and uncertain electricity landscape.

- Outcome-based performance incentives can reward utilities for improvements in quality of service, such as enhanced resiliency, reduced distribution losses, and improved interconnection times.

- Incentives for longer-term innovation are needed to accelerate investment in applied R&D and demonstration projects and learning about the capabilities of novel technologies and practices that may have higher risk or longer-term payback periods.

The structure of the electricity industry should be carefully reevaluated to minimize potential conflicts of interest.

- Network providers, system operators, and market platforms constitute the critical functions that sit at the center of all transactions in electricity markets. Properly assigning responsibilities for these core functions is thus critical to an efficient, well-functioning electricity sector. It is also critical to establish a level playing field for the competitive provision of electricity services by traditional generators, network providers, and distributed energy resources.

- As experience with restructuring in the bulk power system has demonstrated, structural reform that establishes financial independence between distribution system operation and planning functions and competitive market activities would be preferable from the perspective of economic efficiency and would facilitate more light-handed regulation.

- If financial independence is not established, several additional measures are critical to prevent conflicts of interest and abuses of market power. These include stricter regulatory oversight of distribution network planning and operation, legal unbundling and functional restrictions on information exchange and coordination between distribution system operators and competitive subsidiaries, and transparent mechanisms for the provision of distribution system services (such as public tenders or auctions).

- Maintaining a data hub or data exchange may constitute a fourth critical function. Such a hub or exchange would serve several purposes: securely storing metered data on customer usage, telemetry data on network operation and constraints, and other relevant information; allowing non-discriminatory access to this data to registered market participants; and providing end customers with timely and useful access to data on their own usage of electricity services. Responsibility for this function should also be carefully assigned, with priority given to data security and customer privacy considerations.

Wholesale market design should be improved to better integrate distributed resources, reward greater flexibility, and create a level playing field for all technologies.

- Wholesale markets should enable transactions to be made closer to real time to reward flexible resources and to enable better forecasting and control of variable renewable resources and electricity demand.
- Wholesale market rules such as bidding formats should be updated to reflect the operational constraints of novel resources such as demand response and energy storage, as well as new patterns of operation of conventional power plants.
- More efficient pricing of reserves can help wholesale markets function better, improve price signals for energy and operating reserves, and strengthen the link between these two services.

Widespread connection of distributed energy resources, smart appliances, and more complex electricity markets increases the importance of cybersecurity and heightens privacy concerns.

- Robust regulatory standards for cybersecurity and privacy are needed for all components of an interconnected electricity network.
- To keep pace with rapidly evolving cybersecurity threats against large and complex electric power systems, electric utilities, vendors, law enforcement authorities, and governments should share current cyber threat information and solutions quickly and effectively.

Better utilization of existing assets and smarter energy consumption hold great potential for cost savings. At the same time, economies of scale still matter, and the distributed deployment of solar PV or energy storage is not cost-effective in all contexts and locations.

- The value of some electricity services can differ substantially depending on where within the power system that service is provided or consumed. This variation in "locational value" underscores the importance of locationally granular prices and charges and makes it impractical to define a single value for any distributed resource.
- Distributed energy resources can be sited and operated to provide services in those areas of the power system where their services are most valuable. Understanding the specific services that have locational value is thus critical to understanding how distributed resources can create value in power systems.
- Unlocking the contribution of resources that already exist — such as flexible demand, electric vehicles, power electronics, or distributed generation that is already deployed — can be an efficient alternative to investing in electricity generation and network capacity.
- Economies of scale still matter, even for distributed energy resources. For resources that can be deployed at multiple scales, such as solar PV and battery energy storage, incremental costs associated with failing to exhaust economies of unit scale can outweigh locational value. This can result in a "distributed opportunity cost," making distributed deployment of these resources inefficient. Trade-offs between the incremental costs and additional locational value associated with deploying distributed resources on a smaller scale must be considered in each context.
- For resources that exhibit significantly higher unit costs at smaller scales, such as solar PV and battery energy storage, distributed deployment is likely to be inefficient in many locations. Exceptions may include areas that have heavily congested networks or that are experiencing rapid growth in electricity demand. In these areas, locational value may be significant.
- New innovations may transform economies of unit scale for solar energy or storage technologies, enabling more ubiquitous distributed deployment of these resources.

PART 1: UNDERSTANDING ELECTRICITY SERVICES AND HOW DISTRIBUTED ENERGY RESOURCES AFFECT THE DESIGN AND OPERATION OF POWER SYSTEMS

01

A Power Sector in Transition

1.1 The Facts of Today

Electric power systems in the United States, Europe, and several other parts of the world are experiencing an unprecedented set of changes driven by the intersection of several key trends: the increasing decentralization of power systems, epitomized by the growing penetration of distributed generation (and more recently, energy storage) and more active and price-responsive energy "consumers"[1]; a proliferation of information and communications technologies (ICTs) that enable energy to be produced, transmitted, and consumed more intelligently and efficiently by agents of any size; the growth of variable renewable energy sources such as wind and solar energy; the decarbonization of the energy system as part of global climate change mitigation efforts; and the increased interconnectedness of electricity with other critical infrastructure — such as communications and transportation — which enhances the importance of electricity in modern economies. These changes all give rise to a central question: How will the electricity services that are today primarily provided in a centralized, top-down manner be provided in the future?

The Utility of the Future study focuses on a central question: How will the electricity services that are today primarily provided in a centralized, top-down manner be provided in the future?

We begin by highlighting the changes that are occurring in the power sector, regardless of their drivers. However, as we will highlight throughout the report (including later in this chapter), subsidies, regulations, market designs, power sector structures, and technological and business model innovations have all played key roles in driving change in the power system. Later chapters review the efficacy and cost-effectiveness of existing and proposed implementations of many of these policies, regulations, and market designs and propose efficient upgrades to electricity regulation and markets where needed.

1 As distributed resources are adopted, many consumers are becoming what some authors call "prosumers" — that is, they produce energy at some times and consume it at others. Rather than distinguishing among producers, consumers, and prosumers, this report will refer simply to network users or agents.

This study defines distributed energy resources, or DERs, as any resource capable of providing electricity services that is located in the distribution system.

1.1.1 The power system is becoming more distributed

Power systems around the world are becoming less centralized as the resource mix integrates distributed energy resources (DERs)[2] and new options for providing and consuming electricity services emerge in the distribution system. In most power systems, DERs remain minor players in the provision of electricity services; nonetheless, smart energy consumption and DER deployment are generally on the rise.

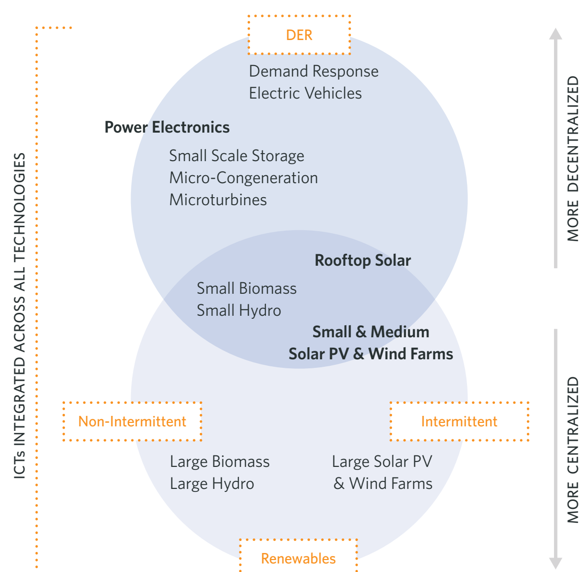
There is an abundant and rapidly growing body of literature on the actual or potential transformation that is taking or may take place in the distribution network, and the term "distributed energy resources" is profusely employed, frequently with different meanings. In this study, a distributed energy resource or DER is defined as any resource capable of providing electricity services that is located in the distribution system. DERs include demand response, generation, energy storage, and energy control devices, if they are located and function at the distribution level. DERs can be understood even without a precise definition of electricity services, which is provided in **Chapter 2**.

Some DERs, including electric vehicles, air conditioners, refrigerators, or a building's thermal storage capacity, exist primarily for reasons other than to provide electricity services. Other DERs, such as solar photovoltaic (PV) panels or electric batteries, are installed specifically to provide such services. Some of these resources, such as electric batteries or solar PV panels, can be deployed at all voltage levels — including large-scale installations at the bulk power level — while others, such as electric vehicles or refrigerators, are intrinsically distributed.

This study examines the evolution of the provision of electricity services, and it is not constrained to the services provided by DERs. Centralized and distributed resources alike can provide these services. What is new is the confluence of: (1) a growing presence of renewable generation, a significant fraction of which is distributed;[3] (2) increasing integration of DERs in the power system, although still modest in most countries; and (3) the proliferation of ICTs, which makes it possible for DERs and flexible demand to participate in the functioning of the power system. The issues related to the integration of renewables at large scale at the level of the bulk power system have been extensively analyzed elsewhere, although much remains to be done (for example, **Chapter 7** discusses market design changes needed to integrate increasing levels of intermittent renewables).

DERs have, in many cases, been conflated with renewable energy resources. **Figure 1.1** illustrates the overlap and the differences between distributed resources and renewable resources. Many renewable resources can, of course, be deployed in both a distributed or centralized form. This study focuses on the potential role of DERs in an evolving power system. We consider DERs that are not renewable, such as micro cogeneration or small, gas-fired, backup turbines. Despite our focus on DERs, we also consider generation that is not distributed, such as a large wind farm connected in high voltage.

Figure 1.1: Illustrative Taxonomy of Distributed and Renewable Energy Resources



2 "Decentralized," "dispersed," and "embedded generation or resources" are also terms commonly used to refer to DERs.

3 For example, in 2014, roughly 29 percent of new solar energy capacity additions in the United States were distributed, constituting roughly 8 percent of all US generating capacity additions. In 2015, 41 percent of new solar energy capacity additions in the United States were distributed, constituting roughly 11 percent of all US generating capacity additions (EIA 2016a).

Germany is perhaps the most striking example of the growing impact of DERs on power systems. There, 98 percent of all solar PV — representing nearly 40 gigawatts of installed capacity — is connected to low- and medium-voltage distribution grids, and 85 percent of German solar capacity is produced by PV installations smaller than 1 megawatt in capacity (Fraunhofer ISE 2016). But Germany is not alone. Nearly one in five customers in the US state of Hawaii and one in 10 single-family homes in California now has a rooftop solar PV system (CSI n.d.; Trabish 2016a). To date, these states are largely outliers in the United States, where only about 1 percent of homes have rooftop solar, although this fact too is changing. In 2015, distributed solar PV alone accounted for nearly 11 percent of all new power generation capacity in the United States (EIA 2016a).

Distributed solar PV is not the only distributed resource of interest. According to one market research firm, combined heat and power (CHP) units and fuel cells accounted for 8 percent of all US generation capacity in 2015, and this capacity is expected to grow (Munsell 2016a). Furthermore, 75 percent of backup generation capacity[4] in the United States is fueled by diesel or natural gas (The Brookings Institution 2011). Onshore wind farms are often connected at distribution voltage levels as well, and they have reached significant capacity in many power systems (Global Wind Energy Council 2016).

Meanwhile, thermal energy storage, lithium-ion batteries, and other energy storage resources (such as flow batteries) are becoming more competitive. These storage systems are increasingly being deployed in a distributed fashion. To date, battery energy storage projects in the United States are largely located within distribution systems or customer premises and average roughly 2.8 megawatts (MW) in capacity.[5] Heating, ventilation, and air-conditioning (HVAC) systems; water heaters; and batteries now collectively account for over 80 percent of the demand resources providing regulation reserves for PJM, a US regional transmission organization, in its market (McAnany 2016). And, while the total market size for energy storage resources remains small and localized, the US market for non-pumped hydro energy storage grew by more than 240 percent in 2015 (Munsell 2016b).

1.1.2 The power system is becoming increasingly digitalized, enabling more active and price-responsive demand

The digitalization of the power system through the deployment of information and communications technologies (ICTs) is proceeding in parallel with — and in part enabling — the decentralization of electricity resources. This digitalization is making it easier to compute and communicate the value of electricity services with finer temporal and spatial granularity. This is, in turn, enabling energy demand to become increasingly responsive to changes in the prices of these services and participate actively in their provision. Digitalization in combination with the new energy resources introduced above is enabling networks to become more actively managed, potentially ending the passive network management paradigm, in which networks are sized to meet the aggregate peak demand of passive consumers.

The late MIT Professor Fred Schweppe, in his seminal 1978 essay "Power Systems 2000," imagined a world in which demand actively participates in the provision of critical electricity services (Schweppe 1978). That vision is fast becoming a realistic possibility today. In the PJM[6] market in the eastern United States, nearly 11 gigawatts (GW) of demand-side resources cleared in the capacity market for delivery in 2019 and 2020. In addition, demand-side resources bid an average of nearly 1.5 GW of capacity daily in the synchronized reserve market (Monitoring Analytics LLC 2016). In all, flexible demand-side resources in PJM earned roughly $825 million in revenues from participating in PJM's various markets in 2015 (McAnany 2016). While PJM has proven to be a leader in activating demand in energy and

4 By "backup generation capacity" we mean generators installed at customer sites with the purpose of ensuring continuous and reliable power supply during times of peak demand, peak pricing, and/or during periods of bulk system blackout.

5 As of September 23, 2016, the Department of Energy (DOE) Global Energy Storage Database contained 471 electrochemical energy storage projects in the United States, accounting for 1,311 MW of capacity. The majority of these projects (266/471) have rated capacities of less than 250 kilowatts (US DOE and Sandia National Laboratories n.d.).

6 PJM is a regional transmission organization that operates electricity markets, coordinates bulk system dispatch, and operates the transmission infrastructure in all or parts of the US states of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, and West Virginia plus the District of Columbia.

capacity markets, it is not alone. Demand resources provide more than 1 GW of capacity in the NYISO[7] market (Kasera 2016).

The increasing digitalization of the power sector through the deployment of ICTs is also embodied in the rollout of advanced metering infrastructure and other network-sensing infrastructure in the United States and Europe. In the United States, roughly 59 million smart meters have been deployed, covering over 40 percent of metered sites (EIA 2016b). In the European Union, advanced meter deployments are expected to reach 72 percent of consumers by 2020 (European Commission 2016).

The increased digitalization of the power sector is enabling countless new opportunities. Dozens of businesses — ranging from nimble new ventures to powerful incumbents — are offering increased monitoring and control of power networks. As ICTs proliferate throughout electricity networks, regulatory agencies, utilities, and aggregators are increasingly utilizing new tariff structures — indeed, in the second quarter of 2016, 42 of 50 US states took action to change tariffs or address DERs in some manner (North Carolina Clean Energy Technology Center 2016). These new tariff structures range from technology-specific demand charges applied to those with distributed resources to time-of-use rates and three-part tariffs with real-time energy prices, demand charges, and fixed charges.

Beyond tariffs, increased network sensing and digitalization is also taking active network management from theory to practice. For example, in the United Kingdom, UK Power Networks now signs “flexible” interconnection agreements with wind developers, speeding up the time to interconnection by actively managing the output of wind farms and curtailing generation when necessary to manage network constraints (Anaya and Pollitt 2015). In the United States, entrepreneurs are developing software and hardware that enables network operators to control the impedance or topology of their networks in real time to optimize power flows,[8] and utilities are piloting programs to aggregate dispatchable demand response and distributed generation to defer or avoid conventional network upgrades.[9]

However, the increased digitalization of the power system has created new vulnerabilities. Protecting a nation’s electricity grid from cyber attacks is a critical national security issue and an important priority for electric utilities (Campbell 2015; Bipartisan Policy Center 2014). As the December 2015 cyber attack on the Ukrainian power grid demonstrated (Electricity-ISAC and SANS Industrial Control Systems 2016), electric utilities are vulnerable to attack and will become more so in the next decade as utility systems employ more digital controls and as operations and metering and resource-management systems become more interconnected and complex. The widespread connection of solar, wind, demand response, and other distributed energy resources with two-way digital controls increases cyber vulnerabilities and requires more widespread and intensive cybersecurity protection. Utilities throughout the world are therefore focusing on resilience and preparation to contain and minimize the consequences of cyber incidents. The increasingly widespread collection and, in certain markets, dissemination of energy production and consumption data is already causing privacy concerns and raising questions about who should own and manage this data.

1.1.3 The resource mix is becoming more renewable and intermittent

The growth of distributed resources is taking place against the backdrop of a transition to a more renewable and intermittent resource mix. Worldwide, the electricity resource mix is being transformed by the growth of renewable energy resources such as onshore and offshore wind, solar PV and concentrated solar power, biomass, small and large hydro, geothermal, and marine energy (REN21 2016). Five headline-catching events from within the span of a few months in 2016 alone

7 NYISO is a regional transmission organization that operates electricity markets, coordinates bulk system dispatch, and operates the transmission infrastructure in the US state of New York.

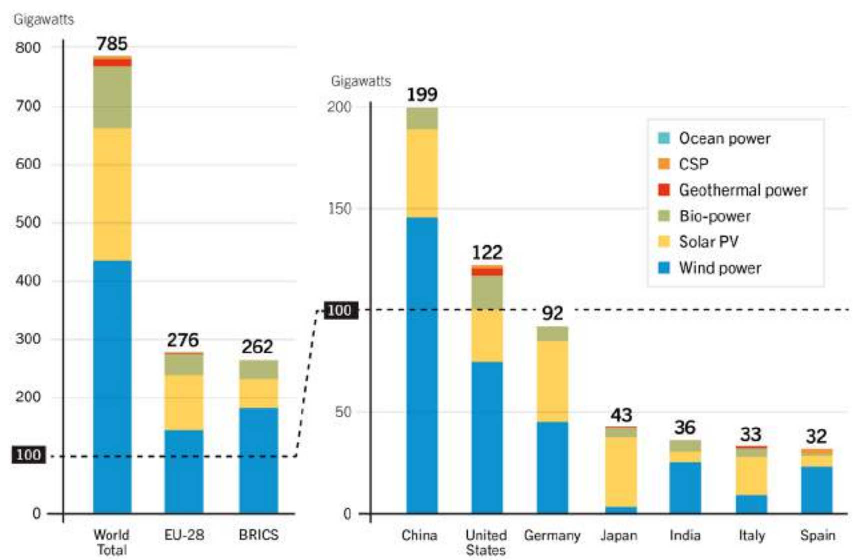
8 See Smart Wires Inc. (www.smartwires.com) for an example of an impedance control technology and NewGrid (newgridinc.com) for an example of a topology control technology.

9 See the Brooklyn-Queens Demand Management program in the US state of New York for an example.

highlight the significant impact of renewable resources on power systems. (1) From May 7 to May 11, Portugal met its electricity demand entirely with renewable energy sources. (2) Costa Rica generated 100 percent of its electricity from renewable sources, including hydropower, geothermal, and wind energy, for 76 straight days beginning June 17 (ICE 2016). (3) Renewable power sources (including large hydropower plants) provided 40 percent of electricity demand in California in the United States on May. (4) For a brief period on May 15, renewable resources met nearly 100 percent of electricity demand in Germany, Europe's largest economy. (5) Finally, solar PV plants produced more electricity for one week in May 2016 than coal-fired power plants in the United Kingdom, one of the birthplaces of coal-fired electricity production.

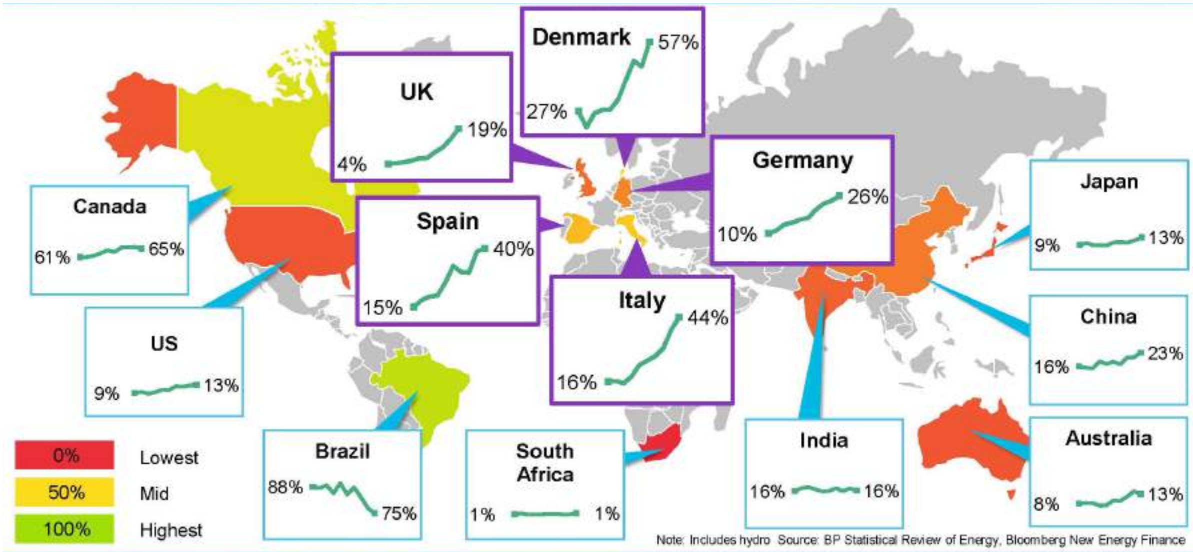
While these short-lived events drew headlines, the broad growth of renewable energy in global electricity markets is no anomaly. Globally, renewable energy resources added 213 terawatt-hours (TWh) in 2015, an amount roughly equal to total 2015 electricity demand growth (BP 2016). Also, 2015 was a record year for investment in renewable energy: In total, nearly $286 billion was invested globally to deploy roughly 134 GW of renewable energy resources, excluding large hydro, representing nearly 54 percent of all new power capacity (Frankfurt School-UNEP Centre and BNEF 2016). Looking forward, the world's largest electricity markets, including the United States, European Union, China, India, Brazil, and Mexico are all planning to expand renewable energy generation significantly in the coming decade.

Figure 1.2: Global Renewable Energy Deployments, Year End 2015



Source: REN21 2016

Figure 1.3: Renewable Energy's Share of Power Generation, 2004–2014



Source: Liebreich 2016. Reprinted with permission from Bloomberg New Energy Finance (BNEF); figure from a presentation given at BNEF Summit: New York, April 5, 2016.

This global shift to renewable resources is not without unintended consequences and growing pains. Power prices in Germany have reached as low as negative 320 euros per megawatt-hour (MWh) and frequently remain below zero for hours at a time.[10] The share prices of E.ON, RWE, and EnBW — three of Germany's largest utilities — have collapsed by 45 percent to 66 percent over the five years preceding the writing of this report, even as the DAX index of German stocks grew steadily. Collectively, Western European utilities have lost hundreds of billions of dollars of market value in the past decade. Further, in 2014, German distribution and transmission system operators curtailed nearly 1.6 TWh of renewable electricity — a 200 percent increase over 2013, and in 2015, curtailment rose a further 69 percent to 2.7 TWh. Many US states are experiencing similar challenges with negative electricity prices and curtailment; certain renewable generators in the Electric Reliability Council of Texas faced negative prices in nearly 18 percent of generating hours in 2011, and many California Independent System Operator units faced negative prices for nearly 6 percent of generating hours that same year[11] (Schmalensee 2016). Indeed, California's market operator, facing unpredicted net demand[12] ramps associated with the penetration of solar PV, is introducing new market products — such as the Flexible Ramping product — to encourage more flexible resources to enter the market (California ISO 2016). Challenges remain in integrating greater shares of renewable energy, but the trend is clear: Renewable energy resources are no longer a niche resource in many global power systems, but rather one of the largest sources of new generating capacity.

1.1.4 Power systems are decarbonizing

The growth of renewable energy is occurring in parallel with (and in part due to) a mounting focus on decarbonizing electric power systems. Climate change presents an urgent global challenge, and in December 2015, 195 nations came together to negotiate the Paris Agreement, which commits the world's nations to limit global average temperature increases to less than 2 °C above preindustrial levels. Almost every nation in

10 This phenomenon will be discussed in more detail in **Chapter 7**. Negative prices can emerge for a number of reasons, including production subsidies and operational constraints for certain power plants.

11 The California Independent System Operator (CAISO) does not make reliability-based renewable energy curtailments publicly available. What data are available indicate that forced curtailments are increasing in CAISO.

12 Net demand here refers to electricity demand minus any intermittent renewable generation.

the world is now focused on reducing greenhouse gas emissions, and the power sector is frequently the linchpin of climate mitigation efforts (DDPP 2015). In August 2015, the US Environmental Protection Agency released Clean Power Plan regulations, targeting a 32 percent reduction in power sector carbon emissions relative to a 2005 baseline by 2030 (EPA 2015). The European Union member states are collectively committed to cut greenhouse gas emissions 20 percent below 1990 levels by 2020 and 40 percent by 2030 (European Commission 2014). Even rapidly growing middle-income countries like Mexico and China have pledged to peak and then reduce their total emissions in the coming decade. The multifaceted transition under way in the world's electricity sectors thus includes the drive toward a lower-carbon energy supply.

1.1.5 Power systems are becoming increasingly integrated with other key sectors and critical infrastructure

The final trend is the increased interconnectedness and interdependence of electricity and other key sectors and critical infrastructure, such as communications, natural gas, heat, and transportation. Very few industries would function without the steady supply of electricity, making reliable, secure, and affordable electric power systems a cornerstone of modern economies. As the US Department of Homeland Security notes, the energy sector — and electricity in particular — is "uniquely critical because it provides an 'enabling function' across all other critical infrastructure" (US DHS n.d.).

Natural gas is gaining prominence in the energy mix of many countries. Some view natural gas as a transition fuel to reduce CO_2 emissions and help integrate intermittent renewables. The installed capacity of natural gas combined-cycle power plants has grown dramatically in recent decades in both the United States and Europe, and the power sector has become a more substantial consumer of natural gas. Gas consumption for electricity generation in the United States has grown from 4 percent to 33 percent of total electricity consumption during the last 25 years (EIA 2016a). Indeed, for the first time in US history, natural gas generated more electricity than coal in 2016. Moreover, gas-based DERs, including fuel cells, have become promising technologies to provide both electricity and heat at the end-consumer level. According to the US Department of Energy, almost 25 percent of Fortune 100 companies now use fuel cells to generate clean, efficient, and reliable power to power data centers, cell phone towers, corporate buildings, retail facilities, or forklifts (US DOE 2015b). The growth of electricity generation from natural gas at both the power plant and end-user level is increasing the dependence of power systems on gas infrastructure. Nevertheless, dependencies run in the other direction as well. Compressor stations, key valves, and regulators are critical elements that consume electricity to transport gas, and some gas-based DERs (e.g., internal combustion engines) also need electricity to start up. As noted by the US Federal Energy Regulatory Commission and North American Electricity Reliability Corporation after a particularly severe cold weather event in February 2011, the reduction in natural gas supplies contributed to the loss of electric service, and the loss of electric service in turn contributed to natural gas curtailments (Schatzki and Hibbard 2013). The electricity and natural gas sectors of many countries are thus fundamentally interdependent, and they will become even more so over time.

In contrast, transportation has historically been one of the few key economic sectors not directly dependent on electric power, but that too is beginning to change. Electric vehicles (EVs) represent an important new class of electricity users — a growing segment of electricity demand that is mobile on timescales not previously seen in the electric power sector. This segment can act as both a significant, flexible, or schedulable demand and potentially even a flexible and distributed supply resource. EV penetration is low today but rising in the majority of markets. Globally, 2015 EV demand increased by more than 80 percent over 2014. EVs could reach 20 percent of new vehicle sales worldwide by 2030 and 35 percent by 2040 (BNEF 2016). Extrapolating that average adoption rate to the US market, roughly 16 million EVs could be traveling US roads by 2030, with on the order of 1,000 GWh of battery storage capacity. While forecasts vary widely, nearly all point toward an increasing penetration of electric vehicles. Given that EV adoption is likely to be concentrated in some markets, average adoption numbers may underestimate the technology's impact — indeed, in Norway, nearly 25 percent of all new light-duty vehicle registrations in 2015 were EVs (Jolly 2015).

1.2 Implications for the Power Systems of Tomorrow?

The facts of today cannot be disputed. The penetration of renewable energy — and more recently, energy storage — has increased over the past decade in nearly every major economy on the planet. Further, these resources are increasingly being deployed in a distributed fashion, disrupting the traditional "top-down" structure of the power sector. A diversity of new agents — including those traditionally described solely as "demand," aggregators, and other new energy solutions providers — are now supplying services to energy, capacity, and ancillary services markets, enabled by the deployment of more affordable and ubiquitous ICTs. These same ICTs are aiding and enabling regulatory reform and increasingly active network management, while creating new concerns about the security of the power system.

Nonetheless, intermittent renewables, storage, and active demand represent only a small fraction of the global resource mix. Indeed, despite years of impressive growth, excluding large hydroelectric resources, renewables still meet only 2.8 percent of global energy demand (BP 2016) and 6.7 percent of electricity production. If the system were to stop changing today, the facts outlined above could potentially be ignored as marginal. Nonetheless, many academics, industry analysts, utilities, new ventures, and other system stakeholders believe that these facts signal the beginning of a more substantial upheaval of the power sector.

For example, Bloomberg New Energy Finance (BNEF) predicts that global fossil-fuel use for electric power generation will peak in 2025 — less than one decade away — and decline inexorably thereafter (BNEF 2016). Renewable resources and energy storage are the culprits in this expected decline. Lest we forget about the changing nature of demand, BNEF predicts that electric vehicles could account for as much as 25 percent of new light-duty vehicle sales in the United States by 2030 (Liebreich 2016). Navigant Research predicts that by the end of the current decade, more decentralized power generation capacity will be added than centralized power generation capacity (Trabish 2016b). As we will discuss in the remainder of this report, if these trends materialize, they will have dramatic effects on the structure and operations of the power sector.

This opinion is not shared only by industry analysts. Audrey Zibelman, chair of the New York State Department of Public Service — the electricity regulatory body for the United States' third most populous state — recently wrote: "Rooftop solar, energy storage (from household batteries to electric vehicles), smart energy management technology, and the aggregation of demand are all areas where demand, rather than generation, can become the state's primary energy resource" (Zibelman 2016). In one of the documents of the "Energy Union Package" defining the European Union's electricity strategy over the coming decade, the European Commission states that, to accomplish the European Union's energy goals, the European Union must "move away from an economy driven by fossil fuels, an economy where energy is based on a centralised, supply-side approach and which relies on old technologies and outdated business models" (European Commission 2015). These are just samples of the key policy makers and policy-making bodies expecting a more renewable and distributed future — a future in which demand-side resources and energy "consumers" play an active role in the investments and operations of the power sector. Similar statements of the potential promise of a more decentralized power system can be found in recent documents by the US National Renewable Energy Laboratory (Zinamen et al. 2015), the UK Office for Gas and Electricity Markets (OFGEM 2015), and the EU Agency for the Cooperation of Energy Regulators (ACER 2014).

The previously highlighted financial strains on some incumbent utilities, when combined with the changing dynamics of the power sector, have caused many in the power system to re-examine the basic economic functions of the agents involved. Theodore Craver Jr., the CEO of Edison International, one of the United States' largest utility holding companies, notes, "It would be foolish to dismiss the potential for major changes in the utility business model." Craver is not alone. A 2015 survey of global electric power sector executives found that 97 percent expect medium or high levels of disruption in their market segments by 2020, with 86 percent of executives in North America and 91 percent in Europe projecting major changes to the market model in which they operate by 2030 (PricewaterhouseCoopers 2015). One of the defining strategy documents produced by German utility company RWE claims that "conventional power generation, quite frankly, as a business unit, is fighting for its economic survival." It is unsurprising then, that Eurelectric, the trade group for European electric utilities, predicts a 6 billion-euro decline in wholesale generation value between 2013 and 2020, and a 10 billion euro increase in downstream market opportunities — including DERs, services, and power flow optimization in Europe alone (Eurelectric 2013). These trends have driven the restructuring of many major incumbent utilities, spurred the launch of many new ventures, and led to mergers, acquisitions, and initial public offerings.

Some envision that the changes seen today will continue unchecked. They envision a future in which centralized resources, and perhaps even transmission and major portions of distribution, will become relics of the past — fossils of the fossil fuel age. In this future, consumers meet the majority of their energy needs by producing on site, transacting for any remaining energy needs with their nearest neighbors; many defect from the grid entirely; the power system is completely upended; and the roles of energy and network suppliers, power system operators, and regulators are wholly redefined.

Others claim that distributed resources are a blip on the energy radar, their adoption driven by subsidies and regulatory and market flaws that shift costs to less fortunate network users. In this view, the lower carbon power systems of the future will be structured much as they are today, but will be reliant primarily on large-scale renewables, nuclear energy, and/or fossil fuels with carbon capture and storage.

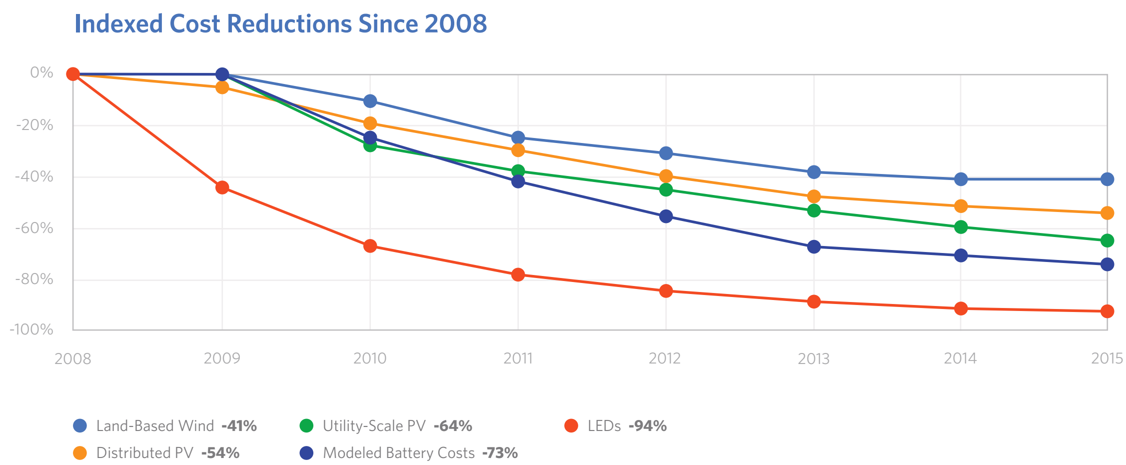
These viewpoints represent two ends of the spectrum, with a continuum of possible futures in the middle. Predicting the pace of eventual change is difficult.

1.3 Drivers of Change

Indeed, many of the technologies and ideas driving these changes aren't necessarily new. Referring again to Fred Schweppe's "Power Systems 2000" essay, we see that even in the late 1970s, some scholars believed that demand-side resources could play a central role in balancing supply and demand and operating power systems in the future. The potential for the emergence of distributed generation (at that time, primarily CHP) was one of the primary motivators of the 1970 Public Utilities Regulatory Policies Act in the United States, and some analysts at the time heralded the imminent arrival of a more distributed and renewable future (Lovins 1976) — a future that only four decades later may now be a realistic possibility. A degree of healthy scepticism about pronouncements of sweeping change would thus not be unfounded.

At the same time, three converging drivers are accelerating the rate of deployment of distributed and renewable resources today.

Figure 1.4: Cost Declines in Key Technologies, 2008–2014



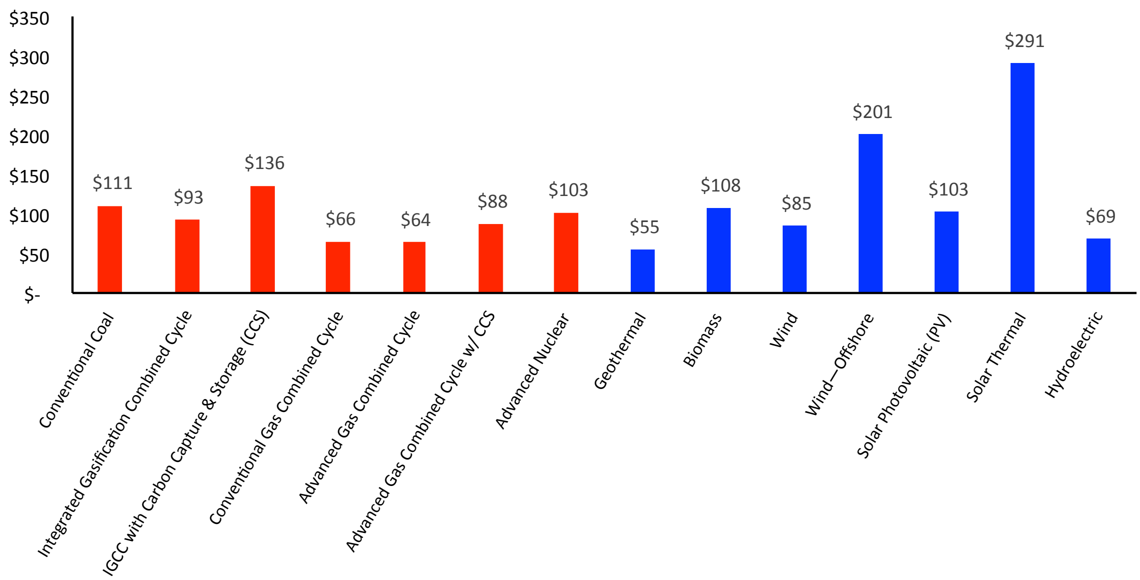
Source: US DOE 2016

First, technological innovation has driven dramatic cost declines in a number of technologies. The cost of wind and solar PV — the two leading non-hydro renewable energy technologies globally — have decreased by 40 percent and 60 percent respectively just between 2008 and 2014 (see **Figure 1.4**). Many solar PV developers and industry analysts expect the installed cost of utility-scale solar PV to fall below $1 per watt before the end of this decade (Wesoff 2015), and experts foresee a further 24 percent to 30 percent reduction in wind energy costs by 2030 (Wiser et al. 2016). Even more impressive is the cost reduction in light-emitting diodes (LEDs), which have plummeted roughly 90 percent since 2008. These cost declines are enabling the widespread adoption of LEDs, a solution that may dramatically reduce the 20 percent of electricity consumed in lighting. LEDs are but one example of the improvements in the technologies that enable homes, businesses, and industries to consume energy more flexibly and efficiently. Energy storage technologies are also progressing at a rapid rate. Lithium-ion batteries, now the common technical basis for both electric vehicle batteries and stationary energy storage, have reached gigawatt-scale markets, driving approximately 14 percent annual declines in battery costs between 2007 and 2014 (Nykvist and Nilsson 2015). One major US automaker projects that lithium-ion battery cell costs will drop below $100 per kilowatt-hour by 2022 — an order of magnitude less costly than 2010 costs (Wesoff 2016).

These recent innovations do not necessarily mean that these technologies are or will be ubiquitously cost-competitive today or in the near future. Many comparative levelized cost of energy[13] analyses exist and frequently highlight that, despite recent innovations, DER technologies still require subsidies or other support to compete with incumbent resources in many markets (see **Figure 1.5** for an example). We discuss the relative economics of various DERs and centralized resources in **Chapter 2** and **Chapter 8**.

13 Levelized cost of energy is a metric that divides the total capital and operational costs of a given energy resource by the cumulative energy output of that resource over its lifetime. This metric is useful for simple comparisons of relative technology costs, but it provides an incomplete picture when comparing technologies with very different utilization patterns in a power system — comparing dispatchable technologies with intermittent renewables, for instance.

Figure 1.5: US National Average Levelized Costs of Electricity for New Generation Entering Service in 2022*



Source: National Academies of Sciences, Engineering, and Medicine 2016, Appendix B.

* Assumes $15 per ton of carbon dioxide emitted and adjusts for intermittency, the potential need for transmission investments, and the cost of criteria air pollutants.

Second, policies — particularly those related to the deployment of renewable energy technologies and decarbonizing the power sector — have created favorable investment environments for many of these emerging technologies, building a positive feedback loop with technological innovation that has, to date, resulted in continued cost declines. Nearly every developed economy, and many developing economies, have created policy mechanisms to encourage the deployment of renewable generation and increasingly storage as well. In 2014, subsidies for the deployment of renewable energy technologies amounted to $112 billion globally (the consumption of fossil fuels is also subsidized — to the tune of $490 billion globally) (IEA 2015). In 2013 in the United States, renewable energy resources received 51 percent of all energy-specific federal subsidies, and wind and solar collectively accounted for 64 percent of federal electricity production subsidies (US DOE 2015a). These policies materialize in specific regulations or market rules, including direct support mechanisms such as feed-in tariffs, auctions, or utility purchase obligations (e.g., renewable portfolio standards), as well as indirect support schemes such as priority in the dispatch, omission from certain market-bidding procedures, or waivers of balancing costs. In other cases, regulations may fail to capture the increasingly complex realities of today's power sector, providing unintended incentives for DER adoption. Most notably, the recovery of fixed network costs through flat, volumetric tariffs, together with the practice of net metering[14] of demand and behind-the-meter generation, results in skewed incentives for network users with embedded generation.

Consumer choice and preference is a third and final driver of change. DERs may bring an unprecedented level of choice to agents that were formerly passive consumers

14 "Volumetric tariffs" refers to the practice of pricing electricity services on a dollar per kilowatt-hour basis. "Net metering" refers to the netting of energy supplied locally with energy supplied via the bulk power system. Net metering is typically performed over a specific time period; that is, all of the energy produced locally over, say, a one-month period is netted with all energy supplied via the power system in that month.

Failing to create a level playing field for DERs and centralized resources will result in significant costs borne by electricity consumers.

of electricity. Increasingly, electricity customers will be able to express their preferences and values through their decisions about the consumption and provision of electricity services (e.g., minimize cost, reduce environmental impact, take control of energy production, express dissatisfaction with the incumbent provider, project status, etc.). In this study, we do not attempt to perform a detailed analysis of the personal preferences or values of consumers and instead focus on outlining a set of regulatory reforms and efficient prices and charges for electricity services that will enable this expression of preference in an efficient power system. However, the ability of consumers to choose not only from whom they procure their electricity services but also from what resources and which new services beyond electricity they might desire, may be another important driver of change with significant implications for future electricity systems.

1.4 The Focus of This Study

This study focuses primarily on the potential for increased decentralization of power systems, with an emphasis on European and North American systems and, by extension, many other power systems with similar designs. This study does not attempt to predict what the future of utilities will be, which technologies or resources will eventually dominate, or the final outcome of the changes that power systems are now experiencing. We do not begin with the assumption that power systems of the future will be dominated by DERs or by centralized resources. Instead, our recommendations attempt to create a level playing field upon which DERs can efficiently and fairly compete with networks and centralized resources to provide electricity services. As we will demonstrate, failing to create a level playing field for DERs and centralized resources will result in significant costs borne by electricity consumers.

This approach aims to "future-proof" power systems such that the system's fundamental needs can be met efficiently however technologies or policy objectives may evolve in the future.

This study offers a framework for proactive regulatory reform, including improvements to the pricing of electricity services, incentives for distribution utilities, and recommendations for power sector structure and electricity market design. This framework is intended to be robust to the uncertain changes now under way and capable of facilitating the emergence of an efficient portfolio of resources — both distributed and centralized — to meet the needs of a rapidly evolving electricity sector.

The objectives of this study are fourfold:

1. To introduce how DERs may affect the design and operation of power systems and create new opportunities for the provision of electricity services (**Chapters 2–3**).
2. To present a framework for proactive regulatory reforms — including a comprehensive system of prices and charges for electricity services, upgraded regulation and incentives for distribution utilities, restructuring of the power sector organization, and improvements to the design of wholesale and ancillary service markets — that collectively remove inefficient barriers and skewed incentives that are impeding the evolution of the power sector (**Chapters 4–7**).

3. To offer — assuming the right reforms, prices, and charges are implemented — a set of insights about the value of novel technologies and an evaluation of what factors are likely to determine the portfolio of cost-effective solutions, both distributed and centralized, that may emerge in different power system contexts (**Chapter 8**).

4. Finally, to collect the regulatory recommendations contained in the study in the format of a toolkit for regulators and policy makers (**Chapter 9**).

Each of our recommendations stems from the fundamentals of power system economics and regulation. Each recommendation is accompanied by quantitative analysis and techno-economic modeling, where possible. In each case where regulatory constraints or factors other than economic efficiency guide decision making, we attempt to provide the rationale for and costs of deviating from the economically efficient frontier.

The suite of models used directly in the production of the results in this study is shown in **Figure 1.6**. These models range in detail and scope — with some representing large swaths of the power sector with appropriate levels of abstraction and others representing smaller portions of the power system with greater detail. A more detailed description of these models can be found in **Appendix A**.

This study focuses on technologies and recommendations that may prove impactful in the 2025–2030 time frame. It also focuses on the power sector but has clear implications for the gas, building, and transportation sectors as well, as these are increasingly intertwined with the electricity sector.

1.5 A Guide to Reading This Report

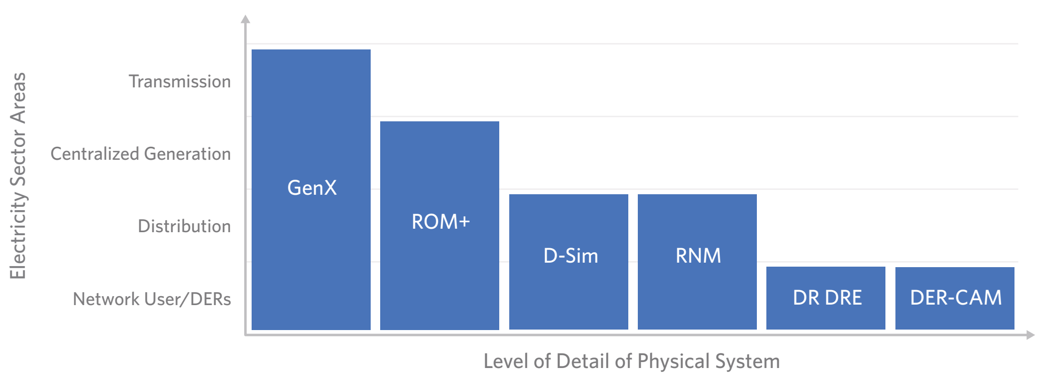
The following sections briefly outline the report, offering a guide to readers to the content ahead.

Part 1: Understanding electricity services and how distributed energy resources affect the design and operation of power systems

To understand the full potential impact of novel resources, including DERs, we must first understand the range of electricity services and the many ways in which DERs may provide these services. **Chapter 2** defines electricity services and introduces the distinction between locational services — those for which value depends significantly on the location of service provision within the power system — and non-locational services — more fungible services that are equally valuable across wide geographic areas. The chapter also discusses the range of new options for the provision and consumption of electricity services, focusing centrally on understanding the distinguishing features of DERs.

DERs, given their novelty, have attracted significant attention from analysts and academics in recent years. They have been lauded as potentially revolutionary in their ability to drive change in power systems operations and planning. Outlining possible visions of the future of power systems can be useful in galvanizing industry stakeholders to explore and create better outcomes.

Figure 1.6: Suite of Models Deployed in MIT Utility of the Future Study



Chapter 3 follows in the footsteps of prior attempts to describe the potential impact of DERs, highlighting some of the ways that DERs may dramatically change the power system, providing examples of how the groundwork for the future of power systems is being laid today, and painting a plausible sketch of a futuristic vision of the power system some years from now.

Part 2: A framework for an efficient and evolving power system

DERs have many characteristics that are fundamentally different from the centralized fossil, hydro, and nuclear resources that have historically dominated the power sector — primarily, that they are often co-located with load, sited in distribution systems, owned or controlled by a multitude of disparate actors, and exist or are installed for reasons other than to provide electricity services. Historically, network users — especially residential as well as small and medium commercial users — were assumed to have no ability or desire to respond to different electricity prices (i.e., were assumed to be price-inelastic) As such, the structure (and even the magnitude) of the prices and charges for electricity services sent to these users were considered an afterthought to cost-recovery and other regulatory considerations such as simplicity and equity. However, today's network users are faced with a plethora of options to change, reduce, or eliminate their consumption of electricity services — including, but not limited to, DERs. If confronted with inaccurate economic signals, these users can make decisions that drive additional costs and decrease the welfare of all system users. Aligning consumer investment decisions with efficient power system outcomes is therefore essential if society is to transition cost-effectively to an affordable and low-carbon energy system (whether this efficient outcome is distributed or centralized). **Chapter 4** highlights the "first-best" fundamentals of a comprehensive system of efficient prices for electricity services and charges for network costs and other regulated costs, identifies and comments on the implementation challenges, and discusses some ways to address these challenges.

At the same time, today's network regulations and the business models deployed by network utilities were built on assumptions developed in the 20th century. These assumptions — that power flows are unidirectional, that electricity demand is price-inelastic and will grow indefinitely, and that new equipment is the only tool for alleviating network stresses — are unfit for today's reality. The regulations governing today's power networks must be as innovative as the businesses to which they are applied. Specifically, network remuneration must account for the new options for service delivery created by DERs while maximizing the incentives for operating and building infrastructure in an efficient manner, managing the increasing uncertainty in network usage, and incentivizing the development and adoption of innovative solutions that lower cost in the near and long term. Furthermore, regulation must account for new risks such as those posed by cyber attacks on utilities or risks to consumers' privacy. **Chapter 5** develops proposals for the future of distribution network regulation and the distribution utility business model.

In the 1990s and 2000s, the emergence of efficient and modular combined-cycle gas-fired generation, escalating costs of utility generation investments, improved interconnections, and liberalization trends in the economy (among other factors) drove heated debate over the monopoly characteristics of bulk power generation and, in some locations, of retail supply. These debates led to the partial or complete restructuring and liberalization of many power markets. Similarly, the emergence of DERs is driving a new wave of debate over the proper structure of the power system, mostly at the distribution level. While there are many close parallels to the restructuring debates of the past, contemporary challenges require extending restructuring discussions "all the way to the bottom" — involving not only distribution network owners and operators, but also end consumers, aggregators (including retailers), and new businesses. Without this restructuring, incumbent utilities may be presented with inappropriate incentives and unfair competitive advantages vis-a-vis new DER businesses or may not be properly incentivized to take full advantage of the new capabilities offered by novel resources and service providers. In addition, DERs are increasingly blurring the lines between transmission and distribution systems, creating the need for new coordination regimes linking the operators and planners of these assets. **Chapter 6** carefully reconsiders the structure of the electricity industry — that of market platform, network provider, system operator, and data hub — and it also introduces several functions critical to an efficient power system, proposes possible structures to assign responsibility for these functions to avoid potential conflicts of interest and abuses of market power, and discusses the range of competitive business models that have emerged to harness DERs for electricity service provision.

The electricity markets that emerged from past restructuring processes have been continuously refined over the past two decades. Even still, the presence of DERs and large amounts of renewable, zero-marginal cost resources may require additional upgrades to nearly every aspect of these markets. Thus, **Chapter 7** proposes design changes and principles for energy and ancillary services markets, capacity remuneration mechanisms, and clean energy support mechanisms with two objectives: (1) to allow the participation of non-conventional distributed resources in existing markets, and (2) to refine market designs to ensure their efficient functioning under high penetrations of DERs and variable renewable resources.

Part 3: Insights on the economics of distributed energy resources and the competition between centralized and distributed resources

Chapter 8 builds upon the framework presented in **Chapters 4** through **7**, which addresses how to establish a level playing field for the provision of electricity services by an increasingly diverse set of resources, both centralized and decentralized. Through the lens of economic efficiency, this chapter analyzes the value of novel technologies such as DERs and explores what factors are likely to determine the portfolio of cost-effective solutions that may emerge in different power system contexts in the future. This entails considering the key trade-offs between developing distributed and centralized electricity resources and providing some insight into what will drive the optimal deployment of DERs. Key questions to be examined include how centralized and distributed resources will compete in the long run in terms of economic efficiency and whether DERs will be able to augment or replace the services traditionally provided by electricity networks.

Part 4: A Policy and Regulatory Toolkit for the Future Power System

Finally, **Chapter 9** presents a concise summary of the regulatory framework proposed by this study. The aim of this chapter is to provide regulators, policy makers, and power sector stakeholders with a concrete toolkit — a set of clear recommendations — that can facilitate the efficient evolution of the power sector.

References

ACER. 2014. *Energy Regulation: A Bridge to 2025 Conclusions Paper*. Ljubljana, Slovenia: Agency for the Cooperation of Energy Regulators.

Anaya, Karim L., and M.G. Pollitt. 2015. "Options for Allocating and Releasing Distribution System Capacity: Deciding between Interruptible Connections and Firm DG Connections." *Applied Energy* 144.

Bipartisan Policy Center. 2014. "Cybersecurity and the North American Electric Grid: New Policy Approaches to Address the Evolving Threat." February 2014.

BNEF. 2016. *New Energy Outlook 2016: Long-Term Projections of the Global Energy Sector*. London, United Kingdom: Bloomberg New Energy Finance.

BP. 2016. *BP Statistical Review of World Energy 2016*. London, United Kingdom: BP.

California ISO. 2016. *California Independent System Operator Corporation Docket No. ER16-___-000 Tariff Amendment to Implement Flexible Ramping Product.*

Campbell, R.J. 2015. *Cybersecurity Issues for the Bulk Power System.* Washington DC: Congressional Research Service. R43989; 7-5700. June 10, 2015. www.fas.org/sgp/crs/misc/R43989.pdf.

CSI. 2016. "California Solar Statistics." Accessed January 9, 2016. California Solar Initiative.

DDPP. 2015. "Pathways to Deep Decarbonization." Deep Decarbonization Pathways Project.

EIA. 2016a. "Wind Adds the Most Electric Generation Capacity in 2015, Followed by Natural Gas and Solar." *Today in Energy*. Washington, DC: US Energy Information Administration. www.eia.gov/tools/faqs/faq.cfm?id=108&t=3.

EIA. 2016b. "How Many Smart Meters Are Installed in the United States, and Who Has Them?" *Frequently Asked Questions.* Washington, DC: US Energy Information Administration. www.eia.gov/todayinenergy/detail.cfm?id=25492.

Electricity-ISAC and SANS Industrial Control Systems. 2016. "Analysis of the Cyber Attack on the Ukrainian Power Grid." Electricity-Information Sharing and Analysis Center. ics.sans.org/media/E-ISAC_SANS_Ukraine_DUC_5.pdf.

EPA. 2015. 40 CFR *Part 60: Carbon Pollution Emission Guidelines for Existing Stationary Sources: Electric Utility Generating Units.* Washington, DC: US Environmental Protection Agency.

Eurelectric. 2013. *Utilities: Powerhouses of Innovation.* Brussels, Belgium.

European Commission. 2014. "2030 Climate & Energy Framework." *Climate Action*. ec.europa.eu/clima/policies/strategies/2030/documentation_en.htm.

European Commission. 2015. *Energy Union Package: A Framework Strategy for a Resilient Energy Union with a Forward-Looking Climate Change Policy.* Brussels, Belgium.

European Commission. 2016. "Smart Metering Deployment in the European Union." ses.jrc.ec.europa.eu/smart-metering-deployment-european-union.

Frankfurt School-UNEP Centre, and BNEF. 2016. *Global Trends in Renewable Energy Investment*. Frankfurt, Germany.

Fraunhofer ISE. 2016. *Recent Facts about Photovoltaics in Germany*. Freiburg, Germany.

Global Wind Energy Council. 2016. *Global Wind Report: Annual Market Update 2015*. Brussels, Belgium.

IEA. 2015. *World Energy Outlook 2015*. Paris, France: International Energy Agency.

Jolly, David. 2015. "Norway Is a Model for Encouraging Electric Car Sales." *The New York Times*. October 16, 2015. www.nytimes.com/2015/10/17/business/international/norway-is-global-model-for-encouraging-sales-of-electric-cars.html.

Kasera, Akshay. 2016. "New York Market Orientation Course: Demand Response." Rensselaer, NY.

Liebreich, Michael. 2016. "In Search of the Miraculous." Bloomberg New Energy Finance Summit. New York, New York: Bloomberg New Energy Finance. April 15, 2016. https://data.bloomberglp.com/bnef/sites/4/2016/04/BNEF-Summit-Keynote-2016.pdf.

McAnany, James. 2016. *2015 Demand Response Operations Markets Activity Report: May 2016*. PJM. May 9, 2016.

Monitoring Analytics LLC. 2016. 2015 *State of the Market Report for PJM*.

Munsell, Mike. 2016. "US CHP and Fuel Cell Capacity Set to Grow 11GW by 2026." *Greentech Media*. May 19, 2016. www.greentechmedia.com/articles/read/us-chp-and-fuel-cell-capacity-set-to-grow-by-11-gw-by-2026.

Munsell, Mike. 2016. "US Energy Storage Market Grew 243% in 2015, Largest Year on Record." *Greentech Media*. March 3, 2016. www.greentechmedia.com/articles/read/us-energy-storage-market-grew-243-in-2015-largest-year-on-record.

National Academies of Sciences, Engineering, and Medicine. 2016. *The Power of Change: Innovation for Development and Deployment of Increasingly Clean Electric Power Technologies*. Washington, DC: The National Academies Press.

North Carolina Clean Energy Technology Center. 2016. *50 States of Solar: Q2 2016 Quarterly Report*. Raleigh, NC.

Nykvist, Björn, and Måns Nilsson. 2015. "Rapidly Falling Costs of Battery Packs for Electric Vehicles." *Nature Climate Change* 5: 329-332.

Ofgem. 2015. *Non-Traditional Business Models: Supporting Transformative Change in the Energy Market.* London, United Kingdom: Office of Gas and Electricity Markets.

PricewaterhouseCoopers. 2015. *A Different Energy Future: Where Energy Transformation Is Leading Us*. 14th PwC Global Power & Utilities Survey.

REN21. 2016. REN21 *Renewables Global Status Report*. Paris, France: Renewable Energy Policy Network for the 21st Century.

Schatzki, Todd, and Paul Hibbard. *Assessment of the Impact of ISO-NE's Proposed Forward Capacity Market Performance Incentives*. Boston, MA: Analysis Group.

Schmalensee, Richard. 2016. "The Performance of U.S. Wind and Solar Generators." *The Energy Journal* 37(1): 123–151.

Schweppe, Fred C. 1978. "Power Systems 2000: Hierarchical Control Strategies: Multilevel Controls and Home Minis Will Enable Utilities to Buy and Sell Power at Real Time Rates Determined by Supply and Demand." *IEEE Spectrum* 15.7: 42–47.

The Brookings Institution. 2011. *Assessing the Role of Distributed Power Systems in the U.S. Power Sector*. Washington, DC.

Trabish, Herman K. 2016. "17% of Hawaiian Electric Customers Now Have Rooftop Solar." *Utility Dive*. February 1, 2016.

Trabish, Herman K. 2016. "How Storage Can Help Solve the Distributed Energy 'Death Spiral': A New Navigant Paper Lays out Three Examples of How Storage Can Facilitate the Transition to a More Distributed Grid." *Utility Dive*. June 21, 2016.

US DOE. 2016. *Revolution… Now: The Future Arrives for Five Clean Energy Technologies—2016 Update*. Washington, DC: US Department of Energy. September 2016.

US DOE. 2015a. *Direct Federal Financial Interventions and Subsidies in Energy in Fiscal Year 2013*. Washington, DC: US Department of Energy.

US DOE. 2015b. S*tate of the States: Fuel Cells in America 2015*. 6th Edition. Washington, DC: US Department of Energy.

US DOE and Sandia National Laboratories. "DOE Global Energy Storage Database." www.energystorageexchange.org/projects.

US DHS. "Critical Infrastructure Sectors: Energy Sector." US Department of Homeland Security. Accessed March 11, 2016. www.dhs.gov/critical-infrastructure-sectors.

Wesoff, Eric. 2015. "First Solar CEO: 'By 2017, We'll Be Under $1.00 per Watt Fully Installed'." *Greentech Media*. June 24, 2015. www.greentechmedia.com/articles/read/First-Solar-CEO-By-2017-Well-be-Under-1.00-Per-Watt-Fully-Installed.

Wesoff, Eric. 2016. "How Soon Can Tesla Get Battery Cell Costs Below $100 per Kilowatt-Hour?" *Greentech Media*. March 15, 2016. www.greentechmedia.com/articles/read/How-Soon-Can-Tesla-Get-Battery-Cell-Cost-Below-100-per-Kilowatt-Hour.

Zibelman, Audrey. 2016. "REVing Up the Energy Vision in New York: Seizing the Opportunity to Create a Cleaner, More Resilient, and Affordable Energy System." IEEE *Power & Energy Magazine* 14(3): 18-24. May-June 2016.

Zinamen, Owen et al. 2015. *Power Systems of the Future*. Golden, CO: National Renewable Energy Laboratory.

PART 1: UNDERSTANDING ELECTRICITY SERVICES AND HOW DISTRIBUTED ENERGY RESOURCES AFFECT THE DESIGN AND OPERATION OF POWER SYSTEMS

02

New Options for the Provision and Consumption of Electricity Services

This report is not about the future of today's utility. Rather, it is about how the services provided by electric utilities today will be provided in the future — by the utility of the future. Since enabling the efficient provision and consumption of such services is the major focus of this study, we begin here by defining what we mean by electricity services. Understanding these services is critical to realizing why the new means of providing such services — mainly distributed generation, energy storage, and demand response — can be particularly impactful. Furthermore, understanding and communicating the value of the electricity services that a given agent consumes is a precursor to allowing that agent to be more price-responsive and efficient.

2.1 Power System Services

Power systems are rapidly evolving to integrate participation by many agents playing diverse roles; furthermore, these agents may adopt different roles at different times (e.g., consumer at one hour, producer at another). In the context of the power system, an agent's role is characterized by the electricity services it provides to and obtains from other agents.

Electricity services are activities performed in the context of a power system that have economic value for some agents, regardless of whether this value is monetized. Electricity services are distinguished from other more generic services in that they create economic value by enabling the consumption of electrical energy, lowering the costs associated with consuming electrical energy,[1] or both. Distributed energy resources (DERs) can provide electricity services, as can any other energy resource.

1 Note that these costs include both private costs internalized in market transactions and public costs or externalities not monetized in market transactions.

Electricity services are activities performed in the context of a power system that have economic value for some agents, regardless of whether this value is monetized. Electricity services create economic value by enabling the consumption of electrical energy, lowering the costs associated with consuming electrical energy, or both.

Electricity services vary widely in their nature: They may be physical, financial, or information-based, and they may vary in format, with different durations, levels of commitment, cost allocation methods, and economic implications.

This section logically defines electricity services. We do not attempt to define the most efficient, fundamental, or mutually exclusive and collectively exhaustive set of services. Rather, we provide a definition that is robust to different regulatory, market, and technology paradigms that can be used to understand how DERs — or any other energy resource — may create value in power systems.

2.1.1 Power systems operated and planned to maximize social welfare

Power systems are incredibly complex constructs, often with millions of agents interacting under a multiplicity of physical laws, policies, regulations, and contracts. To better understand the nature and origin of services, it is useful to conceptualize the planning and operation of power systems within an optimization framework.

Within this framework, the objective of all planning and operations is to maximize social welfare — that is, to maximize the sum of all of the consumer surplus created by electricity consumption, minus the sum of the costs of all of the actions required to supply the consumed electricity.[2] These costs emerge from producing, consuming, storing, transmitting, or transforming other forms of energy into electricity (and vice versa), and they include both private costs internalized in market transactions as well as public costs or externalities not monetized in transactions. This optimization is subject to multiple constraints that emerge from the aforementioned physical laws, policies, regulations, and contracts.

Viewed through the framework of optimization, a binding or active constraint (whether imposed by physical laws, regulations, policies, or any other source) creates a shadow price that reflects the marginal value of either relaxing (i.e., changing the value of) the constraint or finding alternative ways to meet it.[3] That is, this shadow price reflects the increase in social welfare if the constraint is relaxed at the margin. An active constraint therefore creates the opportunity of taking action to relax or relieve it, increasing the welfare created by the use of electricity in the power system; this action is an electricity service, by our definition.

2.1.2 Electrical energy

At the core of all electricity services is electrical energy. Electrical energy is indisputably the most fundamental service in the power system and the reason for its existence. Consumers of electrical energy value the useful work that it can provide (e.g., heating, cooling, lighting, and powering electronics and transport). That is, consuming electrical energy creates utility for these consumers. Of course, some of the value created by electrical energy could alternatively be provided by other forms of energy, such as thermal energy.

At the core of all electricity services is electrical energy. Electrical energy is indisputably the most fundamental service in the power system and the reason for its existence.

2 See Chapter 2 of *Regulation of the Power Sector* (Perez-Arriaga et al. 2013) for this mathematical formulation and its implications.

3 In the optimization literature, this price is referred to variously as a “shadow price,” “dual variable,” or “Lagrange multiplier” of the constraint.

Given that this study focuses on developed power systems in Europe and the United States, we will focus on alternating current (AC) electrical energy. Instantaneous AC electrical power results from the multiplication of sinusoidal current and voltage waveforms, which are typically out of phase by a phase angle and can be conveniently decomposed into two components — that of active and reactive power.[4] At any given point in the system and at any moment of time, only current, voltage, and phase angle can be measured. However, these measurements permit the computation of active and reactive power, which is useful in the management of power systems, as discussed below and in **Chapter 4**.

AC power systems have two key physical features that are fundamental to understanding electricity services. First, the supply and demand of power must be in balance at all times.[5] In the previously described optimization framework, the balancing of supply and demand creates a value (or price) for electric power at a given time. When the power waveform is conveniently decomposed into active and reactive power, we can consider values for both.

Second, electrical energy must be delivered from the site of production to the site of consumption. This is done through the use of electrical networks, or grids (whether these grids are very large or very tiny). Electricity and electricity networks are subject to physical laws and constraints: Conductors and transformers heat up as current passes through them, producing resistive losses in the form of waste heat. Moreover, conductors and transformers must be maintained below specified temperatures. In addition, the frequency of the voltage and current waveforms must be maintained within tight bounds. The interaction of the power balance constraint with these physical network constraints (and the various other classes of constraints introduced next) creates unique values of power at unique points and times in the network (in fact, unique values of real and reactive power). Mathematical and computational techniques exist such that the price of electricity can be computed at any given time and location within the network (Bohn et al. 1984). These prices are defined according to the well-established theory of spot pricing, which, conceptually, can be extended to every corner of the electric grid through locational marginal prices.[6]

Alternating current power systems have two key physical features that are fundamental to understanding electricity services. First, the supply and demand of power must be in balance at all times and at all locations. Second, electrical energy must be delivered somehow from where it is produced to where it is consumed through the use of electrical networks subject to a number of physical laws and constraints. The interaction of the power balance constraint with these physical network constraints creates unique values of power at unique points and times in the network.

4 See Appendix B: Electric Power System Basics in *The Future of the Electric Grid* (MIT 2011), for an introduction to the basic concepts of AC power systems.

5 Whatever produced electricity is not consumed must be stored or dissipated. Electricity can be stored in other forms of energy (e.g., chemically in a battery, mechanically in a compressed gas in a cavern or in a pumped hydro reservoir, or thermally in molten salt, etc.) and released again as electric energy, with some losses. These energy storage devices must be operated such that the instantaneous power balance is met — that is, the presence of storage does not eliminate the need to meet power balance constraints, it simply provides new options for helping meet this constraint.

6 We ignore here that there may not be a unique way of computing the price of electricity, since, even if the difficulties of computing the spot price of electricity everywhere are overcome, many other challenges remain (such as how to account for the nonlinearities in the startup costs and the heating rates of thermal power plants) (Gribik et al. 2007; O'Neill et al. 2005). In many power systems, the spatial differentiation of prices due to network losses and network constraints is ignored. Nodal prices (i.e., locational marginal prices, or "LMPs") are computed for transmission nodes in only some power systems, mostly in North and South America and Australia. Zonal prices, with very large zonal definition (e.g., France as one zone and Germany plus Austria and Luxemburg as another) now exist in Europe. To our knowledge, nodal prices are presently not used at the distribution level in any power system. LMPs used in the distribution system are referred to here as distribution-level LMPs, or "DLMPs."

In theory, in the absence of uncertainty, the only service that needs to exist is the provision of power (conventionally decomposed into active and reactive) at a location and a point in time. The prices that are determined in real time could, in theory, coordinate the consumption and production of power at all points to respect the physical power balance and network constraints and motivate actors to invest sufficiently (with the important exclusion of networks, which are not fully compensated by spot prices[7]).

However, in practice, many challenges emerge. First, given that electricity supply and demand must be balanced at all times, the behavior of the many thousands of agents supplying electricity services and the many millions of agents consuming these services must somehow be coordinated. Second, there are inherent uncertainties in predicting the availability of supply and the level of demand at all times. Finally, these coordination and stochasticity problems are exacerbated by many of the technical characteristics of power systems that make them challenging to operate. Thus, in practice, regulators, system operators, or policy makers frequently place constraints on the planning and operation of the power system — above and beyond existing physical constraints — that define additional services. These constraints are intended to transform complex decision-making processes into simple formulations that can be easily implemented by power system operators and planners (e.g., "the amount of secondary operating reserves should be equal to the rated capacity of the largest generation unit in operation plus 3 percent of the expected demand," "the power flow in a certain line cannot exceed 80 percent of its thermal rated capacity," or "the amount of firm-installed capacity must exceed estimated peak demand by 10 percent"). These constraints are simplifications, and in some cases they emphasize security of supply over pure economic efficiency. Such constraints take two primary forms: coordinating constraints and policy constraints.

7 While marginal pricing of electricity can encourage adequate investment in generation capacity (ignoring again the challenges of computing accurate marginal prices for electricity), economies of scale combined with the discrete nature of network investments and other factors make marginal pricing an insufficient method for encouraging adequate investment in network infrastructure (Pérez-Arriaga et al. 1995).

The goal of the imposition of these constraints should be to coordinate operational decisions in a manner that represents the optimal trade-off between the total system costs for planning and operation and the costs associated with losing energy supply (i.e., the costs of blackouts due to any number of reasons, from lack of operating reserves to network failures, etc.). In practice, constraints may simply reflect the effect of risk aversion or incentives on the regulator or the system operator.

In practice, regulators, system operators, or policy makers frequently place constraints on the planning and operation of the power system — above and beyond existing physical constraints — and these constraints define additional services. Intended to transform complex decision-making processes into simple formulations that can be easily implemented by power system operators and planners, these constraints take two primary forms: coordinating constraints and policy constraints. Examples include operating reserves, firm capacity, and electricity network capacity margins.

2.1.3 Coordinating and simplifying constraints

Some constraints are placed on power systems to help simplify or coordinate complex investment, operational, planning, or market decisions. For example, various forms of operating reserves are created when a constraint is placed on the amount of capacity that must be held "in reserve" to help meet power balance constraints in the case of unexpected failures of power plants or lines or of errors in forecasted demand or variable renewable energy supply. These reserves take the form of forward commitments to provide power at a later time purchased by system operators (on behalf of electricity consumers). The power balance is constantly changing: Agents switch devices on or off, power plants or lines fail, wind and solar production vary, etc. Agents selling electric energy have an inherent economic interest in being prepared to respond to these changes: This readiness could allow them to sell more electricity at a profit. Thus, some level of reserve is provided for "naturally."

However, given the realities of free riders, imperfect markets, imperfect information, the extremely rapid time scales involved (e.g., fractions of a minute, in some cases), and the critical nature of electricity in modern economies, regulatory intervention has been universally adopted to ensure that sufficient reserve capacity is procured. System operators, tasked by regulators and policy makers with ensuring that the power system operates within secure limits, set constraints so that the agents in the power system are ready to respond to imbalances between electricity supply and demand with a prescribed total volume in a prespecified time range. The generation mix of power systems is very diverse in different parts of the world, and system operators may therefore establish very different technical requirements for these "operating reserves." Many definitions of reserves exist — see Ela et al. (2014), ENTSO-E (2013), and Pérez-Arriaga (2011) for reviews of different reserve definitions. Nonetheless, the service emerges in the same way in these various contexts: by adding a minimum level of reserves (a constraint) and by specifying the characteristics of these reserves.

Similarly, left to their own devices and based only on the current and expected future demand and prices for power and reserves, certain agents may decide to build new power-generating capacity (or other energy resources) in anticipation of future profits. However, should a regulator or policy maker decide to intervene to ensure that a level of capacity will be available to meet demand at some future date, a new service is created: firm capacity. There is no conceptual agreement on whether firm capacity should or should not be a service in electric power systems. Experts continue to debate the issue,[8] although in practice many power systems have some kind of intervention in this regard, resulting in a commodity and an associated price.[9]

The previous two examples discussed constraints placed to ensure that a minimum amount of energy resources will be available at future dates (with reserves representing an option purchased to ensure the demand balance in real time, and firm capacity representing an option purchased to ensure the demand balance can be met at some expected future date). Other similar coordinating constraints may be implemented. We focus here not on enumerating all of the possible constraints, but on expounding the logic behind the origin of services.

The same types of simplifying or coordinating constraints can be applied to the functioning of networks. For example, in many systems, strict transfer capacity limits are placed on transmission or distribution lines — that is, system operators dictate that the instantaneous power transfer on a line cannot exceed a certain amount. These constraints can emerge from thermal, stability, or voltage issues. However, in many cases, the imposed constraints are not based on the real-time physical limits of a piece of equipment, but rather on heuristics that simplify complex physical realities and enable safe and reliable operations given uncertainty. For example, operators

8 See Hogan and the Association for the Cooperation of Energy Regulators (2013) for recent reviews of theory and a discussion of the implementations of firm capacity markets.

9 The format of these interventions is very varied (capacity payments, capacity markets, strategic reserves, reliability options, additions to energy market prices, etc.). These interventions are generically termed "capacity remuneration mechanisms" (CRMs). There are few truly "energy-only" markets. The Australian National Electricity Market appears to be the perfect example of an energy-only market, with the peculiarity that basically all the demand is hedged against the very high prices that sometimes occur in this market by call option contracts that are signed by the retailing companies or suppliers. Texas is another interesting example. In the Texas market, as in the UK "electricity pool" in the 1990s, a regulated component is added to the day-ahead energy price to "make it right" (Newell et al. 2012) despite reserve margins declining with load growth and retirements, investment appears to have stalled. Many projects have been postponed or cancelled and no major new generation projects are starting construction. As a result, ERCOT projects that reserve margins will fall to 9.8 percent by 2014, substantially below its current reliability target of 13.75 percent. Reserve margins will decline even further thereafter unless new resources are added. Generation investors state that a lack of long- term contracting with buyers, low market heat rates, and low gas prices in ERCOT's energy-only market make for a uniquely challenging investment environment. In response to these concerns, the Public Utility Commission of Texas (PUCT). This kind of regulatory intervention may be justified to better approximate the true value of energy. However, if these top-up payments go beyond this energy price extension with the purpose of promoting resource adequacy, they in essence become a CRM.

may approximate the transfer capacity of a line by estimating its thermal state without either measuring the conditions of the line directly or accounting for all of the environmental and system factors that might influence its true thermal state (and therefore transfer capacity). This transfer capacity constraint is a simplifying constraint made to ensure the safe operation of the power system.

Reliability constraints simplify complex economic and technical calculations to develop measures of failure and maximum System Average Interruption Duration Indices (SAIDI) or System Average Interruption Frequency Indices (SAIFI). In North America, these constraints are placed on the transmission system by the North American Electric Reliability Corporation, and at the distribution system by state regulators and other more local bodies.

Another common class of services is created by constraints implemented to internalize externalities or achieve other policy objectives.

Network owners may also include some sort of "network capacity margin" — that is, they may impose a constraint on the minimum margin between network transfer capacity and the peak expected power flows through that network infrastructure. This may take the form of a network upgrade heuristic or planning margin. The possibility of this constraint becoming binding creates an opportunity to provide a service by either building more network capacity or securing commitments for future actions that can reduce peak power flows (e.g., through the installation and operation of DERs, contracting with flexible demand, etc.).

2.1.4 Policy and other constraints

The first two categories of services emerged from constraints imposed by the physics of power systems and those imposed to simplify, coordinate, and ensure the safe operation and planning of these systems. Another common class of services is created by constraints implemented to internalize externalities or achieve other policy objectives. Policy makers may create constraints such as mandates for maximum levels of carbon emitted or minimum levels of renewable energy resources or DERs used in a power system. For example, maximum carbon emissions constraints create a price for emitting carbon, creating a service for carbon emissions reduction. Typically, these constraints are created to internalize externalities that are not currently accounted for in power systems operation, planning, and markets or to further other policy objectives. In the carbon emissions example, the constraint enables carbon emitters and/or consumers to internalize the "cost" that carbon emissions impose on society. In cases like this, it would be economically inefficient for regulators to impose constraints if the cost of imposing the constraint exceeded the value of the externality or public good.[10]

2.1.5 Services without constraints

Thus far we've highlighted the various ways that constraints can create electricity services and how the activation of these constraints creates opportunities to create value by providing a service. However, our definition of electricity services is not restricted to the creation of services by constraints. Any action that creates value by enabling the consumption of electrical energy, lowering the cost of consuming electrical energy, or both, is an electricity service. For example, conservation voltage reduction (CVR) — a technique that reduces energy consumption by reducing the voltage of the consumed energy — is a service that can, in certain cases, increase social welfare by lowering the total cost of energy delivery. However, the economic value of this service does not emerge directly out of an activated constraint; rather, it is measured by the cost of implementing the CVR solution and the value of the

10 This is a simple characterization of the use of constraints and market mechanisms to internalize externalities. Externalities can also be internalized with price instruments such as taxes (e.g., a carbon tax for internalizing the cost of carbon emissions) or incentives (e.g., clean energy credits for internalizing the value of clean energy technologies). These pricing mechanisms often implicitly represent a target for the taxed or incentivized activity or resource.

energy saved (and any other potential benefits). Similarly, actions that reduce losses may create value (loss reduction as a service), but this value does not emerge from an activated constraint. Thus, the placement of constraints always creates a service, but a service does not always emerge from a constraint.

2.1.6 Summarizing services

In summary, we define electricity services as those activities that are performed within a power system and that create economic value for some agents by enabling the consumption of electrical energy, lowering the cost of consuming electrical energy, or both. Any energy resource, either centralized or distributed, can provide electricity services. The origin of many of the most common electricity services is the constraints that the laws of physics, policy makers, and regulators impose on the operation and planning of the power system.

2.1.7 Economic value and monetized value

Note that here we focus on economic value — the ability to increase welfare by enabling greater electrical energy consumption and/or lowering the total cost of energy consumption. This economic value may or may not be accurately monetized and reflected in the price of or compensation for various services or commodities in the power system. As a result, we must distinguish between the intrinsic economic value of a given service and the way it is monetized and compensated in practice. In **Chapter 4** of this study, we discuss how services should be monetized to lead to efficient outcomes, and we address the various intricacies associated with this monetization. The goal of the comprehensive system of cost-reflective prices and charges introduced in Chapter 4 is to ensure that the marginal cost to an agent of consuming a service or the marginal value for the agent of providing that service is equal to the marginal increase or reduction of the total social welfare created by that action. That is, we attempt to align the cost or benefit of consumption or production for an agent (the private value for that agent) with the cost or benefit accrued to the entire power system (the system value). Distortions are created when the private value created by an action is less than or greater than the system value of that action.

Finally, we revisit the optimization framework described herein in **Chapter 8**, where we explore how DERs and conventional energy resources interact and compete to create economic value in different contexts.

2.2 Understanding DERs and the New Ways of Providing Electricity Services

We now turn our attention to understanding one of the primary new methods of providing electricity services: distributed energy resources (DERs). This section expands upon the definition of DERs provided in **Chapter 1** and highlights some of the key characteristics of DERs. We now focus on the factors that distinguish distributed resources from other resources.

As noted in Chapter 1, this study defines DERs broadly: DERs are any resources capable of providing electricity services and located in the distribution system. DERs are not a new phenomenon. Combined heat and power plants, controllable water heaters, district cooling plants, and dozens of other types of DERs existed long before the global implications of DERs on the electricity industry was a topic of serious discussion. As a result, many definitions and taxonomies of DERs are available.[11] In contrast to centralized resources (e.g., conventional power plants or pumped-hydroelectric energy storage), DERs are characterized by relatively small capacities (a few kilowatts to a few megawatts), and are connected to lower voltage electricity distribution grids (as opposed to transmission and high-voltage distribution systems).

DERs can be divided into two distinct classes: DERs installed specifically to provide electricity services (e.g., energy storage devices, solar photovoltaic systems with smart inverters, power electronics, or distributed fossil generation);[12] and resources that exist primarily for reasons other than to provide electricity services but that

11 See Ackermann et al. (2001) and Pepermans et al. (2005) for often-cited academic definitions. See DNV GL (2014) for a more updated survey.

12 Of course, not every solar PV system or distributed generation system is installed for the sole or even primary purpose of providing electricity services. Some PV systems, for example, may be installed simply because the systems look "cool" or because their owners want to feel independent from the grid. Nonetheless, for most investors in solar PV or energy storage, the savings or earnings associated with the sale of electricity services are an important motivator.

can be harnessed for this purpose (e.g., flexible demand and electric vehicles). Cost-reflective prices and charges for electricity services — which will be discussed further in **Chapter 4** — can create the conditions to incentivize DERs of the first class to be installed only when DERs will add value. Furthermore, these cost-reflective prices and charges enable existing demand and DERs to create value by exposing the owner to the potential benefit of engaging in the market.

We note that demand is both a consumer of electricity services and a potential supplier of services, and we thus consider demand a potential resource. However, it is important here to distinguish between demand providing a service and the more efficient consumption of electricity services. The act of deciding to consume or not consume electrical energy services based on the price of such services does not itself provide an electricity service. Responding to high electricity prices by curtailing consumption is simply an expression of the degree to which a given agent values the consumption of electrical energy (i.e., their willingness to pay). This flexible demand can be an important source of improved efficiency in power systems but constitutes the economically efficient consumption of a service, not the provision of a service. We refer to demand that serves as an economically efficient consumer of electrical energy as "price-responsive demand."

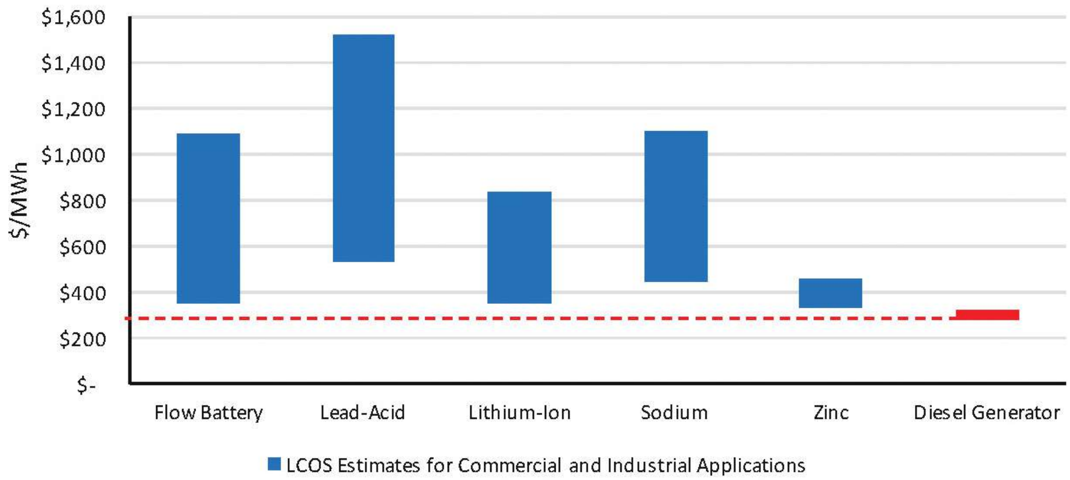
This stems from the fact that demand resources cannot provide energy — they can only consume energy. Of course, agents can procure the option to consume energy (either through a retail supply contract or by participating in forward markets) and sell that option (Borlick 2010; Hogan et al. 2012). Alternatively, an agent with flexible demand may be able to make a forward commitment to provide a future response, where the commitment itself has value beyond the final energy consumption (or lack thereof). A flexible demand agent may commit in advance to stand by and be ready to adjust consumption in order to deliver an additional service, such as operating reserves, firm capacity, or network capacity margin. This commitment itself delivers value distinct from the actual final consumption of energy, may incur additional standby or commitment costs, and may require payment of additional compensation beyond avoiding paying for energy that was not consumed. We consider these cases to constitute demand as a service provider and refer to these cases collectively as "demand response."

Box 2.1: The Limitations of the Distributed Energy Resources Lens

This study focuses on new ways of providing electricity services and on the impacts of such innovation on the power sector. Today, distributed energy resources (DERs) are among the major new sources of electricity services. The rapid proliferation of DERs in many power systems globally is one of the primary motivators for this study, and DERs merit an in-depth discussion. The remainder of this chapter focuses on these resources.

However, it is important to remember that DERs are not the only new energy option. Many new technologies and activities such as network topology control or advanced transformers are capable of providing electricity services, although they do not fit cleanly into existing definitions of DERs. Furthermore, as we highlight throughout the report, the efficient consumption of electricity services may be among the most cost-effective means of improving the success of the power sector in providing welfare. Thus, we encourage readers to keep sight of the broader focus on decentralizing and expanding options for providing and consuming electricity services.

Figure 2.1: Levelized Cost of Storage[14] (Dollar per Megawatt-hour) for Various Technologies in Commercial and Industrial Applications



Source: Lazard (2015). Levelized cost of storage calculations assume 10-year project life, 1 megawatt (MW), 4 megawatt-hour (MWh) battery, 350 cycles per year at 100 percent depth of discharge, and $50 per MWh cost of charging. Diesel generator costs are provided for comparison.

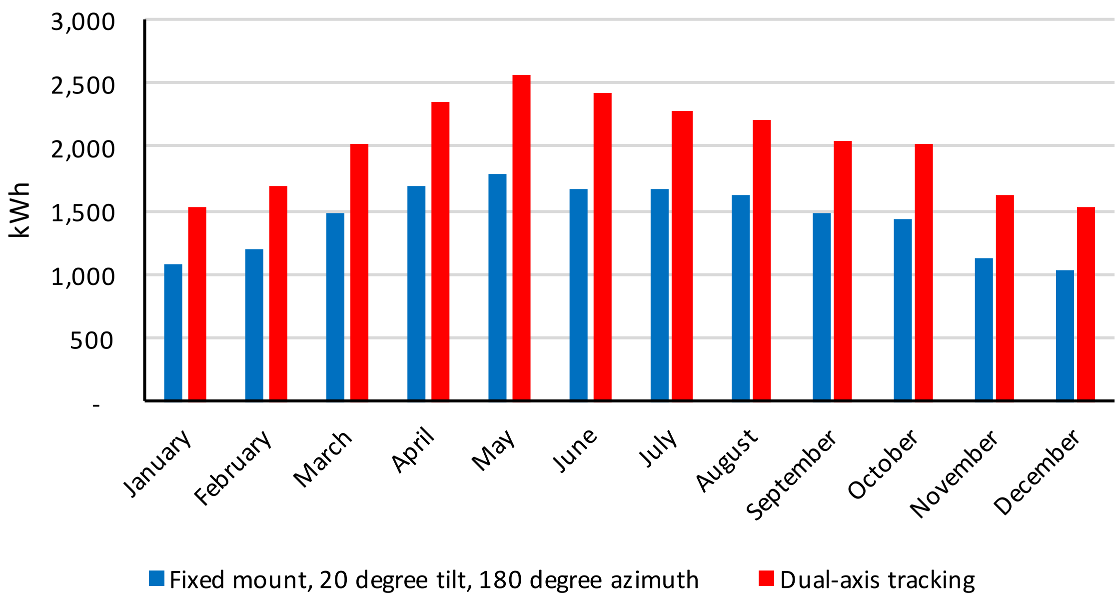
It is important also to note that the technical and economic characteristics of DERs are as varied as the DERs themselves, and there is significant variation in cost and performance among even very similar DERs. This is critical, as it highlights the futility of attempting to define a single value for all DERs.[13] **Chapter 8** explores this topic in detail and provides quantitative evidence.

First, there are cost and performance differences between technologies within a given class of DERs. For example, in the electrochemical battery class, lithium-ion batteries and lead acid batteries have very different performance characteristics. Similarly, within the class of solar PV technologies, thin film and crystalline silicon (c-Si) systems have different characteristics. **Figure 2.1** presents an example of the significant variation in cost of various energy storage technologies deployed at commercial and industrial facilities.[14]

13 This is not a purely academic concept. The proliferation of DERs has spurred a multitude of studies, utility programs, and tariff structures that attempt to define the value of DERs. For example, "value of solar" tariffs — that is, a tariff that attempts to value all of the benefits (and costs, if applicable) of distributed solar and applies only to owners of distributed solar systems — has emerged as a popular regulatory tactic. Spain has implemented a fee, targeted at generation behind the meter from DERs with capacities above 10 kilowatts, to prevent non-recovery of regulated network and policy costs. Consolidated Edison — a distribution utility in New York State — and other utilities have offered bill rebates for customers who purchase smart thermostats and enable the utility to control them during certain periods. All of these efforts are an attempt to define either the value or cost of a specific DER category (e.g., flexible demand or solar PV). In this same vein, a recent draft manual issued by the US National Association of Regulatory Utility Commissioners (NARUC) reviews and evaluates methods to "compensate" DERs by focusing on the value of each technology (NARUC, 2016).

14 The levelized cost of storage (LCOS) is similar to the levelized cost of energy in that it attempts to create a metric for the comparison of costs across technologies and does not attempt to calculate the value of the technology or the application. LCOS is calculated by dividing the total energy provided by the storage system (with appropriate charging and discharging losses) by the total cost of building and operating the system over the system's lifetime.

Figure 2.2: AC Energy Produced by 10-Kilowatt c-Si PV Systems in Phoenix, Arizona, Deploying Different Mounting Systems

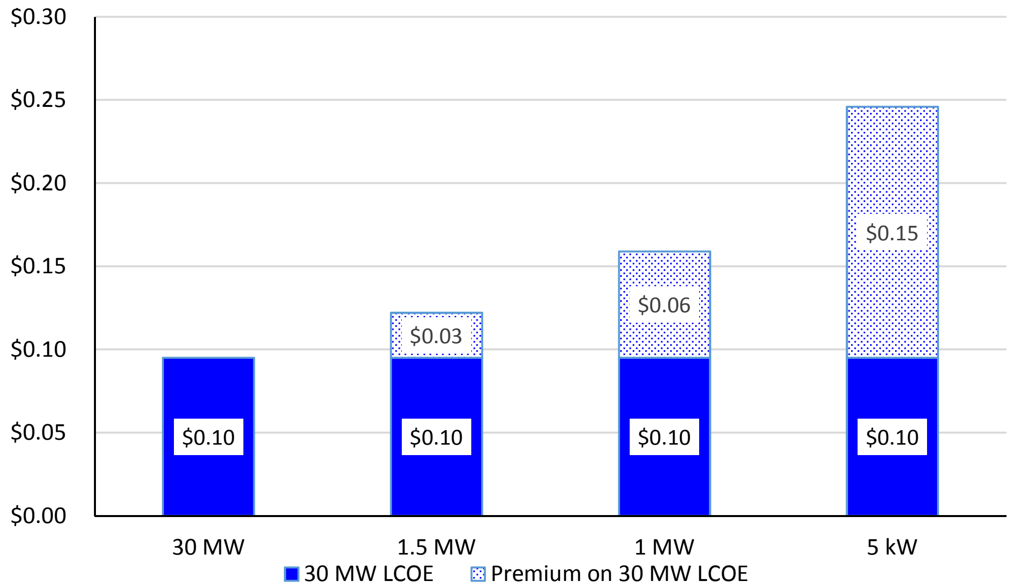


Source: NREL: PV Watts Calculator.

Second, there are cost and performance differences between technologies of a given type. For example, a lithium-ion battery with two hours of storage capacity may cost much less than a lithium-ion battery with four hours of storage (where lithium-ion batteries is the technology type). A c-Si PV module with 15 percent efficiency may likewise cost significantly less than a c-Si module with 20 percent module efficiency (where c-Si PV is the technology type). **Figure 2.2** also illustrates the performance variation within the c-Si PV technology type. Two identically sized (10 kilowatt) and identically located (Phoenix, Arizona) solar PV systems can produce dramatically different energy outcomes depending on orientation and whether or not they employ tracking systems. In this example, PV systems deploying dual-axis tracking systems produced up to 33 percent more energy in a given month than PV systems without tracking.

Finally, there are cost and performance differences within identical technologies at a given scale. For example, the levelized cost of a c-Si PV system installed at the 5-kilowatt scale will be significantly higher than the levelized cost of a c-Si system installed at the 100-megawatt scale. **Figure 2.3** shows the variation in 2015 levelized costs of energy (LCOEs) for solar PV at different scales on Long Island in New York State.

Figure 2.3: LCOEs of Solar PV Installations of Various Scales on Long Island in New York State with 2015 System Pricing, $/kWh[15]



This section introduced some key DER characteristics with the purpose of highlighting the futility of assigning a single value to DERs and the importance of establishing a system of cost-reflective prices and charges for electricity services. The following section explores the factors that uniquely differentiate DERs, shedding light on their role in power systems and further underscoring the dramatic variation in their value.

2.3 What Do DERs Do Differently?

2.3.1 DERs provide locational services

The first and primary differentiating factor for DERs is that their distributed nature enables them to provide services either more effectively, cheaply, or simply in locations inaccessible to more centralized resources. This capability is important, because the value of some services changes with the location of provision.[16] This emerges from the physical characteristics of the networks that connect all power system agents, including losses, capacity limits of network components, and voltage limits at network nodes. DERs, enabled by their distributed nature, can provide services in areas of the power system where these services are most valuable. Understanding the locational nature of DERs allows us to understand the value of investing in a resource or operating it in one location versus another.

Table 2.1 classifies various commonly cited DER values as locational or non-locational. In addition, it distinguishes between values derived directly from the power system and other values not directly associated with the provision of electricity services.

15 Assumes capital costs of $1,500 per kilowatt of alternating current (AC) power at 30 megawatts (MW) AC, $2,000 per kilowatt (kW) AC at 1.5 MW, $2,600 per kW AC at 1 MW, and $4,100 per kW at 5 kW. All systems produce 1,458 kilowatt-hours (kWh) per kW, or a roughly 16.6 percent capacity factor. This capacity factor is consistent with a 1-kW AC system at a fixed 40.8-degree tilt and 180-degree azimuth in Long Island, New York (latitude: 40.78, longitude: -73.1).

16 The value of many services also changes with time, but this is not important to understanding the distinguishing features of DERs.

Table 2.1: Classification of DER Values

	LOCATIONAL	NON-LOCATIONAL
POWER SYSTEM VALUES	Energy Network Capacity Margin Network Constraint Mitigation Power Quality[17] Reliability and Resiliency Black-start and System Restoration	Firm Generation Capacity[18] Operating Reserves[19] Price Hedging
OTHER VALUES	Land Use Employment Premium Values[20]	Emissions Mitigation Energy Security

To make this conversation more concrete, **Figure 2.4** shows the prices of energy throughout the interconnect managed by PJM, a regional transmission organization, in an example hour (July 19, 2013, at 4:05 p.m. ET). As shown, energy was clearly more valuable near New York, New York, than near Cincinnati, Ohio, on this day and at this time. The same concept applies to energy (and other locational services) within a distribution network. If the value of a service is higher as one goes "down" into lower voltages of the distribution system, it may be more beneficial to have DERs provide this service than centralized resources.

The value of locational services can be persistently different — not just different in one hour as presented in Figure 2.4. **Figure 2.5** shows a histogram of the daily average marginal cost of energy at all PJM aggregate nodes in 2015. While most nodes have a daily average locational marginal price (LMP) of roughly $22.5 per megawatt hour (MWh) to $30 per MWh, some nodes have daily average LMPs above $105 per MWh.

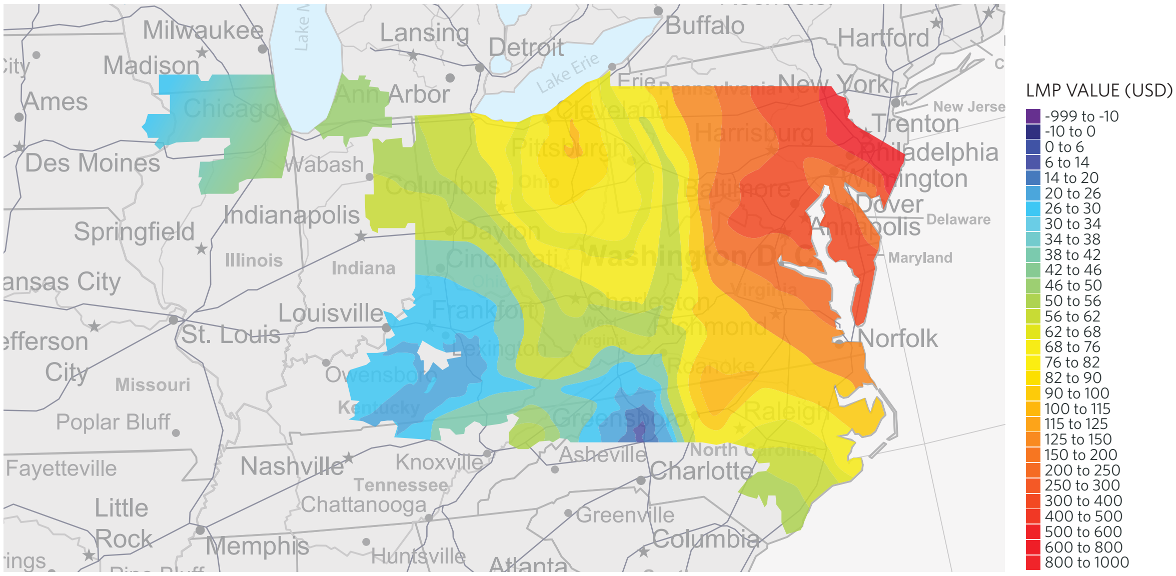
17 Power quality involves the minimization or elimination of subcycle fluctuations or deviations from the theoretically optimal sinusoidal voltage or current waveforms.

18 The value of firm capacity may differ by location when long-term, systematic network constraints exist.

19 The value of operating reserves may differ by location when long-term, systematic network constraints exist.

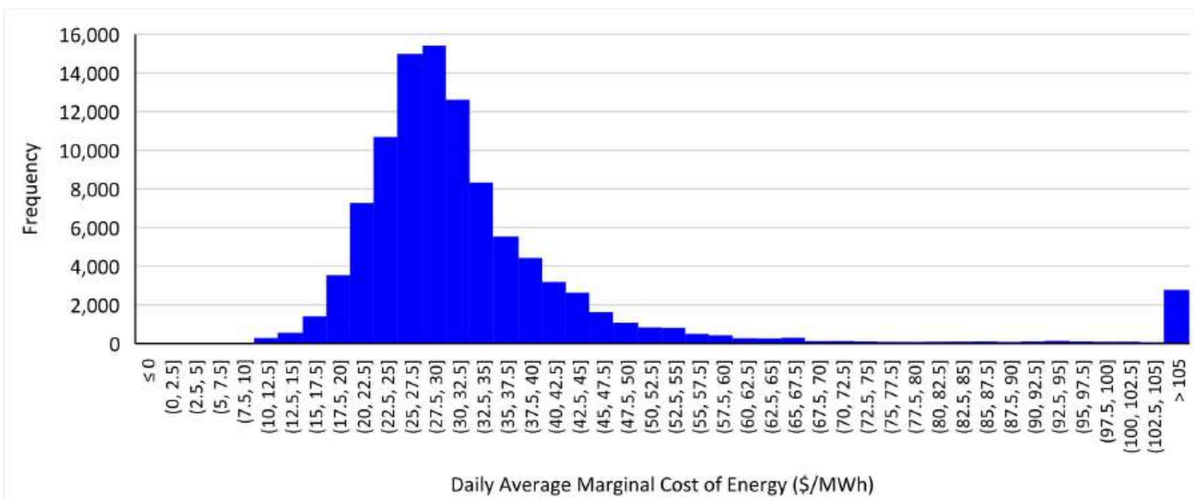
20 "Premium value" is a catch-all that refers to the value that DER owners may derive from factors that are inherent to DERs but that are not directly associated with the electricity services that DERs provide. For example, some consumers may derive value from independently producing their own power or aligning their electricity production with personal values (such as environmental concerns). When such "premium values" are private—i.e., they accrue only to the individual(s) who owns the DER(s)—these premium values should not be internalized by public policies.

Figure 2.4: Contour Map of PJM Nodal Prices on July 19, 2013, at 4:05 p.m. ET



Source: PJM

Figure 2.5: Daily Average Marginal Cost of Energy at All PJM Aggregate Nodes in 2015



Of course, not all services change value based on their location within the service network. For example, the value of mitigating carbon dioxide (CO_2) emissions is the same wherever the CO_2 is emitted in the world, since CO_2 mixes uniformly in the atmosphere over time and each ton of CO_2 contributes equally to global warming. Similarly, operating reserves are deployed to contain frequency deviations that emerge as a result of normal or contingency events. Since, except at very short time scales, frequency is consistent across an entire synchronized interconnected system, the value of controlling that frequency does not change based on the location of frequency regulation within the system.

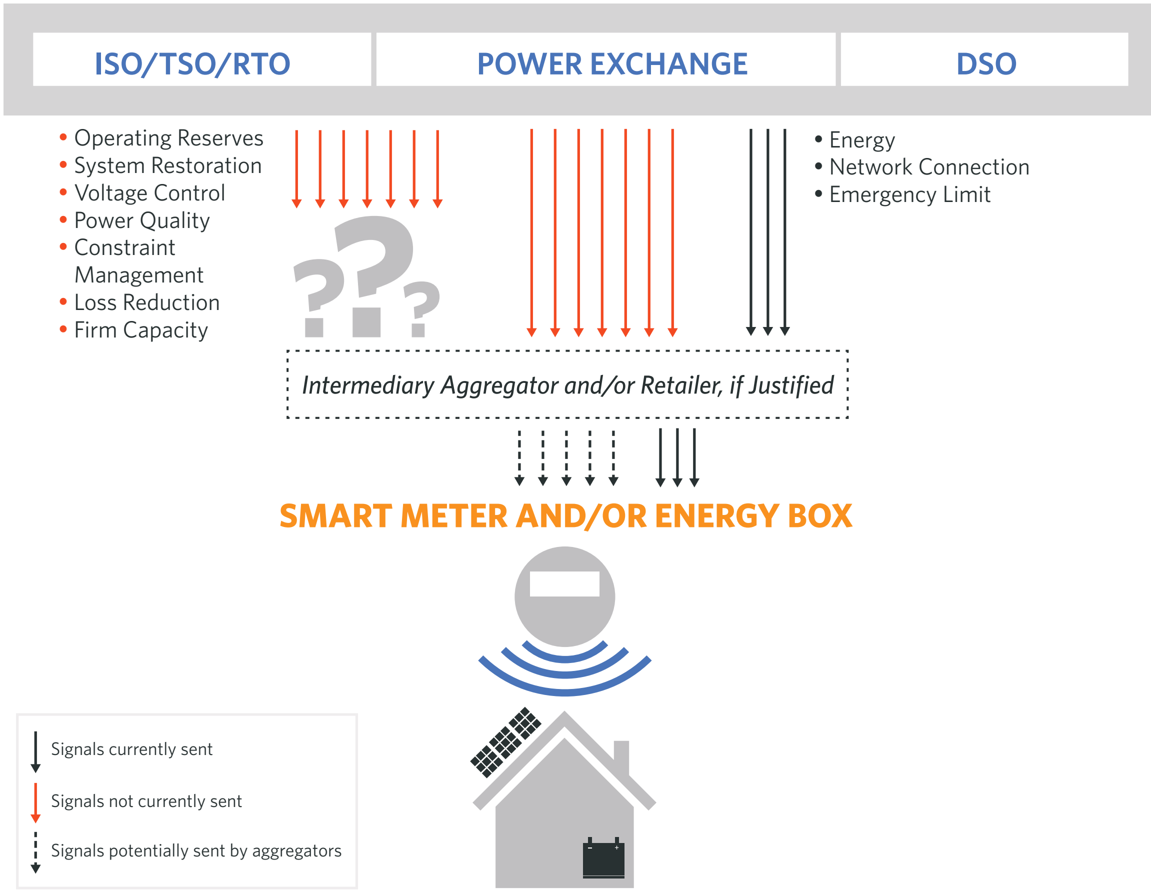
DERs can create value by providing locational services where centralized resources cannot or when the locational value of a service outweighs any added costs associated with providing a service in a distributed fashion (we explore this concept in **Chapter 8**). A large combined-cycle natural gas turbine may not be able to alleviate the strain on an overloaded transformer serving a growing neighborhood in a densely populated city, but demand response in the neighborhood or commercial-scale energy storage could. However, not all values are locational, and the distributed implementation of a resource that can also be deployed at scale (e.g., distributed solar or storage versus centralized solar or storage) should only be considered preferentially when the value of services provided justifies any additional costs due to not fully exhausting economies of unit scale (discussed in Chapter 8).

2.3.2 DERs can be aggregated

Another key distinguishing feature of DERs is that their small scale creates new opportunities for aggregators in power systems. This study adopts a slightly modified version of the definition of an aggregator promulgated by Ikäheimo, Evens, and Kärkkäinen (2010). In the context of this study, "an aggregator is a company that acts as an intermediary between electricity end-users and DER owners and the power system participants who wish to serve these end-users or exploit the services provided by these DERs." Retailers are therefore a special class of aggregators that (historically) served only the function of aggregating small electricity consumers — residences and small commercial entities — and procuring power on their behalf.

Today, aggregators are performing many new functions such as bidding the supply of aggregated DERs into wholesale markets, using loads to provide ancillary services, and more. In many cases, these aggregators are emerging alongside traditional retail aggregators, contracting with network users already under contract with retailers, and creating coordination challenges. Aggregators are enabling more price signals to be sent to agents, allowing new agents to become active in the operation and planning of the power system, and engaging agents in novel ways. **Figure 2.6** portrays the potential role of aggregators as intermediaries that enable more granular price signals to be sent between different power system actors (e.g., the distribution system operator or the transmission system operator) and the agents they serve. Aggregation should not be accepted as valuable without critical assessment. For example, aggregation across locations with different prices could present significant operational challenges. This study examines what value aggregators create, which leads to a better understanding of which agents should be allowed to perform aggregator functions (see **Chapter 6**). This requires exploring the different ways that aggregators can create value (or disvalue) in power systems, building on the classification of values highlighted in **Section 2.3**. This discussion will attempt to lend clarity to pertinent questions related to the role of aggregators such as: Should the power system accommodate many aggregators or only one centralized aggregator? Who can or should be an aggregator (transmission and distribution system operators, retailers, third parties, etc.)? What market design elements may need to be adapted or adopted to accommodate DER aggregators? How should new third-party aggregators be coordinated with existing power system aggregators (e.g., retailers)?

Figure 2.6: Current Power System Information Gaps and the Potential Role of an Aggregator[21]



2.3.3 DERs give rise to novel business models

The final key distinguishing feature of DERs explored in this study is the way in which DERs give rise to new business models for the competitive provision of electricity services. The emergence of DERs is already driving changes to the business models (i.e., the revenue streams, cost structures, customer segments, value propositions, etc.[22]) of regulated distribution utilities, as will be discussed in **Chapter 5**. DERs are also giving rise to entirely new ways of competitively providing electricity services.

In **Chapter 6**, we review the structure of nearly 150 business models engaged in deploying solar PV and energy storage and unlocking demand response. Analyzing these business models further reinforces the importance of properly structuring the electricity industry and markets, in order to: 1) minimize the prevalence of businesses that are built on the creation of opportunistic private value at the expense of system value, and 2) create a level playing field for competition among the many business models capable of providing the range of electricity services desired by markets, utilities, and end consumers. Chapter 6 and **Appendix B** explore these concepts in more detail.

21 Note that a retailer can perform the function of an aggregator, in which case price signals would go directly from system operators to the retailer.

22 This study adopts the definition of a business model put forth by Osterwalder and Pigneur (2010). According to Osterwalder and Pigneur, a business model is comprised of: a value proposition, key partners, customer segments, a cost structure, revenue streams, key activities, key resources, customer relationships, and channels.

References

ACER. 2013. "Capacity Remuneration Mechanisms and the Internal Market for Electricity." Ljubljana, Slovenia: Agency for the Cooperation of Energy Regulators.

Ackermann, T., G. Andersson, and L. Söder. 2001. "Distributed Generation: A Definition." *Electric Power Systems Research* 57: 195–204. doi:10.1016/S0378-7796(01)00101-8.

Bohn, R.E., M.C. Caramanis, and F.C. Schweppe. 1984. "Optimal Pricing in Electrical Networks over Space and Time." *The RAND Journal of Economics* 15: 360–376.

Borlick, R.L. 2010. "Pricing Negawatts." *Public Utilities Fortnightly* 14–19.

DNV GL. 2014. "A Review of Distributed Energy Resources." Arlington, VA: DNV GL.

Ela, E., M.R. Milligan, A. Bloom, A. Botterud, A. Townsend, and T. Levin. 2014. *Evolution of Wholesale Electricity Market Design with Increasing Levels of Renewable Generation*. Golden, CO: National Renewable Energy Laboratory.

ENTSO-E. 2013. "Network Code on Load — Frequency Control and Reserves." European Network of Transmission System Operators.

Gribik, P.R., W.W. Hogan, and S.L. Pope. 2007. *Market-Clearing Electricity Prices and Energy Uplift*. Cambridge, MA: Harvard Electricity Policy Group.

Hogan, W., R.D. Tabors, R.D., and M. Caramanis. 2012. "FERC Order 745 Amici Curae on Demand Response Payments."

Hogan, W.W. 2015. "Electricity Market Design: Energy and Capacity Markets and Resource Adequacy." *Capacity Markets: Gauging Their Real Impact on Resource Development and Reliability*. Cambridge, MA: EUCI Conference.

Ikäheimo, J., C. Evens, and S. Kärkkäinen. 2010. DER *Aggregator Business: The Finnish Case*.

Lazard. 2015. "Lazard's Levelized Cost of Storage Analysis — Version 1.0." New York, NY.

NARUC. 2016. Draft NARUC *Manual on Distributed Energy Resources Compensation*. Washington, DC: National Association of Regulatory Utility Commissioners.

NREL. PV Watts Calculator. National Renewable Energy Laboratory. pvwatts.nrel.gov/.

Newell, S., K. Spees, J. Pfeifenberger, R. Mudge, M. DeLucia, and R. Carlton. 2012. E*RCOT Investment Incentives and Resource Adequacy*. Cambridge, MA: The Brattle Group.

O'Neill, R.P., P.M. Sotkiewicz, B.F. Hobbs, M.H. Rothkopf, and W.R. Stewart. 2005. "Efficient Market-Clearing Prices in Markets with Nonconvexities." *European Journal of Operational Research* 164: 269–285. doi:10.1016/j.ejor.2003.12.011.

Osterwalder, A., and Y. Pigneur. 2010. *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. John Wiley & Sons.

Pepermans, G., J. Driesen, D. Haeseldonckx, R. Belmans, and W. D'haeseleer. 2005. "Distributed Generation: Definition, Benefits, and Issues." *Energy Policy* 33: 787–798. doi:10.1016/j.enpol.2003.10.004.

Pérez-Arriaga, I. 2011. "Managing Large Scale Penetration of Intermittent Renewables." MIT Energy Initiative Symposium. Cambridge, MA: Massachusetts Institute of Technology 1–43.

Pérez-Arriaga, I.J. 2013. *Regulation of the Power Sector, Power Systems*. London, United Kingdom: Springer. doi:10.1007/978-1-4471-5034-3.

Pérez-Arriaga, I.J., F. Rubio-Odériz, J.F.P. Gutiérrez, J.A. Ogando, and J. Marín. 1995. "Marginal Pricing of Transmission Services: An Analysis of Cost Recovery." *IEEE Transactions on Power Systems* 10: 65–72.

PART 1: UNDERSTANDING ELECTRICITY SERVICES AND HOW DISTRIBUTED ENERGY RESOURCES AFFECT THE DESIGN AND OPERATION OF POWER SYSTEMS

03

Envisioning a Future with Distributed Energy Resources

As early as 1978, in an article titled "Power Systems 2000," the late MIT Professor Fred Schweppe anticipated that price-driven demand response by residential and commercial customers, when enabled by efficient price signals and communication technologies, could play a large role in future electricity systems. Nearly forty years later, the electric power sector is beginning to implement what Schweppe foresaw. But while Schweppe only considered demand response, this study contemplates a wide range of new technologies that will shape the electicity industry going forward, including distributed generation, distributed storage, and electric vehicles, as well as many diverse and sophisticated forms of demand response, metering, energy control devices, and communications and computational capabilities, both centralized and distributed. Schweppe's article offered a perspective that was different from the established perspective of the time. He envisioned the power system as a meeting point of supply and demand, where supply and demand play equally important roles and together produce a "homeostatic" equilibrium. The purpose of this chapter is to introduce the reader to the changes that a substantial presence of distributed resources in power systems could bring to the provision of electricity services. Some of these changes have already begun. This chapter consists of four parts. The first part presents a collection of perspectives about the future of the electric power sector that depart from traditional assumptions about the sector's course going forward. The changes discussed in this section may or may not end up having a substantial impact on the future power sector landscape, but they are conceivable. Each of the perspectives we discuss is associated with high penetration and participation of distributed energy resources, and each challenges traditional views about the role of supply and demand in electricity markets.

The second part of this chapter presents a collection of case studies that illustrate the capabilities of distributed energy resources (DERs) when they are able to play an active role in power system operations. The case studies show how DERs can change the power sector;[1] they also illustrate how power systems can function when DERs are present in large quantities and point to the new

1 This change, to the extent that it may happen, could take diverse forms in different power systems, depending on idiosyncratic factors such as the pre-existing industrial base or existing technology-specific support schemes.

challenges this brings. The themes of the case studies vary: Some anticipate the active response of residential and commercial appliances to economic signals; others focus on the new roles, operational functions, and planning requirements of distribution companies and associated costs; still others highlight the challenges and importance of incorporating innovative technologies and processes at the distribution level. Other case studies yield insights into wholesale market prices, bulk power system operations, and the generation technology mix in systems with high DER penetration and explore new dimensions of the relationship between distribution and transmission system operators.

The third part of this chapter focuses on the implications of widespread DER adoption for cybersecurity protection and resilience. Cybersecurity threats and vulnerabilities require new policy approaches, while increased adoption of DERs and the rapidly expanding "Internet of Things" (IoT)[2] raise new challenges for privacy — an important concern for individuals and utilities.

Finally, the fourth part of the chapter presents a vision of how these technologies could transform power systems within a decade and beyond. The purpose of this exercise is not to predict the future, but rather to explore new possibilities and provoke broader thinking about the future of the electric power sector by considering novel arrangements that are not present in current systems.

3.1 Landscapes of Change

This section envisions a future in which DERs are widespread in electricity systems. First, however, we briefly consider the traditional, centralized paradigm from which this hypothetical future may evolve. Most electricity users were once connected, or are still connected, to vertically integrated power systems where a single centralized utility company made (or makes) all decisions concerning the operation of sytem assets and investment in energy supply. In recent decades, this centralized utility model has begun to fray, with the introduction (in many jurisdictions) of competitive generation by a more diverse set of actors. In such systems, companies that own large power plants compete to sell electricity through a number of mechanisms: through centralized trading platforms or bilateral, often long-term contracts; through spot market sales to large customers; or through retailers that buy wholesale power to resell to medium and small customers. At this point in time, many power systems have also introduced retail competition, whereby various intermediaries compete with one another to sell electricity to end customers. In these power systems, retailers must typically compete for the deregulated components of the electricity price against default suppliers with regulated tariffs, where those tariffs cover all the cost components of operating the power system and are established by governments or specialized regulatory agencies. In all cases, the secure operation of the bulk power system is supervised and managed by a system operator and the distribution company performs a similar role at lower voltage levels.

How might the widespread presence of DERs significantly change this paradigm?[3] Some changes could occur as mere continuations of the trends that are already being observed in pioneering systems that have a growing DER presence. In other cases, however, the changes we describe would require a radical deviation from the centralized paradigm that has prevailed for more than a century. Though only now beginning, these changes could become commonplace in ten years and could lead to a power system that Thomas Edison would not have recognized. What follows is a non-exhaustive inventory of hypothetical but plausible DER-related scenarios that challenge the reader to consider a power sector that is fundamentally different from the one we are familiar with today.

2 The "Internet of Things" refers to the interconnection of physical "smart" devices through external Internet and cloud-based control systems that enable these objects to collect and exchange data. For instance, electric cars in the future could exchange information that might be considered personal, such as detailed information about their location or times of day when they are using energy or being charged, etc.

3 Note that this report is agnostic with respect to the level of DER penetration that is likely to be achieved by any particular date, in any particular power system, or worldwide. But high penetration levels are already a reality in some power systems, as documented in **Chapter 1**.

3.1.1 Distributed provision of services

The multitude of devices that can inject, store, or withdraw power at their connection points — acting either independently or in some coordinated fashion — to provide the electricity services defined in the previous chapter represents a major departure from the conventional, centralized power system paradigm. Specifically, it implies that distributed providers of electricity services may be able to successfully compete and collaborate with centralized supply. Distributed resources may have advantages over centralized ones, and they may also have added (or different) limitations and costs. The conclusion is clear: First, there is a great need to identify and remove any inefficient barriers to DER participation as well as any ill-designed subsidies or supports for specific technologies that could distort investment decisions. Second, it will be important to create a comprehensive system of prices and charges that correctly conveys and monetizes the value of both centralized and distributed resources, all competing under the same rules. Only under these conditions will the full economic and technical potential of all resources be realized.

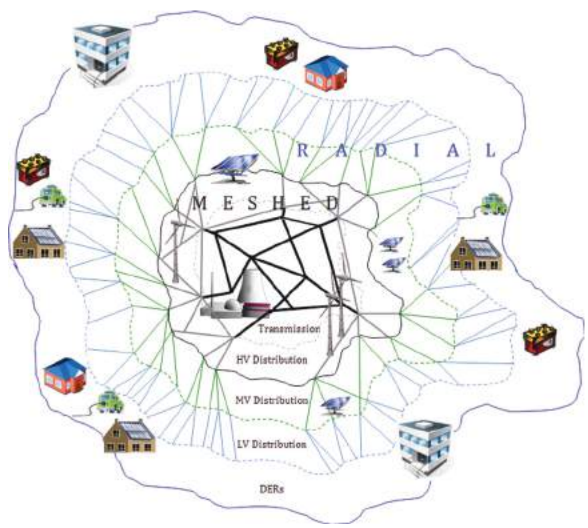
3.1.2 Diversity in patterns of power flow, control signals, and origin of services

Distributed generators connected to the distribution network at different voltage levels may not be well aligned with local demand. This implies that, at any given time, net power flows can occur in either direction (i.e., from the distribution network outward to the bulk power system, or vice versa). In some power systems, mostly in European countries, output from distributed generators sometimes surpasses that of centralized resources across large zones. Moreover, because of DER participation in these systems, the communications, control, and price signals that are required for service provision are also bidirectional. For this reason, and as noted in previous chapters, the "top down" paradigm[4] must be abandoned. **Figure 3.1** conveys a new perspective on the electric power system and its implications. It suggests that many aspects of the system must be reconsidered, including the roles of transmission and distribution system operators, the boundaries between distribution and transmission, procedures for allocating network costs, and technologies and processes for enabling cost effective DERs to provide electricity services and participate in markets that have traditionally been restricted to large actors.

4 By "top-down" paradigm we mean an industry structure in which power flows only in one direction: from large power plants, which provide all services, to passive end consumers, passing from the highest voltage levels to the lowest ones.

The significant presence of DERs — either actual or anticipated — necessitates a reconsideration of how DERs can most effectively provide electricity services. It implies that the electric power sector's customary "top-down" paradigm must be abandoned, and that a fresh look at the design, operation, and regulation of distribution and transmission networks is required.

Figure 3.1: Beyond the "Top-Down" Paradigm



3.1.3 The enabling role of information and communications technologies

Today, powerful and inexpensive information and communications technologies (ICT), with capabilities that could not have been imagined even a few years ago, allow agents in power systems to interact in new ways.

These new modes of interaction will be commonplace a decade from now in many parts of the world. We are not predicting the discovery or emergence of new technologies, which is always a possibility, but rather a transition toward the intelligent use of the full capabilities of existing technologies. Of course, as we explain in later chapters of this study (notably, **Chapters 5, 6,** and **7**), this transition depends on the removal of inefficient barriers that presently exist. Removing such barriers will allow for the efficient adoption of distributed technologies and associated business models; that is, new business models will emerge that create economic value and meet the preferences of agents.

Key technological developments will include the integration of new capabilities from the "Internet of Things," the widespread dissemination of two-way communication devices, and the ability to determine, with increasing granularity in time and space, the economic value of electricity services and appropriate charges for the use of electricity networks. These developments must coincide with the elimination of inefficient regulatory barriers and the adoption of market rules that allow any agent or device to participate efficiently in providing various electricity services with commercial value (again, see **Chapters 5, 6**, and **7**). A description of what the power system could look like in a not-so-distant future is presented later in this chapter.

3.1.4 The need for further restructuring of the power sector

One can expect that high penetration of DERs will be accompanied by the corresponding emergence of new service providers and a diversity of business models to capture the value of distributed resources. However, this is only possible in a power sector with a markedly different industry structure than the current one. DERs — acting autonomously or via aggregators — will be active participants in the operation of the power system. They will provide a variety of electricity services, which system operators at both the transmission and distribution levels will utilize alongside services provided by centralized resources. This will demand a much closer integration of the activities of distribution and transmission system operators. Finally, the role of distribution network owners and operators must adapt to a new and more dynamic context in which diverse new business models offer the full suite of electricity services. Regulators must make sure that the structure and allocation of responsibilities in the power sector is conducive to competition between various types of electricity services providers.

3.1.5 The rise of empowered consumers

DERs may give rise to a potentially revolutionary change in the relationship of network agents to energy users. Formerly passive customers will be able to match their utilization of energy with their preferences and priorities because, for the first time, they will have substantial choice. This is already normal for other types of goods. For instance, food is more than a collection of nutrients that people consume to stay healthy and function properly. On the contrary, consumers can choose among a diversity of food products and experiences. Food is associated with social activities, individual preferences for textures and flavors, status, time and effort devoted to the activities of cooking and eating, etc. The same can be said of other markets where end users have choices, such as markets for cars, personal computers, and mobile phones. Choice in the provision and production of electricity services may lead to a similar outcome in which functionality and cost are not the only factors that influence consumer behavior. Increasingly, electricity customers will be able to express their preferences and values, whether those include minimizing cost, reducing environmental impact, increasing energy autonomy and reducing reliance on the grid, avoiding use of the incumbent provider, etc.

Formerly passive "consumers" will be able to match their utilization of energy with their diverse individual preferences and values because they will have a meaningful level of choice.

This study focuses primarily on the competition between different types of energy resources in terms of cost and technical performance. Hence, we do not perform a detailed analysis of consumer preferences or values. Rather, our objective is to design a regulatory framework and a system of cost-reflective prices and charges that together create a level playing field in which all resources and all agents can combine economic signals with their personal preferences to arrive at their own decisions in terms of investment in, and utilization of, electricity services.

3.1.6 Grid defection, self-generation, and implications for tariff design

Grid defection — that is, the decision of a network agent to self-produce electricity without support from, or connection to, the interconnected power system — may be justified (or not) for economic or non-economic reasons. Among the economic reasons, the widespread availability of increasingly affordable distributed generation and storage resources is lowering the economic barrier to self-sufficiency in the provision of electricity services. In some cases, grid defection may even make economic sense today. Possible non-economic reasons for grid defection include the satisfaction associated with being energy and technology independent, or with relying on a system that is more environmentally friendly than the incumbent's.

For individual customers, the economic case for defecting from the grid is enhanced by the potential to avoid network charges and the costs of various electricity policies. It is an open question whether some of these policy costs — for example, subsidy support for renewables and/or energy efficiency programs, social tariffs, stranded utility costs, and others — should belong in the electricity tariff. Some jurisdictions include these costs in electricity tariffs while others recover these costs via taxes of various sorts. It is also an open question whether some kinds of policy costs should exist at all, such as subsidies to inefficient and carbon-emitting domestic fuels. Moreover, since grid defection only marginally reduces existing network and policy costs, the "left-over" network and policy costs that the grid-defector has avoided will be borne by the customers that remain connected to the network. This dynamic is similar to the monetary transfer that is currently occurring in power systems that use net metering (or similar types of policies) to encourage self-generation. In these systems, self-generation reduces payments to the network provider, but does not reduce existing network and policy costs. As a result, these costs fall on remaining customers that are not engaged in self-generation.

The regulated component of the electricity tariff has been often used for political reasons, in many cases to finance a variety of subsidies and other costs whose presence in the electricity tariff is questionable. The actual or anticipated emergence of significant levels of self-generation, and also perhaps of grid defection, demands urgent attention from governments and regulatory authorities. In particular, governments and regulatory authorities need to reconsider what should and should not be included in regulated electricity tariffs. This is yet another reason to proceed with a redesign of the current system.

3.2 Challenges and Opportunities for DER Integration

DER cost reductions will increase opportunities for end users to self-provide energy services at a competitive cost relative to what they would have paid to procure these services by other means. Furthermore, some consumers may become system service providers — that is, they may begin providing services to the system for use by other agents. However, DER integration also comes with challenges, which require a rethinking of traditional practices and assumptions about the functioning of the power system, as well as changes to established regulatory frameworks. This section discusses some of

> *DER integration requires a rethinking of traditional practices and assumptions about the functioning and regulation of power systems.*

the challenges and opportunities that DER integration poses for end users, distribution networks, and wholesale and retail markets. To illustrate these challenges and implications, we present a series of case studies within the context of US and European markets.

3.2.1 Customer adoption of DERs in residential and commercial buildings

Today, customers who adopt DERs are primarily motivated by the opportunity to reduce their energy bills. This opportunity exists as a result of policies and regulations. An increasing number of DER technologies, including solar photovoltaics (PV), batteries, home energy management systems, micro-combined-heat-and-power (CHP) systems, fuel cells, and others, are now cost-effective in certain applications. This presents many opportunities and challenges for providing the right mix of policy incentives. In addition, the emergence of distributed technologies has the potential to enhance the provision or consumption of electricity services.

This section uses case studies to illustrate how customers may opt to use different DERs depending on the financial benefits they provide — both by avoiding having to procure services from the system and by generating revenues in return for providing services to the system. The first case study involves reducing energy costs by managing building heating and cooling loads in response to hourly electricity prices. In this case study, buildings can also provide ancillary services to the system in competition with traditional providers, such as centralized generators. The second case study builds on the fact that gas and electricity are increasingly becoming substitute fuels for a variety of end-user services such as heating and cooling. It illustrates how various factors, including retail gas and electricity prices, climate conditions, and load profiles, determine the competitiveness of electricity and gas DER technologies. The third and final case study shows that DERs are in competition not only with centralized resources, but also with other DERs. In this case study we demonstrate how battery storage and demand flexibility (e.g., managing air conditioners and water heaters) can provide the same services to reduce electricity bills. The profitability of each technology will depend on the deployment of the other.

CASE STUDY 1: BUILDINGS-BASED DEMAND RESPONSE IN ANCILLARY SERVICES MARKETS

High electricity consumption, large thermal mass, and the ability to control heating, ventilation, and air conditioning (HVAC) loads make commercial buildings good candidates for providing demand response in energy, capacity, and ancillary services markets. In the United States, buildings account for 74 percent of total electricity use and nearly half of this consumption occurs in the commercial sector (EIA 2015). Energy use by HVAC systems accounts for the largest fraction of total commercial electricity consumption and is often the most variable category. Materials used in commercial building envelopes and structures can provide energy storage when strategically heated or cooled. Methods such as model predictive control (MPC) can be used to determine optimal cooling strategies that exploit thermal storage to minimize energy costs in addition to potentially providing other services, such as operating reserves. Increasing knowledge and interest in acquiring enabling technologies and software, combined with HVAC systems that can facilitate more flexibility in building operation, may prompt commercial building owners or operators to participate in ancillary services markets.

To explore this potential, we simulate a medium-sized office building, located in Boston, Massachusetts, that participates in ancillary services markets while optimizing hourly energy consumption to reduce total costs.[5] In this simulation, the building is subject to either wholesale locational marginal prices (LMPs), as established by the independent system operator for New England (ISO-NE), or flat retail electricity rates.[6] At the same time, the building can generate revenues from ISO-NE's day-ahead markets for frequency regulation and spinning reserves. **Table 3.1** displays results from two price scenarios for a typical July day: Scenario A features higher ancillary services prices throughout the day, relative to Scenario B. In addition, in each scenario, the medium-sized office building is subject to either hourly LMPs or a flat retail rate of 14.5 cents per kilowatt-hour (¢/kWh) for its energy purchases (reflecting average prices for July 2015). The office building is able to recover 122–141 percent of its HVAC energy costs when it pays the LMP, resulting in a negative net operating cost. If subject to the average retail rate for electricity, the office building saves 13–26 percent of electricity costs in both scenarios. In this case, the flat rate prevents smarter energy consumption that could enable the building to optimize its usage in response to the changing value/cost of electricity throughout the day.

5 Further description of the methodology and results of this case study can be found in a paper by Nora Xu titled "Describing commercial buildings' thermal response and optimal cooling strategy for provision of electricity services." Utility of the Future Memo Paper: energy.mit.edu/uof.

6 Note that office energy cost accounts only for electricity consumed at the wholesale LMP and unlike the retail rate, does not include any additional network or regulated costs.

Table 3.1: Monthly Weekday Cost and Revenue Estimates for Electricity Services Associated with a Medium-Sized Office Building in Boston, Massachusetts

	SCENARIO A LMP	SCENARIO A RETAIL RATE	SCENARIO B LMP	SCENARIO B RETAIL RATE
OPTIMAL ENERGY COST ($)	403	1,436	323	1,422
OPTIMAL REGULATION REVENUE ($)	503	352	369	178
OPTIMAL SPINNING RESERVES REVENUE ($)	68	27	26	0
REDUCTION IN ENERGY COST ($)	571	379	395	178
OPTIMAL NET OPERATING COST ($)	-168	1,057	-72	1,243

Note: The table shows monthly weekday scaled estimates for optimal energy cost, ancillary services revenue, and net operating cost for two July price scenarios in which a medium-sized office pays either the LMP or average retail rate for electricity.

Ancillary services revenue by itself may not be a sufficient economic motivator for owners of small office buildings to invest in certain HVAC systems and associated technology and control systems. However, facilities with metering and required building automation systems (BASs) might see additional benefits in other demand response or energy cost minimization programs.

Building loads, such as loads for heating and cooling, can be managed to reduce electricity bills by reacting to hourly energy prices and providing operating reserves.

As this example illustrates, the effect of tariff design is an important determinant of the quantity of ancillary services provided by office buildings. Under flat volumetric retail rates for electricity, medium-sized office buildings provide a lower quantity of ancillary services than under dynamic energy market prices (LMPs). This is because under a flat volumetric retail rate, the energy consumed by the building is valued at the retail tariff while the regulating energy provided to the wholesale market is valued at the lower wholesale LMP.

CASE STUDY 2: THE INTERPLAY BETWEEN GAS AND ELECTRICITY FOR SPACE CONDITIONING

Distributed energy resources for space conditioning comprise a set of varied technologies, ranging from mature well-established systems — such as furnaces, boilers, and air conditioning (AC) units — to emerging ones such as micro-CHPs, reversible heat pumps, and hybrid gas-electricity conditioning systems.

Commercial and residential customers, among others, consider the economics of these technologies when deciding to adopt one technology over another. Performance characteristics such as efficiency and heat-to-power ratio, as well as economic characteristics — like capital and operating costs and energy prices and their associated tariff structure — are expected to have a major impact, not only on the profitability of these technologies but also on how competitive they are with other technologies for meeting consumers' energy needs.

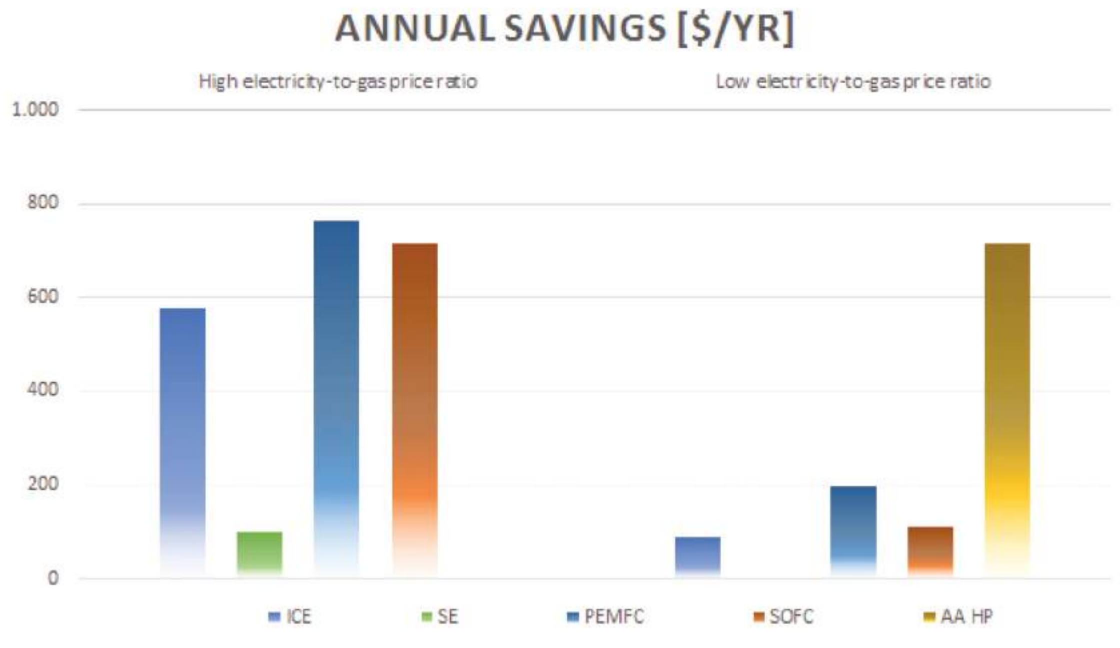
Figure 3.2 shows annual savings and internal rates of return (based on ojected technology costs[7]) for six different DER technologies used to provide space conditioning in a single-family house, under cold climatic conditions and under high and low electricity-to-gas price ratios.[8] The analysis considers five different micro-CHP technologies: internal combustion engine (ICE), Stirling engine (SE), two kinds of fuel cells — polymer electrolyte membrane fuel cell (PEMFC) and solid oxide fuel cell (SOFC) — and an air-to-air heat pump (AA HP). Annual cost savings for each of these technologies are calculated with respect to a reference case based on a high-efficiency gas-fired condensing boiler. We observe that energy prices have a large impact on potential energy savings and the upfront costs of the various technologies significantly determine their economic viability. The trade-off between electricity costs and fuel costs is important, as systems with a high electricity-to-gas price ratio are clearly favorable to micro-CHPs, whereas electric heat pumps are economically favorable in markets with lower electricity-to-gas price ratios. The profitability of DER technologies also depends on climatic conditions, with cold climates favoring cogeneration systems. Highly efficient gas and electric heat pumps are plausible alternatives for reducing primary energy consumption and energy costs.

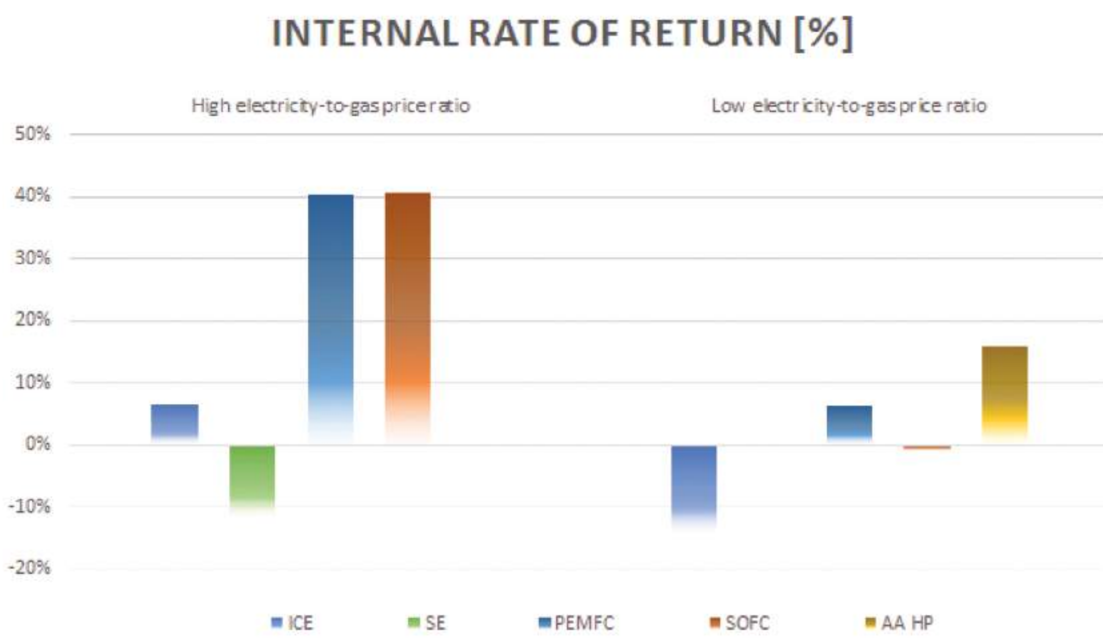
Retail gas and electricity prices, climate conditions, and load profiles are key determinants of the competitiveness of gas- and electricity-fueled DERs that can provide the same end-user services.

7 For this analysis we use projected costs from various US agencies. Projected total costs for heat pumps are based on the US Energy Information Administration (EIA) report, Updated Buildings Sector Appliance and Equipment Costs and Efficiencies, April 2015 (available at: www.eia.gov/analysis/studies/buildings/equipcosts/pdf/full.pdf). Projected costs for residential combined heat and power engines are based on ARPA-E's "Generators for Small Electrical and Thermal Systems - GENSETS Program Overview" (available at: arpa-e.energy.gov/sites/default/files/documents/files/GENSETS_ProgramOverview.pdf). Finally, projected costs for residential CHP fuel cell systems come from the US Department of Energy's "Fuel Cell Technologies Office Multi-Year Research, Development, and Demonstration Plan - 3.4 Fuel Cells," Updated November 2014 (available at energy.gov/sites/prod/files/2014/12/f19/fcto_myrdd_fuel_cells.pdf).

8 Further description of the methodology and results can be found in Dueñas, P. and K. Tapia-Ahumada, "Interplay of Gas and Electricity Systems at Distribution Level," Utility of the Future Memo Paper at: energy.mit.edu/uof.

Figure 3.2: Annual Savings and Expected Internal Rates of Return for Different DERs Under Cold Climatic Conditions





Note: Annual savings and expected internal rates of return are shown for two electricity-to-gas price ratio scenarios (we assume high and low price ratios of 3.6 and 1.9, respectively) and are calculated with respect to a reference case based on a high-efficiency gas-fired condensing boiler. Calculations consider technology cost projections,[10] not current costs. ICE — internal combustion engine; SE — Stirling engine; PEMFC — polymer electrolyte membrane fuel cell; SOFC — solid oxide fuel cell; and AA HP — air-to-air heat pump.

However, residential micro-CHP systems continue to face very high upfront costs.[9] Although reciprocating engines are a mature and established technology for large and medium applications, the cost of this technology remains high for smaller applications. Moreover, the use of micro-CHP systems gives rise to some environmental concerns. Given their high electrical efficiencies and low primary energy consumption, fuel-cell-based systems are promising. However, their high equipment costs continue to be a barrier to further deployment. Cost reductions are crucial to make these technologies competitive in the near future.

CASE STUDY 3: COMPETITION BETWEEN BATTERY STORAGE AND FLEXIBLE DEMAND

This case study illustrates how two of the most prominent DERs — demand flexibility and battery energy storage — compete with each other at the residential scale.[10]

We simulate the adoption of DER technologies in response to electricity tariffs, climate conditions, technology cost, and performance parameters, in the case of a single-family household. The key inputs for this simulation are the extent of flexible demand and the upfront cost of batteries.

Figures 3.3 and **3.4** show the impact of different levels of demand flexibility and upfront battery costs on the profitability — expressed as internal rate of return (IRR) — of the batteries.[11] Demand flexibility is increased when the temperature setting of cooling and heating loads is managed to modulate energy consumption in particular periods by expanding the comfort temperature dead-band. In addition, demand flexibility is enlarged by engaging more energy end-uses, such as air conditioners (AC) and water heaters (WH). **Figures 3.3** and **3.4** consider five stages of flexibility: no flexibility, air conditioning dead-band temperatures of 1 °C and 2 °C, and air conditioning and water heating dead-band temperatures of 2 °C.

In both cases, demand flexibility has a significant negative impact on the profitability of batteries. As the amount of flexibility increases, the cost of batteries must significantly decline for batteries to be profitable. These simulations show that demand flexibility has the potential to diminish battery revenue streams.

A noteworthy difference between the cases is the availability of the thermal resource. The hotter climate in Texas leads to greater use of air conditioning, which in turn means greater potential for smart energy management strategies to reduce customer bills. Without flexible demand, batteries are profitable at a higher upfront cost in Texas but, as in New York, any amount of flexibility quickly diminishes the revenue opportunities. Therefore, the size of the thermal resource can significantly affect the outcome.

DER technologies may compete to provide the same electricity services. The profitability of one technology (e.g., storage) strongly depends on the deployment of others (e.g., demand response).

9 Refer, for example, to Hawkes, A., E. Entchev, and P. Tzscheutschler. 2014. "Impact of Support Mechanisms on Microgeneration Performance in OECD Countries." On behalf of International Energy Agency — Energy in Buildings and Communities Programme (IEA- EBC Annex 54) published by Technische Universität München, Germany, October 2014. In this reference, the technology cost of an internal combustion engine (ICE) of 1.2kWe is around $8,000, the cost of a Stirling engine (SE) of 1.0kWe is approximately $6,500, the cost of a polymer electrolyte membrane fuel cell (PEMFC) of 0.75kWe is over $11,500, and the cost of a solid oxide fuel cell (SOFC) of 0.70kWe is above $15,000.

10 Assumed technology costs are as follows: IC-$4,500/kWe; SE-$4,500/kWe; PEMFC-$2,500/kWe; SOFC-$2,500/kWe; AA H-$4,000/kWe.

11 Further description of the methodology and results can be found in Huntington, S. "Case study: Battery Storage vs Flexible Demand." Utility of the Future Memo Paper at: energy.mit.edu/uof.

Figure 3.3: The Impact of Flexible Demand and Battery Cost on Battery Profitability in New York

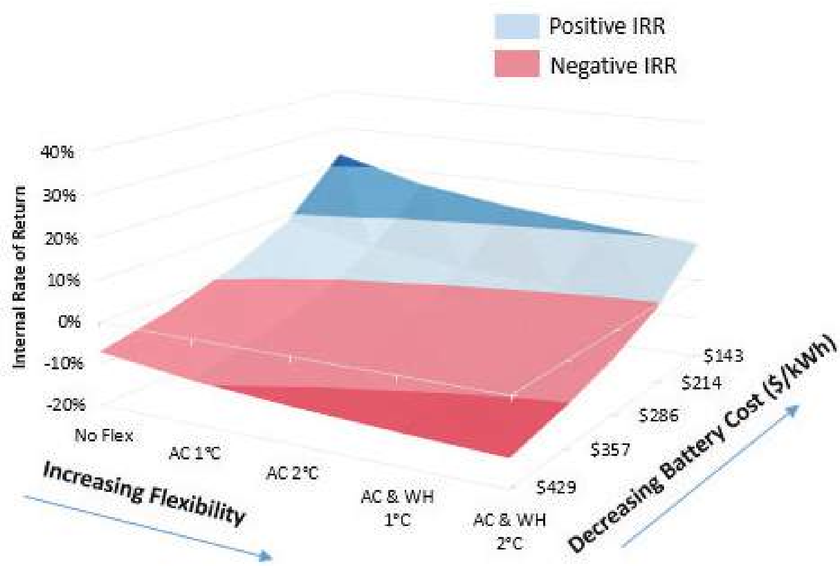
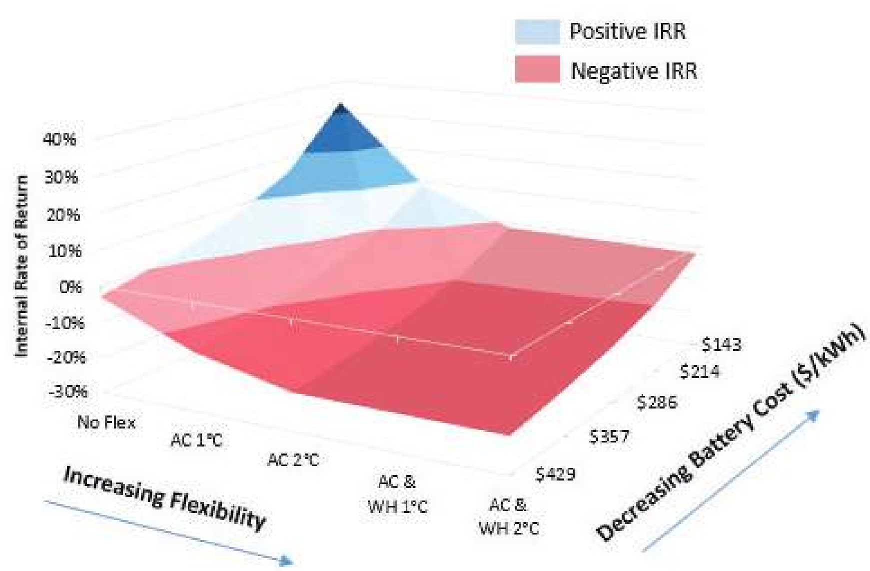


Figure 3.4: The Impact of Flexible Demand and Battery Cost on Battery Profitability in Texas



3.2.2 DERs at the distribution network level

Significant deployment of DERs requires revisiting conventional approaches to the planning and operation of distribution networks. It also requires defining new roles and functions for distribution companies. In this section, we illustrate how passive grid operation and a fit-and-forget approach to connecting new network users is bound to lead to significant cost increases in the face of significant DER deployment. To minimize costs over the long term, regulators will need to incentivize distribution companies to develop innovative network management approaches and to migrate to active network management approaches in which locational network services can be provided by DERs.

In addition, this section discusses potential new roles for distribution companies as system operators and neutral market facilitators. Some of these new roles may include managing data to help facilitate retail competition, providing local and upstream markets and services with non-discriminatory access to DERs by facilitating interaction between DERs and other stakeholders (e.g., market operators, transmission or independent system operators, suppliers, and aggregators), and acting as system operators to acquire DER flexibility services. Finally, distribution companies may have a role in deploying innovative technologies such as distributed storage, electric vehicle (EV) recharging infrastructure, and smart metering. However, distribution companies should perform this latter role only if adequate rules are in place to prevent undue market dominance and exploitation of market power. In particular, distribution companies, which are regulated players, should not compete with deregulated players to provide the same services. Though briefly covered here, all of these topics are discussed in much greater detail in **Chapter 6**.

Fit-and-forget versus active network management

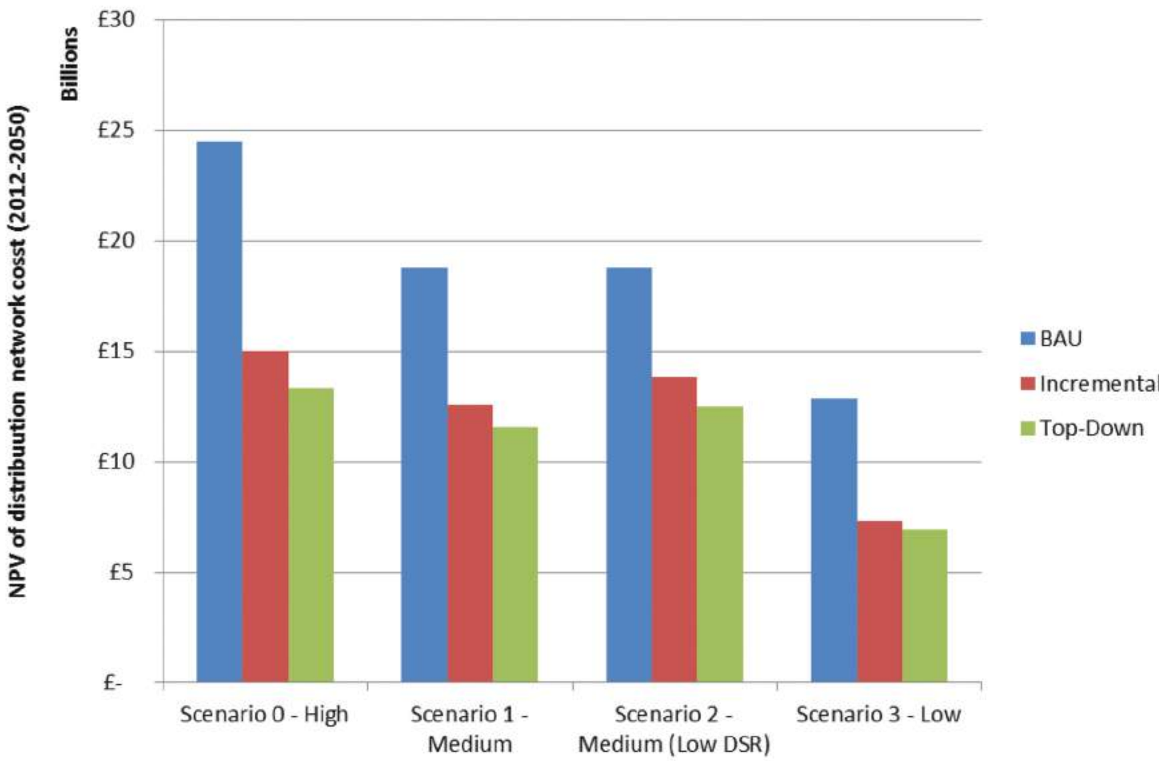
Distribution network planning and operation have historically been carried out as two nearly fully independent tasks. Long-term network planning typically entailed forecasting regional and local peak demands over a planning horizon and reinforcing the grid accordingly. The main goal of this process was to ensure that no physical constraints were violated during the real-time operation of the system. In this paradigm, the grid could be passively operated with very low levels of monitoring and control. Demand could be assumed to be unresponsive and inelastic in operational time scales, and the network included no distributed resources capable of efficiently meeting operational constraints.

The connection of DERs has so far been managed under this same paradigm: The network is reinforced whenever existing grid capacity is insufficient to ensure that even the most extreme conditions can be accommodated. This network management paradigm is often referred to as fit-and-forget and it has proven to be relatively effective and cost-efficient in the context of a conventional centralized power system. However, the deployment of DERs raises questions about the suitability and efficiency of the fit and forget approach.

For example, a recent analysis conducted by the UK Smart Grid Forum (with participation from public authorities and industrial and other stakeholders) estimated the levels of distribution network investment that would be required to accommodate DERs in different scenarios for the year 2050 (EA Technology 2012). A business-as-usual (BAU) scenario, where only conventional "wires" investments are made, was compared to two "smart" distribution grid investment strategies:[12] (1) a top-down strategy, where there is an upfront investment in smart grid technologies and further investment follows as and when networks reach their capacity limits, and (2) an incremental strategy, where investment only occurs as and when networks reach their capacity limits. **Figure 3.5** shows that the implementation of smarter distribution grid solutions that facilitate active network management leads to much lower costs compared to the conventional fit-and-forget BAU paradigm.

12 Among the most attractive smart grid solutions found in the UK Smart Grid Forum study are dynamic network reconfiguration, distribution FACTS, demand side response (DSR), local smart EV charging infrastructure, and permanent meshing of networks (EA Technology 2012).

Figure 3.5: Network Investment to Connect Low-Carbon Technologies Under Different Network Management Paradigms



Note: Levels of network investment are shown for connecting different penetrations of low-carbon technologies to the British distribution system by 2050. Source: EA Technology (2012)

The UK analysis illustrates how passive grid operation and a fit-and-forget approach to network planning lead to unnecessarily large costs, particularly in systems with high DER penetration. In such systems, costly grid reinforcements will be triggered by situations that occur relatively infrequently (e.g., a few times per year) such as very high injections or withdrawals from solar DERs. In these cases, active management or curtailment over short periods of time can often reduce the costs associated with such peaks. Moreover, the need for grid reinforcements can result in long lead times for connecting new network users. To avoid such inefficiencies and to facilitate an efficient transition to a distributed system, distribution companies will need to adopt innovative ways to manage their networks.

This necessarily implies coupling network planning and operation so that network constraints are managed not only at the planning and connection stages but also during real-time operation. Active network management in turn relies on the extensive utilization of ICTs to enable the advanced monitoring and control capabilities of grid assets and network users. One of the greatest challenges for regulating distribution companies in the future will be managing their close interaction with DERs that they do not directly control, but that assist in the operation of the distribution grid. The flexibility associated with DERs may become essential for day-to-day network operation. Distribution companies will have to become true "system operators" in the sense that they acquire network services, such as forward commitments of DER capacity to stand by to meet voltage control or congestion management requirements — just as various resources commit capacity to bulk power system operators for different classes of operating reserves.

Early examples of this transformation can be seen in certain countries. For example, since 2012 German distribution companies have had the capability to remotely limit power injections from certain PV installations (specifically, installations with capacity above 30 kilowatts) to mitigate local network constraints in exchange for financial compensation. Smaller PV units can either follow the same instructions as larger installations or permanently limit their injections to 70 percent of their rated generation capacity.[13] Regulators in some other jurisdictions are proposing similar changes to planning and operation practices. In the UK, for example, distribution companies are required to justify their long-term investment plans based on benefit–cost analyses that include innovative grid solutions (OFGEM 2013).

13 These parameters are set forth in the German Renewable Energy Sources Act or EEG.

Box 3.1: The Impact of Solar PV Generation on Distribution Network Costs and the Benefits of Distributed Storage

This box illustrates how high penetrations of rooftop solar PV increase distribution network expansion costs and how distributed storage may offer solutions to mitigate these incremental costs in the future.[14] Our example considers a distribution network in which the installed capacity of rooftop PV (and therefore total solar generation) is expected to increase over a certain period. For context, we assume that this growth is driven by subsidies, such as net metering or a feed-in tariff.

Following the aforementioned fit-and-forget approach, a distribution company could respond to increasing PV penetration by increasing the capacity of the wires, installing additional voltage regulators, and investing in other network equipment, with consequent increases in total network cost.

As shown in **Figure 3.6**, in some cases the total economic impact of meeting a significant (greater than 20 percent) fraction of load with distributed PV generation can be large — in the most extreme case, even doubling the cost of the network.

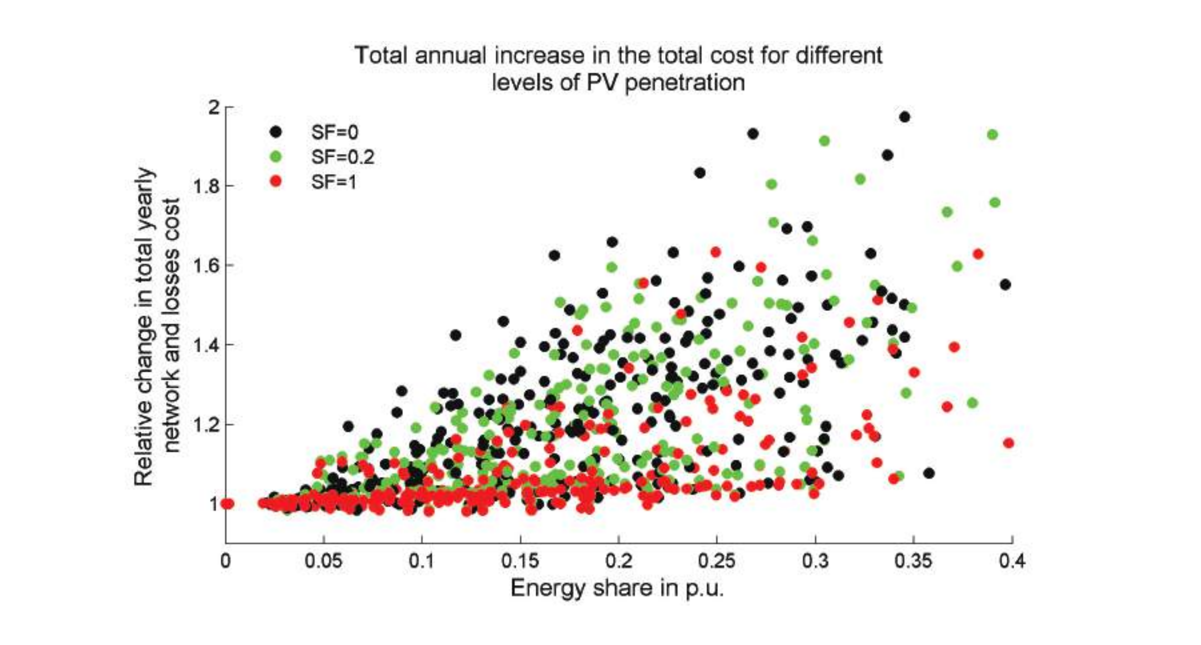
This raises important questions. How should these costs be allocated? If the network users that generate the costs are required to pay for them, would they still decide to invest in PV generators? Are there “non-wires” solutions that can lower total system costs?

14 For more details about the methodology used, we suggest reading Chapter 7 of MIT’s *Future of Solar Energy* study.

Figure 3.6 also shows that the installation of behind-the-meter distributed storage mitigates the impact of solar PV and enables large quantities of PV capacity to be connected to the grid without significantly increasing (or in some cases reducing) network costs. The level of penetration of distributed storage is modeled using a "storage factor" (SF) parameter, which is proportional to the reduction in maximum net PV injection to the grid obtained by adding storage.

Despite this cost-mitigating effect, procuring distributed storage is not necessarily the most cost-effective approach to manage the network impacts of distributed solar, for two reasons. First, it is not clear whether storage is cheaper than conventional network solutions (at current battery prices, storage is almost certainly not cheaper than conventional network solutions in the majority of cases). Second, there could be other, more cost-effective solutions such as PV curtailment or demand response. As we propose in this study, a cost-reflective system of prices and charges would reveal the least-cost solutions to meeting network challenges.

Figure 3.6: Impact of Distributed Solar PV on Network Cost with Different Levels of Energy Storage



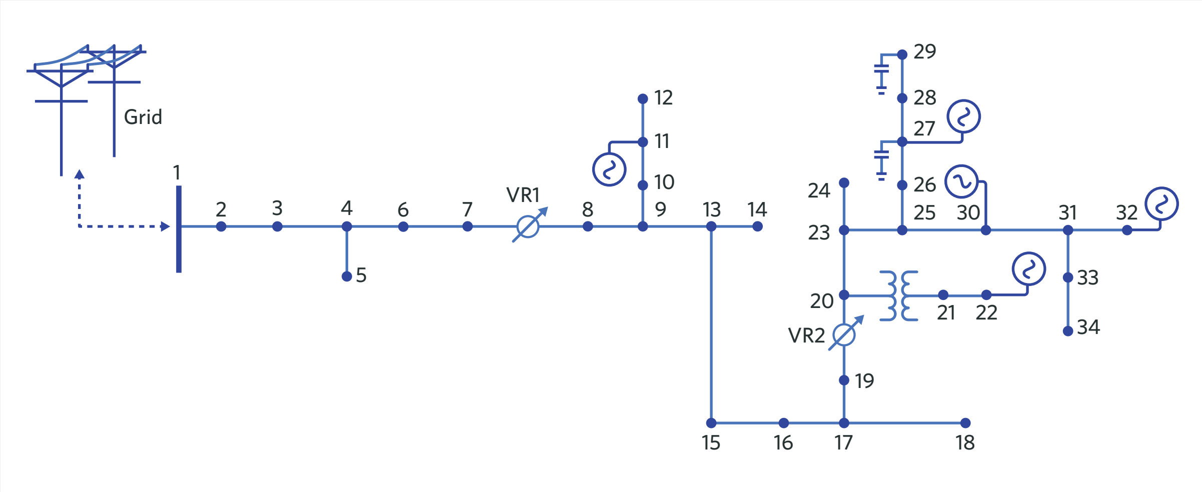
Note: The chart shows the cost-mitigating effect of energy storage (represented by a storage factor, SF). Source: MIT (2015)

Box 3.2: PV Inverters with Appropriate Control Technologies Can Enhance Distribution Network Capabilities

Combining distributed generation, such as solar PV, with flexible volt-var inverter control strategies can increase distribution network capabilities. The distribution system operator would optimize reference set points for local PV inverter controllers in order to minimize energy losses and maintain voltages within operational limits along the feeders.

PV inverters are a class of technologies that provide voltage support by changing the reactive power injection and absorption levels depending on the active power output of the solar PV system. When coupled with PV inverter control strategies, line losses associated with solar PV generation could be significantly reduced, while line transfer capacities and the loadability of networks (that is, the maximum load that can be supplied at peak hours) could be significantly increased.[15]

Figure 3.7: A 34-Node IEEE Test System (Feeder) with Five PV Generators



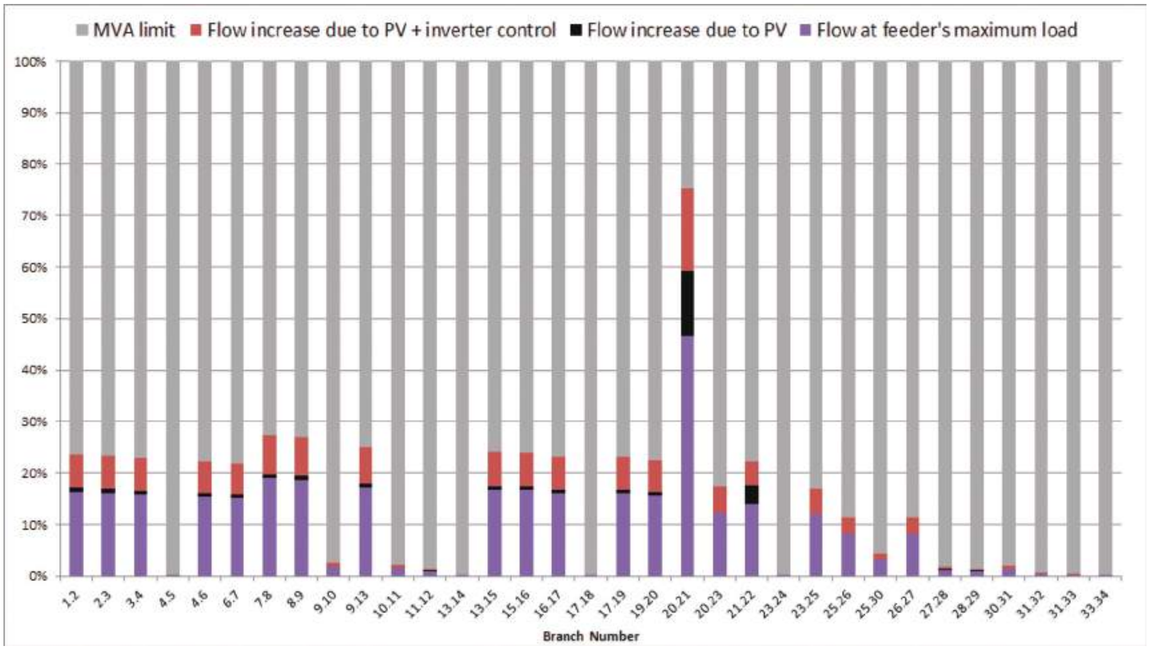
Note: The installed PV capacity at each node containing a generator is the peak demand at that same node.

Figure 3.7 shows a 34-node IEEE test system or feeder with five PV generators at different nodes. The maximum transfer capability of each branch or line at the maximum feeder load without PV is limited by voltage constraints in the feeder.

15 For more details about the methodology used, please refer to Abdelmotteleb I., T. Gómez, and J.P. Chaves. 2016. "Benefits of PV Inverter Volt-Var Control on Distribution Network Operation." IIT-Comillas Working Paper. www.iit.comillas.edu/html20/publicacion/mostrar_publicacion_working_paper.php.en?id=284.

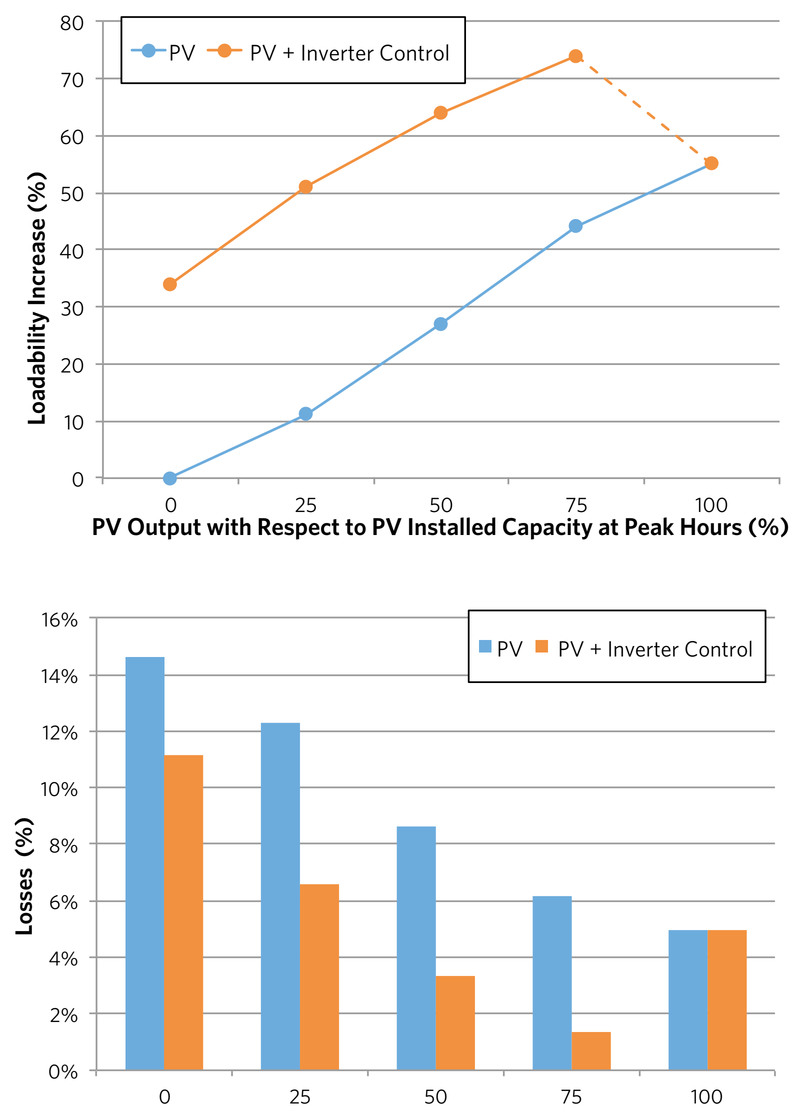
As illustrated in **Figure 3.8**, all the branches are significantly below their thermal limits (100 percent). The integration of PV generators with volt-var control strategies improves transfer capabilities in some of the feeder branches.

Figure 3.8: Variation in the Transfer Capabilities of Feeder Branches with PV Integration



Another characteristic of solar PV is its ability to help meet demand depending on solar output at peak hours. The maximum load that can be supplied at peak hours is known as feeder "loadability." Increases in feeder loadability can reduce the need for network reinforcements. As shown in **Figure 3.9a**, loadability is highest when solar PV is coupled with volt-var inverter control strategies. Moreover, when coupled with such strategies, network losses associated with solar PV can be significantly decreased (**Figure 3.9b**).

Figure 3.9: Network Loadability and Losses with Solar PV



Note: In **Figure 3.9a** (top) network loadability is measured as the percent demand increase that can be supplied at peak hours before the feeder reaches its transfer capacity. In **Figure 3.9b** (bottom) network losses are shown as a percent of demand, and at different levels of PV generation with and without PV volt-var inverter control strategies.

In sum, the integration of distributed generation (PV in this example) with volt-var inverter control strategies, in distribution networks with transfer capabilities that are mainly limited by voltage constraints (such as networks in rural areas), can improve system efficiency by reducing energy line losses and increasing network loadability, thereby potentially deferring the need for network reinforcements.

New roles of distribution utilities

The previous section points to the importance of considering DER flexibility when conducting network planning and operation. As the integration of DER capabilities becomes more commonplace, the system operator role of distribution utilities will become increasingly complex. New responsibilities may include managing end-user data and enabling or deploying innovative technologies, such as advanced metering, distributed storage systems, and EV charging infrastructure. (As we discuss in greater detail in **Chapters 5** and **6**, these new roles may be provided by distribution companies or by other market agents or entities.) Regulations must also be adapted to eliminate barriers to the provision of electricity services and to allow DERs to access local and upstream markets and services, thereby facilitating DER interaction with market operators, transmission or independent system operators, suppliers, and aggregators.

The interaction between distribution network utilities and network users has traditionally been limited to the grid connection process, phone calls during supply interruptions, and — depending on specific regulatory requirements — metering and billing processes. For reasons discussed in previous sections, however, the implementation of active network management strategies entails utilizing DER flexibility in network operations and planning.

Passive grid operation and a fit-and-forget approach to network management are outdated; more active network management strategies need to be implemented to reduce network costs and to obtain the most value from DER capabilities.

The regulatory conditions required to enable this degree of interaction largely depend on the organization and structure of the power sector, as extensively discussed in **Chapter 6**. For instance, in California, the regulatory commission has mandated that large investor-owned utilities deploy a certain quantity of storage capacity by 2020. This capacity will be used for several applications, including distribution network support (CPUC 2013). The network utilities themselves can own up to 50 percent of this storage capacity; the remainder would be owned by independent companies. In the European Union, by contrast, where there is retail competition and multiple actors own and control DERs, distribution companies are not allowed to own and operate DERs. In both situations, maximum efficiency would be achieved if a level playing field existed for all types of DERs, regardless of the structure of ownership and control. This requires a neutral market platform to facilitate all commercial transactions. Envisioning this approach, New York regulators have identified distribution utilities as distributed system platform providers, despite the absence of unbundling rules, as part of an on-going reform of the state's electric power sector (NYDPS 2014).

Implementing markets for network services at the distribution level will generate additional revenue streams for DERs and create new business opportunities that combine the flexibilities of a DER portfolio to respond to the needs of distribution companies or other upstream stakeholders.

Another new potential role for distribution companies is in the deployment of innovative technologies such as distributed storage, EV recharging infrastructure, and advanced metering infrastructure. In this case, the main question to be addressed by policy makers and regulators is whether these technologies should be considered as the distribution operator's assets (and thus regulated as part of the system operator's natural monopoly) or treated as assets to be delivered in a competitive fashion. The selection of one of these alternatives will be influenced by considerations of efficiency and economies of scope, market dominance, conflicts of interest, and existing unbundling rules.

Regulators must be vigilant of the potential for the distribution company to gain market dominance during this transition period. Depending on system complexity and the number of tasks to be accomplished by the distribution company, more effective unbundling or regulatory oversight may be required. Given that the roles of distribution utilities are beginning to look more like those of transmission system operators (TSOs), some lessons learned about the independence of TSOs could be relevant in a world where distribution utilities are transitioning to become "distribution system operators" (DSOs). These topics are thoroughly discussed in **Chapter 6**.

3.2.3 DERs and their interactions with the bulk power system

This section focuses on how DERs interact with the design, operation, and regulation of the bulk power system. Large influxes of DER capacity, namely variable and intermittent renewable energy such as solar and wind connected to low- or medium-voltage distribution networks, present several challenges to the functioning of wholesale electricity markets. Intermittent resources contribute to the need for rapid cycling of thermal resources and responsive demand to maintain a balance of generation and demand, and this increases operational costs. Moreover, zero variable cost renewable generation reduces average market prices and, due to high variability, leads to higher price volatility, which may potentially be offset by responsive demand. In addition, uncertainty about production schedules may contribute to more frequent or significant divergence of electricity prices between sequential markets, such as day-ahead and intraday or real-time markets.

Conversely, barriers exist that create difficulties for DER participation in wholesale markets, relative to conventional centralized technologies. For example, market participation requirements related to the size of a resource, as well as specific market product definitions, should be revisited. Moreover, the introduction of locational signals — e.g., locational marginal prices at the transmission and distribution levels that include the marginal cost of losses — will be relevant in terms of promoting efficient responses from network users. Because average, flat retail prices do not reflect temporal and locational differences in the value of electricity, they will lead network users to make economically inefficient choices with respect to DER investment and operation. Advantages and complexities derived from the implementation of locational energy prices are discussed in **Chapter 4**.

Finally, since DERs can contribute to the provision of electricity services at the transmission and distribution levels, much closer coordination between DSOs and TSOs will be required. Significant coordination will be needed in information exchange; monitoring and analytic capabilities; computation of prices for electricity services; forecasting, scheduling and activation of resources; and other system operator responsibilities.

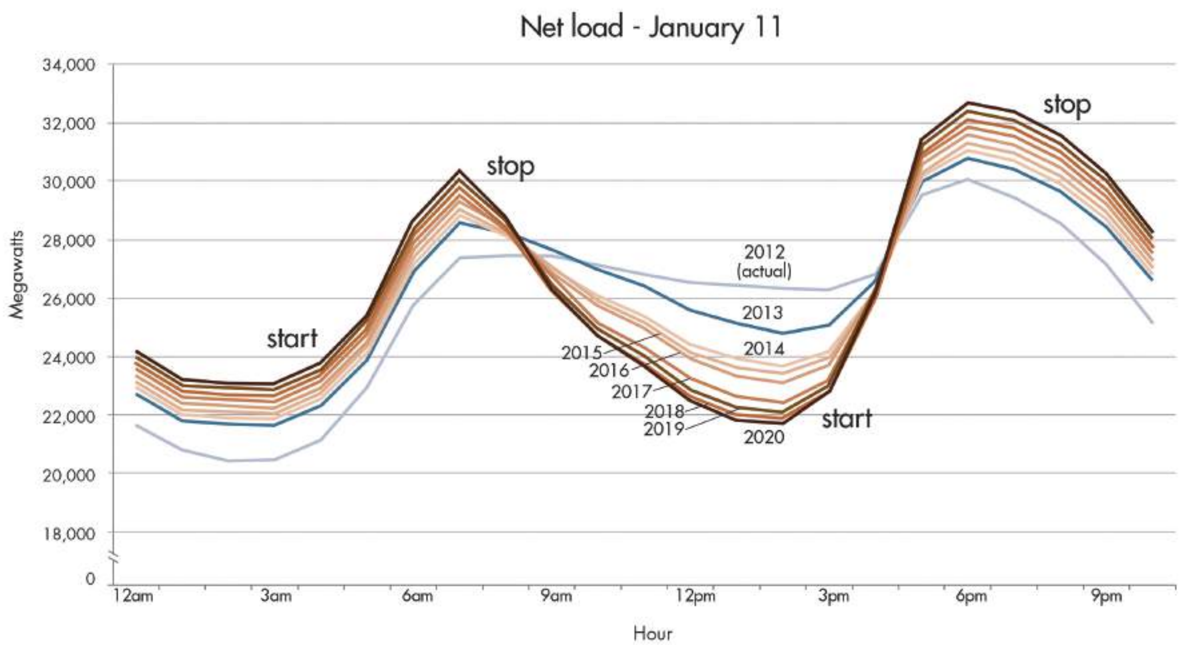
The effects of DERs on wholesale markets

Integrating DERs into wholesale markets presents several challenges related to how DERs affect the functioning of these markets. On the other hand, eliminating current barriers that constrain the participation of DERs relative to conventional centralized technologies is required for economic efficiency.

DER technologies, particularly variable renewable energy (VRE) technologies such as wind and solar,[16] are likely to experience significant growth in coming years — at both the distribution and transmission levels. The central challenge of integrating VRE technologies into power systems is their intermittency, a characteristic that requires other (usually thermal) resources to rapidly adapt their power output to maintain the instantaneous balance of generation and demand. Variations in VRE power output (whether expected or unexpected) will increase the need for flexible generation capacity in power systems. The fact that a rapid change in VRE generation can sometimes be predicted does not eliminate the need for fast-ramping resources. This is well illustrated by California's widely known "duck curve" (**Figure 3.10**). Increasing solar penetration in California's power system has led to a net load curve that necessitates significant ramping of thermal generators in the evening and drastic output reductions by those same generators during the daytime. The duck curve is not unique to California — it also occurs in other jurisdictions with high penetrations of solar PV such as Germany, Italy, and others.

16 Variable renewable energy technologies are characterized by volatile, partially unpredictable, and mostly non-dispatchable power output with zero fuel cost.

Figure 3.10: Duck Curve



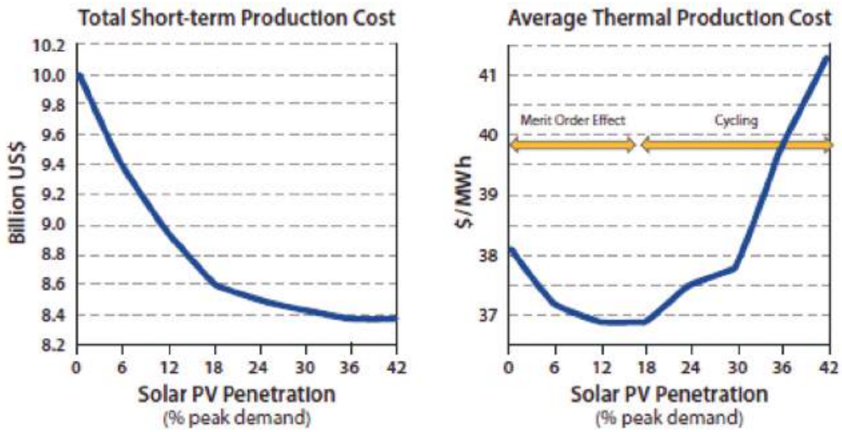
Source: CAISO (2013)

The best-known effect of high levels of VRE penetration is the so-called merit order effect: Zero variable cost generation displaces generators with higher variable costs, thereby immediately reducing wholesale market prices. Because VRE technologies are only intermittently available, the reduction in prices occurs only in those periods when solar or wind generation is available, although it is likely to lead to lower average wholesale prices.

In power systems where most of the variability of VRE generation will be absorbed by thermal generation, these thermal units will be forced to rapidly change their output and to start up and shut down more frequently. Consequently, costs associated with the cycling of thermal plants will increase and, for large penetrations of VRE technologies, may offset any cost reductions associated with the merit-order effect. Because of these thermal cycling costs, overall energy costs can increase. **Figure 3.11** illustrates this dynamic using a simulation of the impact of solar generation in ERCOT, the power system that serves most of the state of Texas.[17]

17 ERCOT stands for Electric Reliability Council of Texas.

Figure 3.11: Production Costs with Increasing Solar Penetration in ERCOT



Note: The figure shows total short-term production cost and average short-term production cost for thermal generators only associated with increasing solar penetrations in ERCOT, a thermal-dominated system. (MIT 2015)

In sum, VRE technologies have a twofold effect on spot prices. Prices can increase due to the cost of thermal generation cycling in certain periods and prices can decrease when VRE production displaces more expensive generation in other periods. In periods when high VRE production displaces costlier production, prices can fall to almost zero, reflecting the fact that there is excess electricity supply. Prices may even become negative when inflexible thermal generators are unable to come offline (because it is technically infeasible or uneconomic), and renewable generation is benefitting from priority dispatch rules and/or production subsidies. This dynamic can have the overall effect of increasing price volatility in spot electricity markets, which, combined with the unpredictability of VRE generation, can lead to electricity price divergence between sequential markets, such as day-ahead and intraday or real-time markets. A review of wholesale market rules is needed to address these issues; we return to this topic in **Chapter 7**.

Increasing penetration of variable renewable energy sources, such as solar PV and wind, will increase the need for flexible demand and generation to adapt to their intermittent output.

Active participation of DERs in wholesale markets

Current electricity wholesale markets present numerous barriers to the participation of DERs. In most cases, these barriers simply result from technology evolving at a faster pace than electricity market rules and regulations. The participation of DERs may be hindered by a lack of clear rules or by rules that were designed for large traditional resources and have not been updated.

Pioneering experiences in DER integration have illuminated the most urgent reforms required to enable DER participation in wholesale electricity markets. This section draws on some of these early experiences to highlight the most relevant barriers that exist today. It also bears noting that reforms to allow for DER participation are only the first step in a necessary redesign of electricity markets. Such reforms must be followed by efforts to ensure the efficient functioning of markets that include significant DER participation. This challenge is briefly introduced here, and more thoroughly addressed in **Chapter 7**.

Definition of market products

In order for DERs (aggregated or not) to compete efficiently with conventional resources they must be able to participate in all relevant markets, from long-term capacity markets to real-time and balancing markets.

Current market rules must be examined to enable the participation of DERs and to level the playing field between centralized and distributed resources.

In markets that allow DER participation, DER involvement may still be hindered by market products that were designed with the capabilities of conventional resources in mind and that therefore fail to fully reflect the needs of the power system or properly take into account the capabilities and limitations of DERs. The remuneration received for delivering these ill-designed (or outdated) market products may not capture the additional value of, for example, the flexibility provided by a DER relative to a thermal power plant.

The challenges associated with DER participation are very different across market segments. For example, in capacity markets, the key challenge is assessing the contribution to firm capacity by resources with very different technical characteristics. Traditionally, many capacity markets only allowed conventional generators to join. More recently, demand response has been allowed to take part in a growing number of capacity markets due to its ability to reduce energy loads and alleviate stress on the grid during emergency conditions.[18] However, assessing the actual capacity value of VRE and storage technologies still constitutes a barrier to tapping their potential contribution toward the overall capacity of the system.

18 Demand response (DR) programs have been the primary way in which DERs — not connected at the bulk power level — have had an opportunity to participate in US wholesale markets.

Box 3.3: Aggregations of Distributed Energy Resources Are Allowed to Bid in the California Wholesale Market

On June 2, 2016, the US Federal Energy Regulatory Commission (FERC) approved a proposal by the California Independent System Operator (CAISO) to allow individual energy resources that are too small to participate in the wholesale market to be grouped together to meet the market's minimum 0.5 megawatt (MW) threshold.

While some DERs were previously able to enter the market in limited ways, CAISO is the first grid operator in the United States to spell out a process for grouping various DERs to reach the threshold for market participation.

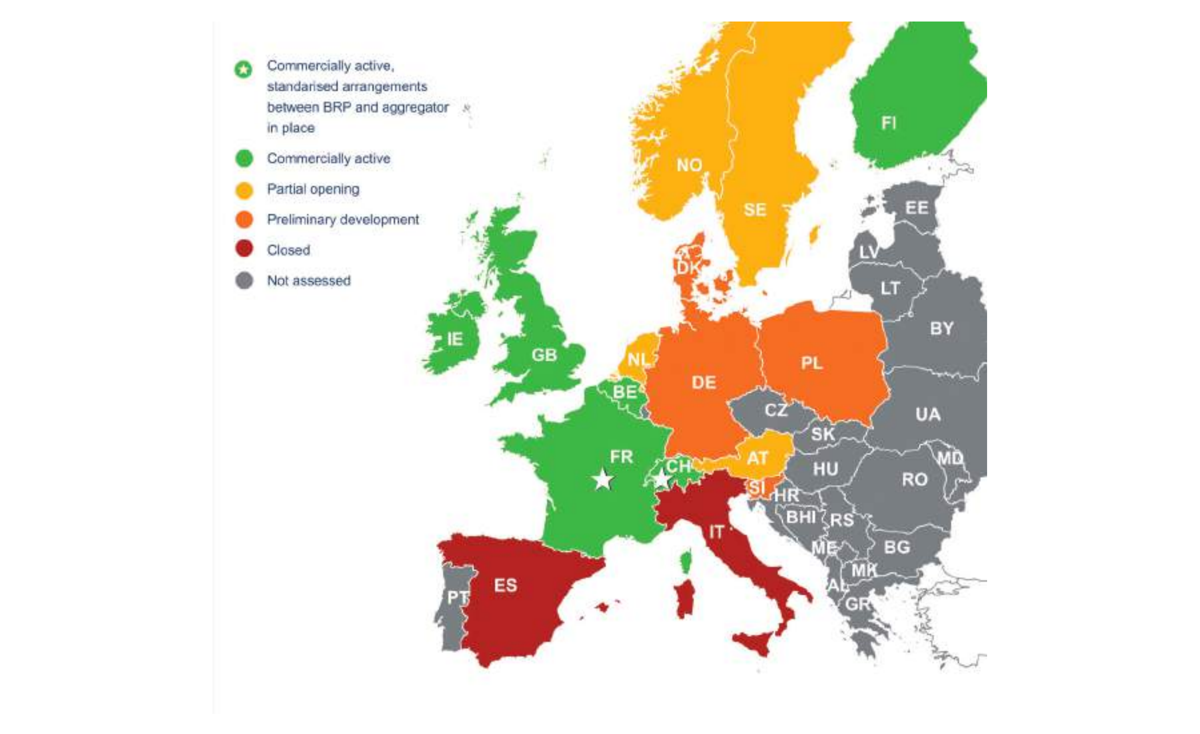
According to CAISO's President and CEO, "It's critical that the ISO collaborate with distribution system operators, regulators and market participants to harness these valuable distributed resources."*

* www.marketwired.com/press-release/california-iso-leads-historic-push-for-distributed-energy-resources-2132153.htm

Box 3.4: Markets for Demand-side Resources in Europe

The Smart Energy Demand Coalition (SEDC) was founded with the goal of "promoting the active participation of demand side resources in European electricity markets." In a 2015 market analysis, SEDC assessed European electricity markets according to whether they: (1) enabled consumer participation and aggregation; (2) have appropriate program requirements; (3) have fair and standardized measurement and verification requirements; and (4) have equitable payment and risk structures. SEDC's conclusions based on this analysis are summarized in **Figure 3.12**.

Figure 3.12: Map of Explicit Demand Response Development in Europe Today



Source: SEDC (2015)

While SEDC has acknowledged progress by some EU member states in implementing network codes that are favorable to demand response, it gave other member states a negative grade based on five remaining regulatory barriers to the deployment of DR resources: (1) demand response is not accepted as a resource; (2) baseline load estimates are inadequate and/or non-standardized, making it impossible to measure the amount of load reduction attributable to demand response; (3) technology-biased program requirements are present; (4) aggregation services are not fully enabled; and (5) standardized processes between balancing responsible parties (BRPs) and demand response aggregators are lacking.

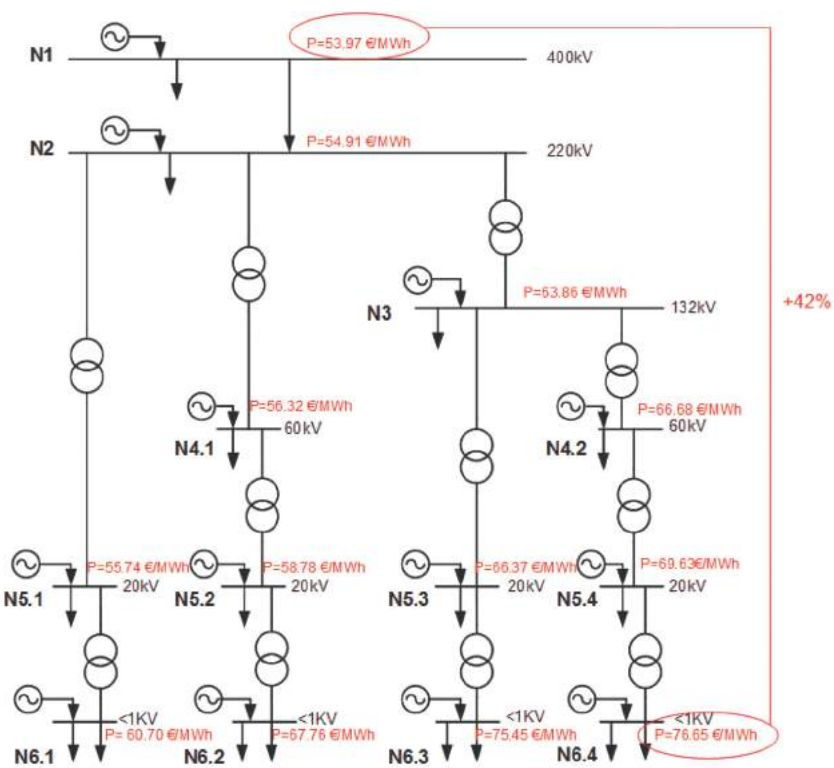
The effects of DERs on transmission and distribution networks

Many aspects of DER deployment — including which types of DERs are favored by investors, how extensively they are deployed, how they are operated (i.e., active participation or passive participation), and what impacts they have on the electric grid — depend critically on the prices and charges DERs face.

Extending the calculation of LMPs to distribution networks is conceptually simple but computationally data intensive because of the sheer number of nodes involved and the potential presence of DERs (Caramanis et al. 2016). The advantages and complexities of implementing LMPs at the distribution level are more thoroughly discussed in **Chapter 4**.

As an example, we use a simplified representation of a Spanish distribution network to show that the effect of losses and local network congestion (when it exists) is significant enough to affect investment and dispatch decisions for DERs. The simplified network in **Figure 3.13** is representative in that the share of losses at different voltage levels is similar to what it would be in an entire network. Using this network to estimate the impact of technical losses, we find differences in LMPs between voltage levels at peak hours of up to 40 percent — sufficient to make a large difference in DER investment and operational decisions. Note that in Figure 3.13, the power flows "downward" (i.e., to lower voltage levels) and therefore lower voltage levels are associated with higher LMPs. The situation would be the opposite in the case of an exporting (rather than importing) feeder. Therefore, pricing signals should reflect differences in operating conditions.

Figure 3.13: Distribution LMPs Using Marginal Loss Values Similar to the Spanish Distribution System



Source: Pérez-Arriaga (2016)[19]

19 For more details on the modeling used to produce this figure, see **Appendix A**.

This example illustrates that current pricing schemes — specifically, the use of average, flat prices that ignore locational and temporal variations in cost/value — have to be reconsidered to promote efficient decisions for DER investment and operation. This topic, along with a proposal for a new system of prices and charges, is discussed in detail in **Chapter 4**.

The marginal effect of energy losses can significantly affect locational marginal prices for electricity at different voltage levels on the grid.

DERs: Jumping over or going through the TSO-DSO fence?

In principle, DERs can contribute to the provision of all electricity services at both the transmission and distribution levels. But, as we have already emphasized, DERs can participate efficiently only if they have the economic incentives to do so, and access to the markets or mechanisms that make trade in these services possible. System operators at transmission (TSO) and distribution (DSO) levels are key institutions for enabling this transition and their coordination is of critical importance. Note that in some jurisdictions these two entities have different names and similar (but not identical) roles. The terms "independent system operator" (ISO) and "regional transmission organization" (RTO) in the United States correspond with the EU term "transmission system operator" (TSO), but ISOs and RTOs cannot own transmission assets while TSOs can. At the distribution level, the terms "distribution network operator" (DNO) and "distribution system operator" (DSO) are synonymous.

Coordinating DSOs and TSOs in an environment of increased complexity, and managing the new roles and functions of DSOs (discussed in an earlier section of this chapter), are critical to enable DERs to participate efficiently in supplying electricity services at the transmission and distribution levels (Ruester et al. 2013). For maximum efficiency, a DSO should be neutral with respect to the provision of services by different market agents. DERs and aggregators should be able to provide services directly to the TSO, if appropriate information is given to the DSO, and the DSO should have the ability to act appropriately to avoid constraint violations on the distribution grid. Conflicts can occur when actions by DERs mitigate a problem or provide a service at the TSO level, but create a problem at the DSO level. Therefore, coordination between the TSO and DSO — including a mutual understanding of their respective priorities — must be established.

Different actions are currently being undertaken around the world to efficiently integrate DERs into the power system and to reform the roles of the agents involved in that transformation. The European Commission (EC) and the New York State Department of Public Service are two examples of institutions that are actively pursuing increased coordination between DSOs (utilities) and TSOs/ISOs. Several organizations, including ENTSO-E (European Network of Transmission System Operators for Electricity), GWAC (GridWise Architecture Council in the US), ISGAN (International Smart Grid Action Networks), CIRED (International Conference on Electricity Distribution), and EDSO (European Distribution System Operators) have established task forces and working groups to investigate future roles and responsibilities, interactions with market players, data management, and coordination requirements for and between network operators (CEDEC et al. 2015; CEDEC et al. 2016; ISGAN 2014; Taft and Becker-Dippman 2015).

DERs can provide services to DSOs and TSOs. As a result, the roles of both types of system operators will continue to evolve as increasing penetration of DERs changes load and generation patterns. Coordination between system operators will also need to expand in areas such as information exchange; monitoring and analytic capabilities; pricing; forecasting, scheduling, and activating resources; and establishing system operator responsibilities. DSOs and TSOs must be able to monitor and engage resources while also studying and sharing information in a timely manner to enable efficient markets and reliable system operations. In sum, effective

The historical separation between transmission and distribution needs to be reconsidered. Because DERs can provide services at both levels of the system, coordination among transmission and distribution system operators becomes more relevant.

coordination between TSOs and DSOs is of utmost importance if the electricity system is to obtain full value from services provided by DERs (Pérez-Arriaga 2016).

3.3 Cybersecurity, Resilience, and Privacy with DERs

Protecting the electricity grid from cyber attacks is a critical national security issue. As the December 2015 cyber attack on Ukraine's power grid, and previous grid failures, such as the multiday blackout that hit the northeastern United States in 2003, have shown, any event that causes prolonged power outages over a large area is extremely costly — both to the economy and in terms of citizens' health and general welfare. Widespread connection of solar, wind, demand response, and other distributed energy resources to energy markets will increase the digital complexity of electric power systems. Therefore, more widespread and intensive cybersecurity protection and planning are required.

In particular, cybersecurity threats to the distribution system can be expected to challenge the industry for many decades. Throughout the world, utilities and non-utilities that interact with the grid need resilient systems and must be prepared to contain and minimize the consequences of cyber incidents. Because an increasing quantity of private and corporate information will be gathered and stored by utilities and their affiliated companies, utilities of the future will need to address privacy challenges. Increased use of Internet-connected devices in homes, offices, and industrial facilities will exacerbate these challenges, especially since many of these devices store their data in the cloud.

3.3.1 Cybersecurity threats, vulnerabilities, and new approaches

Cyber and physical security threats pose a significant and growing challenge to electric utilities. Unlike traditional threats to electric grid reliability, such as extreme weather, cyber threats are less predictable and therefore more difficult to anticipate and address. The ways in which a cyber attack can be conducted are numerous and the growing complexity and interconnectedness of electric grids is increasing the number of potential targets and vulnerabilities (MIT 2011; Campbell 2015; Smith et al. 2016; Nourian and Madnick 2015). The attack surfaces of software environments — that is, the different points where an unauthorized user (the "attacker") can try to enter or extract data — are increasing.

Cyber incidents can cause loss of grid control or damage to grid equipment due to deliberate tampering with data, firmware, algorithms, and communications; false data injection into pricing or demand systems; data exfiltration; and ransom demands to restore access to data. Much like the electromagnetic pulses that can be caused by nuclear explosions and major geomagnetic disturbances, widespread cyber attacks are generally high-impact, low-frequency events. Multiple smaller, lower-impact events may occur more frequently. Attacks on the financial and industrial sectors are typically financially motivated, whereas attacks on critical infrastructure systems tend to be politically or ideologically motivated. Future attacks may feature a mix of cyber and physical attacks and may be paired with social action to instill anxiety and fear.

Electric power systems are comprised of cyber systems, physical systems, and people. Failures can originate from physical or cyber attacks and from people acting mistakenly or purposely (i.e., with intent to harm). For instance, DER nodes can be compromised by strategically manipulating generation set points on a distribution feeder (Shelar and Amin 2016). Software attacks can damage variable frequency drives in electro-mechanical equipment to control motor speed and torque. Threats can be both external and internal to the power system. Traditional supervisory control and data acquisition (SCADA) systems, distributed control systems, and programmable logic controllers were designed as closed systems with limited control interfaces, but these technologies are now becoming digitized and are being designed to include more "intelligent" software and hardware components. This increase in digitization and complexity can create new opportunities for unauthorized outsiders to access, and potentially disrupt, these systems. Future SCADA and distributed control systems may have a secondary diagnostic infrastructure for the purpose of verifying that the system is operating properly and that data are coherent and are not being tampered with. Such a diagnostic center may serve multiple SCADAs in a given region in order to benefit from economies of scale.

Mobile communications connected to utility systems may compound the cyber risks that utilities confront. On December 23, 2015, synchronized multi-stage, multi-site attacks on Ukrainian electric utilities made equipment inoperable and unrecoverable, forcing manual operations to recover and provide power. This served as a "wake up call" to policymakers and regulators to take cyber attacks seriously (Electricity ISAC and SANS Industrial Control Systems 2016). Electric utility cyber incidents, in which systems are breached but not overtaken, continue to occur. Such breaches are necessary precursors to full-fledged cyber attacks, since potential attackers must conduct targeted assessments of specific utility systems prior to launching attacks. The US Cyber Command leadership has stated that cyber attacks by foreign states could cause catastrophic damage to portions of the US power grid (Rogers 2016).

The growth of the Internet of Things, which can improve efficiency and convenience, also expands vulnerabilities if sufficient cybersecurity and encryption have not been "built in" and vulnerable wireless protocols (e.g., ZigBee) are used. Wirelessly connected IoT devices, including smart light bulbs and other electrical components in a "smart home" or sensors or cameras at an industrial facility, are vulnerable to cyber disruptions and attacks, and could spread malicious code.

The usual purpose of malware that is targeted at electric utilities is to obtain control of a utility's systems. The goal may not be to shut down an entire system but rather to make the system less efficient, disrupt certain regions, game pricing models, gain information about a nation's electricity consumption and industrial operations, or prepare for future attacks.

Associated economic, health, and safety impacts can be large. Lloyd's 2015 Emerging Risk Report indicates that a widespread attack on the US grid, if it were to disable 50 out of the 676 large (over 100 megawatt) generators in a region, could have a $243 billion economic impact and incur more than $20 billion of insurance claims in 30 lines of business (Lloyd's 2015).

On a global level, the average annual cybersecurity budget of energy companies was approximately $3.6 million (US dollars) in 2014, and cybersecurity spending accounted for almost 4 percent of energy companies' information technology (IT) budgets in 2014 (Scottmadden 2015).

In sum, emerging energy markets that enable an active role for DERs and that are based on the near-instantaneous, high-volume exchange of digital information greatly increase the exposure of power systems to cyber attacks.

As **Table 3.2** shows, electricity regulatory agencies, electric utility coordinating organizations, and standards agencies all have roles to play in developing cybersecurity standards — for both the United States and Europe.

Table 3.2: Organizations and Standards Relevant for Cybersecurity in the United States and Europe

	UNITED STATES	EUROPE
REGULATORY ORGANIZATIONS	Federal Energy Regulatory Commission (FERC); North American Electric Reliability Corporation (NERC); state public utility commissions and public service commissions	European Commission (EC; including DG Energy); Agency for the Cooperation of Energy Regulators (ACER); Council of European Energy Regulators (CEER); national regulatory authorities (e.g., UK — Ofgem, Germany — Bundesnetzagentur)
COORDINATING ORGANIZATIONS	Electricity Sector Information Sharing and Analysis Center (E-ISAC); Industrial Control Systems — Computer Emergency Readiness Team (ISC-CERT); Electricity Sector Coordinating Council (ESCC); North American Transmission Forum; Edison Electric Institute (EEI)	European Union Agency for Network and Information Security Agency (ENISA); International Council on Large Electric Systems (CIGRE)
SUPPORTING ORGANIZATIONS	Department of Energy (DOE); Department of Homeland Security (DHS)	European Commission Joint Research Centre; European Network of Transmission System Operators for Electricity (ENTSO-E); European Network of Transmission System Operators for Gas (ENTSOG); NATO Cooperative Cyber Defense Centre
RELEVANT STANDARDS	National Institute of Standards and Technology (NIST) standards and cybersecurity framework; SANS Institute CIS Critical Security Controls	European standardization organizations (e.g., CEN, CENELEC, and ETSI) and CEN-CENELEC Focus Group on Cybersecurity; British Standards Institution (BSI)
INTERNATIONAL STANDARDS	International Electrotechnical Commission standards; Center for Internet Security critical security controls	

Box 3.5: NIST Cybersecurity Framework

The US National Institute of Standards and Technology (NIST) has established a cybersecurity framework that includes the following objectives (NIST 2014a):

- Identify (institutional understanding to manage cybersecurity risk to organizational systems, assets, data, and capabilities);
- Protect (implement the appropriate safeguards);
- Detect (identify the occurrence of a cybersecurity event, and enable timely responses);
- Respond (activities, to take action regarding a detected cybersecurity event); and
- Recover (restore the capabilities or critical infrastructure services that were impaired through a cybersecurity event).

Cybersecurity programs in all parts of the electric power system need to identify cyber threats and protect operational and IT networks. Poor architecture, some of which was designed decades ago when cyber threats were not as prevalent, can create vulnerabilities, allowing adversaries to obtain initial access, establish reliable inbound and outbound communications, and maintain a persistent presence inside a network. Effective defenders can try to stop these vectors of attack and vulnerable or poorly engineered systems can be replaced if funding allows. Unfortunately, there is no single strategy that prevents all cybersecurity attacks, or eliminates the possibility that DERs will introduce new cyber vulnerabilities (Smith et al. 2016).

Cybersecurity is important to maintain the integrity and correct operation of the electric grid. Thus, minimum cybersecurity regulatory standards are needed for all components of an interconnected network: the bulk power and transmission systems, distribution systems and distributed energy resources, metered points of connection with network users, and Internet-enabled devices in residential, commercial, and industrial buildings. All entities that interact with and connect to the electric grid (e.g., DERs, microgrids) should adhere to minimum cybersecurity standards, not only those entities, such as utilities, that are registered with the North American Electric Reliability Corporation (NERC). Non-traditional energy providers and electricity service providers, including DER aggregators, should be obligated to address cyber risks because their actions (or inactions) could have a dramatic impact on the overall security of the electric grid.

In 2013, the US government issued Executive Order 13636 on Improving Critical Infrastructure Cybersecurity.[20] Other cybersecurity standards, regulations, and practices in place in the United States include NERC's Critical Infrastructure Protection Standards Version 5, a cybersecurity framework developed by the National Institute of Standards and Technology (NIST), and the US Department of Energy's Electricity Subsector Cybersecurity Capability Maturity Model.

20 President Obama's Executive Order 13636 on Improving Critical Infrastructure Cybersecurity was issued on February 12, 2013. It is available at www.whitehouse.gov/the-press-office/2013/02/12/executive-order-improving-critical-infrastructure-cybersecurity.

In July 2016, the European Commission issued a communication on "strengthening Europe's cyber resilience system and fostering a competitive and innovative cybersecurity industry.[21] The European Commission also created the "Energy Expert Cyber Security Platform" (EECSP) with the aim of providing guidance to the Commission on policy and regulatory direction EU-wide. The first European legislation on cybersecurity, the Network and Information Security (NIS) Directive, was issued in July 2016 and entered into force August 2016. The directive provides legal measures to increase the overall level of cybersecurity in the EU by increasing cybersecurity capabilities in member states, enhancing cooperation on cybersecurity among member states, and requiring operators of essential services in the energy sector to take appropriate security measures and report incidents to national authorities. Once implemented, European consumers, governments, and businesses will be able to rely on more secure digital networks and infrastructure to reliably provide essential electricity services. However, significant coordination on the part of member states is required to reach similar levels of cybersecurity across all of the European Union. Moreover, although the NIS is a promising first step, more explicit cybersecurity regulations and best practices will need to be implemented in Europe.

A better understanding of the costs of meeting future standards for cybersecurity and resilience is required. In the United States, there is currently no single central authority for cybersecurity preparedness. The Federal Energy Regulatory Commission (FERC) and NERC have authority over cybersecurity standards development and compliance for the bulk power system, but there is no formal regulatory oversight of compliance with cybersecurity standards at the distribution system and for smaller aggregations of DERs. Some state public utility commissions have started to address cybersecurity challenges at the distribution level, but more decisive actions are required. Similar cybersecurity regulatory oversight and action is needed in Europe and in other

21 The European Union (European Parliament and the Council) Directive 2016/1148, concerning measures for a high common level of security of network and information systems across the Union, was issued on July 6, 2016. It is available at: eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.194.01.0001.01.ENG&toc=OJ:L:2016:194:TOC

parts of the world. For the millions of DER components that will ultimately be deployed, direct control by utilities is not feasible; rather, a hierarchical approach is necessary for utilities to interact with these widely dispersed cyber-physical systems. DER security specifications should be developed, updated, and used by utilities and other non-traditional utility entities for all types of installations. A National Electric Sector Cybersecurity Resource report titled "Cyber Security for DER Systems" provides common technical security requirements for autonomous cyber-physical DER systems, DER energy management systems, utility- and retailer-operated ICT, and DER interactions with ISOs, RTOs, and energy markets (Cleveland and Lee 2013). Additional cyber guidelines that address DER actors, logical interfaces, and logical interface categories, can be found in the NIST Interagency Report on "Guidelines for Smart Grid Cybersecurity" (NIST 2014b).

The International Electrotechnical Commission and its technical committee have also developed security recommendations and standards for information exchange for power systems. The Commission's 2016 technical report on "Resilience and security recommendations for power systems with distributed energy resources (DER) cyber-physical systems" provides cybersecurity recommendations and engineering and operational strategies for improving the resilience of power systems with interconnected DER systems (IEC 2016).

Electric vehicles may contribute to cyber risk in the distribution system when smart or price-sensitive charging strategies would require bi-directional communications between charging and distribution control systems. Moreover, these systems must support many vehicle types and therefore many communication protocols, which makes filtering and network analysis much more difficult. Additionally, like portable USB drives, electric vehicles can carry mobile viruses between grids or into homes, defeating network separation defenses. Therefore, manufacturers of hybrid or electric vehicles and autonomous driving features will need to integrate cybersecurity in the design phase.

The first step to defend against cyber attacks is to develop a robust cyber risk management culture. Each organization must start by identifying and classifying the risks it faces, and by undertaking a security assessment of its infrastructure. Once risks have been identified and classified, action plans must be developed and periodically reviewed. Well-vetted frameworks and policies for physical disaster management and recovery already exist; these can serve as precedents for approaching cyber risk management.

To address cybersecurity breaches (a precursor to possible cyber attacks), network planners and operators need to understand what constitutes baseline or "within-band" operations, and need to be prepared to detect and respond to anomalous cyber activity. As DER capacity increases, digital automation expands, and more interconnected services are provided, many new cyber "attack surfaces" will be created. Utilities and DER providers need approaches to defend against such attacks, reduce the "dwell time" of attackers, implement multiple layers of cyber defenses ("defense in depth"), and recover from cyber and physical attacks.

As more utility and operational systems employ "cloud-based" services and cloud computing for data storage and processing, enhanced and non-conventional cybersecurity methods will be required. If data from multiple utilities and entities are stored in the cloud, new cybersecurity approaches will be needed. Regular scans for known cyber vulnerabilities and malware at all levels of generation, transmission, and distribution will be required. Relying only on Internet perimeter security will fail to protect distributed networks. In these cases, enhanced encryption for individual software components, hardware, and data will be even more important. Because enhanced analytics and artificial intelligence can help identify "anomalous" or "out-of-band" behavior, these potential solutions merit attention. Both cybersecurity and dynamic cyberattack strategies will evolve, and compliance with NERC, FERC, and European Commission regulations will be only the starting point for more robust and proactive cybersecurity systems. Governance institutions and industry coordinating organizations to ensure continuous improvement in cybersecurity best practices for the industry as a whole will be required. These institutions can serve a role analogous to that of the Institute for Nuclear Power Operations, which promotes high levels of safety and reliability in commercial nuclear reactors.

The US Electricity Information Sharing and Analysis Center coordinates rapid information exchange about cyber incidents, providing coordination for more than 80 percent of the United States, a significant portion of Canada, and parts of Mexico. Electric utilities and other key organizations in Europe, Asia, South/Central America, and other parts of the world would benefit from a wider and more rapid exchange of cybersecurity information. Work is required to develop a framework that balances this sharing of information with the need to respect the individual concerns of participating bodies. To keep up with rapidly evolving cybersecurity threats against large and complex electric utility networks, there is a need for electric utilities, vendors, law enforcement agencies, and governments to share current cyber threat information and actionable intelligence. Established organizations in the United States, such as the Electricity Information Sharing and Analysis Center (E-ISAC), the Industrial Control Systems Computer Emergency Readiness Team (ICS-CERT), and intelligence agencies need to improve communications and continue to conduct emergency response exercises such as GridEx and Cyber Guard. In Europe, established organizations that enable trust-based sharing of security data and information, such as the European Energy Information Sharing & Analysis Centre (EE-ISAC), need an even broader mandate, more robust mitigation strategies, and approaches for coordinated recovery from cyber attacks. Ideally, robust information sharing across agencies and international organizations will generate trust among these agencies and organizations and contribute to the efficient processing of critical information (Choucri et al. 2016).

Providing cybersecurity for the electric grid requires developing a risk management culture; understanding the characteristics of baseline or "within-band" operations; rapid sharing of information about cyber threats; and active, skilled, and coordinated teams to detect and respond to anomalous cyber activity, defend against cyber attacks, reduce the "dwell time" of cyber attackers, and implement layered cyber defenses.

3.3.2 Resilience with DERs

The 9/11 attacks and Hurricanes Katrina and Sandy in the United States, the Tohoku earthquake and tsunami in Japan, the cyber attack on Ukraine's power grid in 2015, and other terror incidents and natural catastrophes demonstrate the critical importance of resilience for energy systems. The concept of resilience has been interpreted differently in different disciplines, such as psychology, physics, ecology, and engineering. In the United States, President Obama's 21st Presidential Policy Directive[22] defines this term in the context of energy infrastructure: "resilience is the ability of critical infrastructure to prepare for and adapt to changing conditions and withstand and recover rapidly from all hazards, which include natural disasters, industrial accidents, pandemics, cyber incidents, sabotage, acts of terrorism, or destructive criminal activity." Many jurisdictions around the world view resilience as an important characteristic of electricity networks, and are incorporating resiliency measures into national and regional legislation — another example is the European Commission's Programme for Critical Infrastructure Protection.[23] Practices that improve resilience at the system level have been in place for decades. In the

22 President Obama's "Directive on Critical Infrastructure Security and Resilience" (Presidential Policy Directive–21) was issued on February 12, 2013.

23 See Communication from the Commission on a European Programme for Critical Infrastructure Protection. COM (2006) 786 final at eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52006DC0786&from=en. Commission Staff Working Document on the Review of the European Programme for Critical Infrastructure Protection. SWD (2013) 318 final. 28 August 2013 at ec.europa.eu/energy/sites/ener/files/documents/20130828_epcip_commission_staff_working_document.pdf

1970s, a classification system was introduced to characterize the operating state of the power grid: normal, alert, emergency, in extremis, or restoration. Which classification applied in a given situation depended on operational conditions (as measured by, for example, frequency, voltage, power flows, etc.) and external conditions (such as weather). Since then, contingency analysis has been used in unit commitment dispatch. Moreover, operating power reserves have been created to respond to undesirable imbalances of demand and supply, load curtailment has been used to respond to severe imbalances, and strategically located generators have provided black-start capability to restore power after a blackout. Other good practices include stockpiling spare transformers and developing detailed response plans.

Power systems are changing and DERs will provide new opportunities to improve resilience system-wide. In future power systems, the current network topology (meshed transmission network and radial distribution network) will extend to integrate microgrids at the customer or community level (Ton and Smith 2012). If designed correctly, this build-out is expected to help isolate failures, provide alternative pathways for avoiding component failures, resolve local failures before the entire network is exposed to instability, and maintain continuity of service even under extremely severe conditions. Microgrids, when combined with DERs, can help provide resilience. With the use of "islanding" operations, microgrids can assist in black-start or continued operations if the broader network is compromised due to a cyber or physical incident.

A small number of success stories in Japan have demonstrated the potential of distributed energy systems to maintain power delivery under extreme conditions. Roppongi Hills is an area in downtown Tokyo that self-provides electricity, heat, and cooling. Service in this area was uninterrupted on March 11, 2011 following the 9.0 magnitude earthquake and tsunami, and Roppongi Hills helped restore service in other areas. Its Sendai microgrid was able to serve most of the nearby university campus as well as critical facilities such as a hospital. Likewise, in the United States, distributed energy systems and microgrids proved advantageous in the aftermath of Hurricane Sandy (Hampson et al. 2013; Marnay et al. 2015).

However, the presence of DERs also raises new challenges for resilience. More DERs entail more complex network configurations and operations. In networks with significant active DER participation, ICT systems will have to coordinate thousands (or hundreds of thousands) of distributed devices alongside hundreds (or thousands) of centralized generators. This alone can increase system vulnerability to extraordinary events. The power system of the future will consist of both cyber and physical assets that are tightly integrated, and all of these assets must be protected. But mere protection is not sufficient. As described in the previous section, a resilient power system must also include the ability of the cyber-physical grid to withstand and recover from malicious and inadvertent cyber and physical attacks. The traditional approach of redundancy does not work if a cyber attack disables all similar units.

Utilities need to carefully plan for extraordinary events. Utility mutual response agreements for responding to natural disasters may not work well for cyber attacks of unknown duration and scope. Increased deployment of DERs will require utilities to operate in a more complex environment, and may lead to common failure modes if the same DER software configurations are used across regions or nations. Utilities need to be able to operate in a partially manual mode, disconnected from the main grid, if digital controls and telecommunications are unreliable following a cyber attack. Spare equipment programs in the United States, such as the Spare Transformer Equipment Program, SpareConnect, and Grid Assurance can improve grid resilience. Similarly, resilience should be formally considered in EU planning, directives, and legislation, and should be improved at the regional level.

To meet the novel challenges posed by cyber threats, utilities are developing a tiered approach to sustain service during an attack and restore service once disruption occurs. These measures include developing control mechanisms that will not provide the full functionality of regular systems but can sustain limited vital operations and maintain "fall-back" mechanical controls (Stockton 2016). Because the control center is the nerve center of the power system, and because its resiliency is extremely important, the computer hardware and software in the energy management system should be designed to withstand failures and degrade gracefully when necessary. The control center must be protected from physical as well as cyber attacks, and a backup control center should be available. Adjacent control centers (for example, the PJM Interconnection and the Midwest Independent Transmission System Operator are adjacent systems in the United States) should partially back each other up (NRC 2012).

Resilience is important in minimizing interruptions of service due to extraordinary events such as cyber attacks. Enhancing resilience requires detailed planning, effective mutual assistance agreements, and thoughtful use of DERs and microgrids. Privacy is an ongoing concern.

3.3.3 Privacy

Privacy is an increasingly important issue for individuals, companies, and utility systems. Hence, the potential for widespread adoption of DERs and IoT technologies to exacerbate privacy concerns is an important topic. Vastly more information will become available with the increasing connectivity of electric and telecommunications devices. Data analytics and the opportunity for outside organizations to improve user experience and generate additional revenue will increase the amount of information that is held by electric utilities, DER service providers, and other service providers. If data are held in cloud storage and are accessible to multiple people and organizations, maintaining privacy will become more challenging. Current privacy restrictions, encryption requirements, and information disclosure requirements vary by country and region (within the United States, disclosure requirements also vary by state) and by the type of data held. If electric utilities begin to collaborate more with device control system aggregators, electric car owners, and vehicle charging aggregators, specific procedures to protect data breaches and information exfiltration will be required. The challenge is to simultaneously protect legitimate customer expectations of privacy, be a good steward of data, and apply analytics to create additional value for consumers.

In Europe, a recent revision of General Data Protection Regulation 2016/679 makes Data Protection Impact Assessments mandatory under certain conditions. These assessments are a key instrument to enhance data controllers' accountability.[24] The European Commission's energy group (DG Energy), the Joint Research Centre, and industry developed a Data Protection Impact Assessment Template for Smart Grid and Smart Metering Systems to help utilities assess smart grids when evaluating privacy and data protection.[25]

3.4 Peering into a Future Power System with Today's Technology

This study does not try to foresee or predict the future, rather our aim is to explore how fast-moving technological and policy developments might change the electric power sector in coming years. In that spirit, this section presents a vision of what a possible future could look like if abundant distributed resources are coupled with the widespread use of currently available ICTs. This vision is focused on the lower voltage distribution and retail levels where most agents are connected and where there is the most opportunity for technological developments. We use residential buildings as our unit of analysis but our findings apply equally to commercial and small industrial buildings.

With the continued development and deployment of IoT technologies, it is conceivable that most new residential electric and gas appliances could contain or will be connected to an Internet-enabled chip or device in the not-so-distant future. This will help enable customers to optimize all the energy services they consume or supply, according to their individual preferences. From a technology standpoint, this functionality could be implemented in many ways, but in our example we

24 See General Data Protection Regulation 2016/679 (Ref. European Union 2016/679) at eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32016R0679).

25 Available at ec.europa.eu/energy/en/test-phase-data-protection-impact-assessment-dpia-template-smart-grid-and-smart-metering-systems.

describe a single possibility[26] that relies on cloud-based computation and also coordination, since it is easily understood.[27] However, similar functionality could be achieved with technologies in which data storage and computation occurs locally: for example, with the use of energy boxes that are placed in individual households. In our cloud-based example, coordinating the response of multiple customers would be the responsibility of aggregators, who would be given permission to access end-user data (which, once again, would be stored on cloud-based servers).[28] Every few seconds or minutes, depending of the nature of the appliance (e.g., thermostat, refrigerator, air conditioning unit, battery, lights, or other), each end-user node would send and share current information, in an encrypted format, about energy use for each appliance.

For instance, a thermostat would send a temperature measurement and its user preferences regarding air conditioning, a refrigerator would send data about its consumption and energy requirements in the following hours, an air conditioning unit would send data about its consumption, a battery would send data regarding its state of charge and overall condition (i.e., the number of cycles it has performed), and so on. The hardware in each appliance would have the capability to receive and respond to external commands, thereby enabling automatic control over temperature, lighting levels, battery charge and discharge decisions, active and reactive power production (e.g., by a solar PV installation), appliance connection and disconnection, and other parameters.

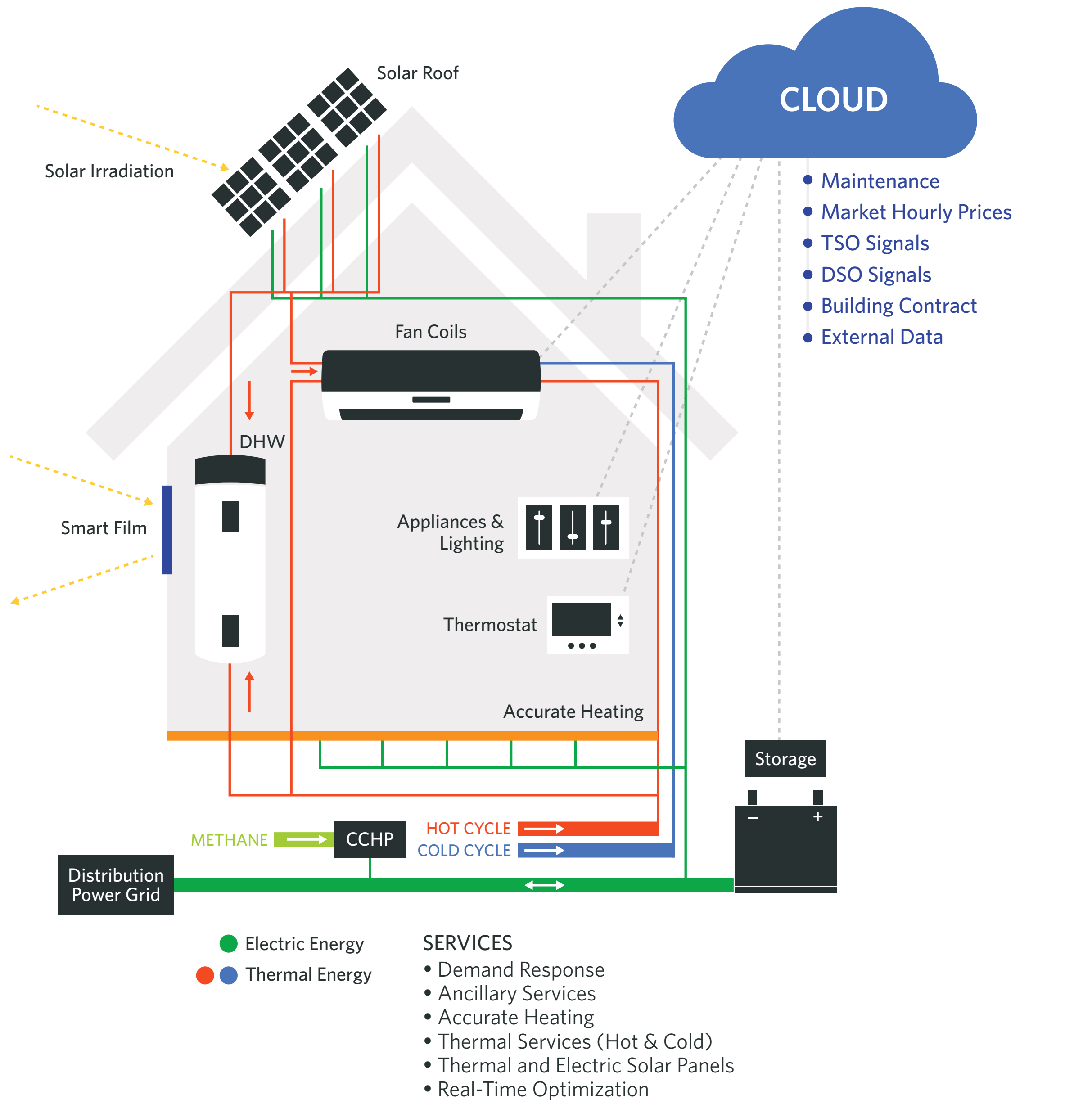
This type of technology configuration is illustrated in **Figure 3.14**, which depicts a smart home equipped with thermal and electric solar panels, electric and thermal storage, and a combined cooling and heating power unit (CCHP). Each of these technologies is connected to a cloud-based data management and computation system. The smart home would be able to manage air conditioning depending on the building's occupancy, send data to cloud-based storage, and control all the house's appliances and equipment based on signals that it receives from the cloud. Optimization algorithms (also computed in the cloud) would be used to generate these signals. Privacy restrictions could be put in place to limit access to end-user information stored in the cloud. The aggregator would have access to energy markets, as well as relevant external data (e.g., weather, prices for relevant electricity services, etc.), and would provide services to the transmission and distribution system operators on behalf of the smart home owner, while also optimizing the home's energy use (again, according to the preferences of the end user).

26 One among many possibilities for implementing energy management systems in buildings is the SPLODER technology, which is a cloud-based energy management system developed by researchers from the Institute of Research in Technology, at Comillas University. See: Calvillo, C. F., A. Sánchez-Miralles, J. Villar, and F. Martín. 2016. "Optimal Planning and Operation of Aggregated Distributed Energy Resources with Market Participation." *Applied Energy* 182 (November): 340–57.

27 One possibility is that all computation occurs in the cloud, individually for each "virtual energy box," but without any coordination. The example assumes that one or more aggregators manage the response of many customers in some coordinated fashion.

28 In the future, as technologies advance, aggregators may not need to play this role. See **Chapter 7**, which further discusses the role of aggregators.

Figure 3.14: Example of a Smart Home with Several DERs Connected to Cloud-Based Data Management and Computation Systems



Data generated at each building would converge (in the cloud) with information about the power system, such as the spot price of wholesale energy, the prices of operating reserves, information about requests for bulk power system support (e.g., to meet expected ramps or maintain appropriate frequency levels), information about the activation of commitments for demand reductions, and information pertaining to feeder-level distribution system markets and operating conditions (e.g., imminent voltage limit violations, imminent transformer capacity limits, etc.). The distribution system operator may also send emergency signals that limit the power that can be injected or withdrawn by any user for a prescribed time interval. However, the use of electricity prices that embed information about the availability of generation or about active network constraints should minimize (though probably not eliminate) the need for these types of emergency signals.

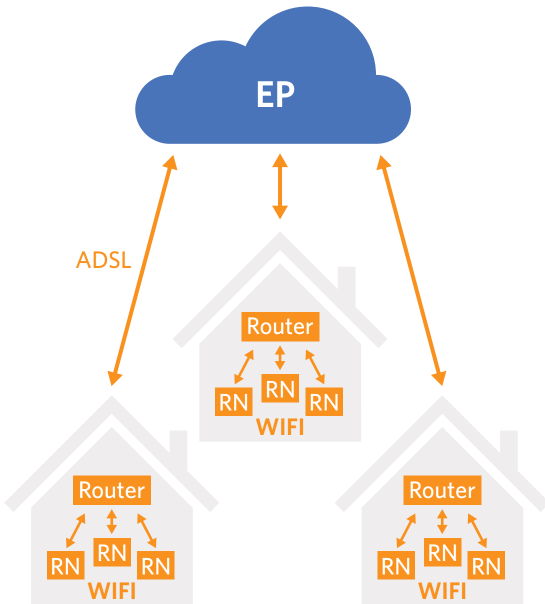
Several technological alternatives are currently available that could serve as the interface between individual smart homes and electricity markets. The conceptually simplest technology is an individual household-level "virtual energy box" that is contained in the cloud and that is able to process information from each home in relation to information from the power system. The energy box would optimize electric and thermal energy use in each home, including optimizing electricity self-generation, electric and thermal storage, air temperature, hot water, lighting, and other appliances — subject to preferences set by the user. The energy box would be able to learn from historical data and could select a model that most accurately describes the thermal behavior of the building, while also predicting (to a high degree of accuracy) the house's energy consumption patterns and profiles. Moreover, the box would be able to use weather and price forecasts to improve on the home's energy plan and could enact real-time decisions on behalf of its residents. The computer model that manages the virtual energy box will have seen comparable patterns and models for many other homes and other utility circuits across numerous intervals per year. This will enable the model to understand the likely performance of each home and feeder better than any individual building operator.

What is the purpose of this type of optimization? The most obvious purpose is to reduce total household energy bills while maintaining comfort and possibly also meeting other end-user preferences such as interest in supporting environmental objectives. In addition to reducing the cost of purchasing electricity services, the energy box would also account for revenues from selling electricity services to the distribution and transmission system operators (and possibly directly to other end users). Moreover, the energy box would choose among incompatible options, such as reducing demand to reduce feeder-level power flow and increasing demand to aid frequency control in the bulk power system, although actions that affect voltage or frequency control would have to be previously analyzed by system operators to make sure they do not have undesirable security implications. Each household would have discretion to choose between commercial providers of the control software used to integrate data from appliances, the power system, and customer preferences. The software would optimize the use of energy in the housing unit and could provide recommendations for improving household energy performance — for example, by buying new appliances, undertaking modifications such as insulation upgrades, and removing inefficient devices. All of these functions would occur automatically and would reflect household preferences and presets, with little or no need for human intervention. The savings that would result from households actively participating in markets for multiple electricity services could greatly exceed the relatively small cost savings that can be achieved today by manually shifting load from periods of high-energy prices to periods with expected lower prices.

It will take time, both for regulators to modify market rules and rate structures so as to enable individual residences to participate in diverse markets and for "virtual energy box" technology to become widespread. In the meantime, new or existing aggregators can contract with households to gain access to household energy data. The aggregator can use its portfolio of households (many of which, presumably, will have DERs) to jointly optimize the purchase and provision of electricity services while retaining a fraction of the cost savings that result.

A typical aggregator would manage the energy usage of households, commercial buildings, and small industrial facilities, with a portfolio that might eventually total tens of thousands, hundreds of thousands, or even millions of customers (**Figure 3.15**). As with a virtual energy box, aggregators' decisions on behalf of customers would be constrained by customer preferences. Such aggregators exist today, but due to current market rules and transaction costs, their customers tend to be large — that is, above a minimum size or response capacity — and include mostly commercial and industrial buildings. A potential benefit of aggregators over a household-level virtual energy box is that aggregators might have higher computational power and more (or better) information to determine price forecasts, weather conditions, and power system conditions. They might therefore be uniquely positioned to optimize the provision of electricity services.

Figure 3.15: Aggregation of Household Energy Management Through Cloud-based Data Management and Computation Systems



Note: RN=remote node; EP=energy platform

The availability of a vast store of data on building-level energy use and power system performance offers an enormous range of possibilities. The applications we have described so far still fall within the classical paradigm of a centralized marketplace, where all players send bids in order to purchase and offer services. In this paradigm, there is an institution that is charged with computing the price for each electricity service at any given time and location. The vision presented in this section simply extends the range of participants from large incumbent power plants to myriad distributed energy resources. However, the data platform that we have described could support additional possibilities that depart from the classical marketplace paradigm described above. For example, one could imagine that, instead of aggregators, so-called market facilitators emerge to enable peer-to-peer transactions between DERs or aggregations of DERs. These facilitators could scan available bids in the data store and propose bilateral matches that would occur independent of the centralized marketplace. Although bilateral trade is generally economically suboptimal because it tends to lead to different prices for the same service at the same time and place, agents may value peer-to-peer transactions for non-economic reasons. Moreover, if existing barriers are not removed soon, if existing tariffs are not redesigned, and if existing market rules (that were designed for large power plants) are not updated, DER users and entrepreneurs will find alternative pathways for market participation. While these alternative pathways may not be economically efficient, they will be sufficient for DERs to realize a fraction of their potential and, possibly, capture large shares of electricity services markets.

The vision presented here is not science fiction. All the technologies needed to achieve it — both the hardware and software — exist today and are ready for deployment. Pilot projects to test the capabilities of these new approaches are already helping to translate the vision into reality — what is needed now is updated regulation to enable these technologies and systems to be deployed at scale.

References

CAISO. 2013. "What the Duck Curve Tells Us about Managing a Green Grid." California Independent System Operator. www.caiso.com/Documents/FlexibleResourcesHelpRenewables_FastFacts.pdf.

Campbell, R.J. 2015. *Cybersecurity Issues for the Bulk Power System*. Washington DC: Congressional Research Service. R43989; 7-5700. June 10, 2015. www.fas.org/sgp/crs/misc/R43989.pdf.

Caramanis M., et al. 2016. "Co-optimization of Power and Reserves in Dynamic T&D Power Markets with Nondispatchable Renewable Generation and Distributed Energy Resources." *Proceedings of the IEEE* 104(4). April 2016.

CEDEC, EDSO, ENTSOE, EURELECTRIC, and GEODE. 2015. "General Guidelines for Reinforcing the Cooperation between TSOs and DSOs." November 2015. www.entsoe.eu/publications/position-papers/position-papers-archive/Pages/Position%20Papers/general-guidelines-for-reinforcing-the-cooperation-between-tsos-dsos.aspx.

CEDEC, EDSO, ENTSOE, EURELECTRIC, and GEODE. 2016. "TSO-DSO Data Management Report." July 2016. www.entsoe.eu/publications/position-papers/position-papers-archive/Pages/Position%20Papers/TSO-DSO-Data-Management-Report.aspx.

Choucri, N., S. Madnick, and P. Koepke. 2016. "Institutions for Cyber Security: International Responses and Data Sharing Initiatives." MIT Sloan School Cybersecurity Interdisciplinary Systems Laboratory. August 2016. web.mit.edu/smadnick/www/wp/2016-10.pdf.

Cleveland, F. and A. Lee. 2013. "Cyber Security for DER Systems." National Electric Sector Cybersecurity Organization Resource. Electric Power Research Institute (EPRI). July 2013. smartgrid.epri.com/doc/der-rpt-07-25-13.pdf.

CPUC. 2013. "Decision adopting energy storage procurement framework and design program." California Public Utilities Commission Decision 13-10-040. October 21, 2013.

EA Technology. 2012. "Assessing the Impact of Low Carbon Technologies on Great Britain's Power Distribution Networks." Prepared for Energy Networks Association on behalf of Smart Grids Forum — Work Stream 3, Version: 3.1. Report No: 82530. July 31, 2012.

EIA. 2015. *Annual Energy Outlook 2015*. Washington, DC: US Department of Energy, Energy Information Administration. April 2015.

Electricity-ISAC and SANS Industrial Control Systems. 2016. "Analysis of the Cyber Attack on the Ukrainian Power Grid." Electricity-Information Sharing and Analysis Center. ics.sans.org/media/E-ISAC_SANS_Ukraine_DUC_5.pdf.

Hampson A., T. Bourgeois, G. Dillingham, and I. Panzarella. 2013. *Combined Heat and Power: Enabling Resilient Energy Infrastructure for Critical Facilities.* ICF International Report for Oak Ridge National Laboratory. March 2013.

IEC. 2016. "Resilience and Security Recommendations for Power Systems with Distributed Energy Resources (DER) Cyber-physical Systems." IEC TR 62351-12:2016. International Electrotechnical Commission. April 2016.

ISGAN. 2014. "TSO-DSO Interaction: An Overview of Current Interaction Between Transmission and Distribution System Operators and an Assessment of their Cooperation in Smart Grids." ISGAN Discussion Paper Annex 6 Power T&D Systems, Task 5. International Smart Grid Action Network. September 2014.

Lloyd's. 2015. "Business Blackout: The Insurance Implications of a Cyber Attack on the US Power Grid." Lloyd's Emerging Risk Report. Lloyd's and University of Cambridge Center for Risk Studies. www.lloyds.com/~/media/files/news%20and%20insight/risk%20insight/2015/business%20blackout/business%20blackout20150708.pdf.

Marnay C., H. Aki, K. Hirose, A. Kwasinski, S. Ogura, and T. Shinji. 2015. "Japan's Pivot to Resilience: How the Microgrid Fared after 2011 Earthquake." *IEEE Power & Energy Magazine* 13(3): 44-57. May 2015.

MIT. 2011. *The Future of the Electric Grid*. Chapter 9. Cambridge, MA: Massachusetts Institute of Technology. December 2011.

NIST. 2014a. Framework for Improving Critical Infrastructure Cybersecurity Version 1.0. National Institute of Standards and Technology. February 12, 2014. www.nist.gov/cyberframework/upload/cybersecurity-framework-021214.pdf.

NIST. 2014b. *Interagency Report on Guidelines for Smart Grid Cybersecurity.* NISTIR 7628: Volumes 1-3. National Institute of Standards and Technology. September 2014.

Nourian, A. and S. Madnick. 2015. "A Systems Theoretic Approach to the Security Threats in Cyber Physical Systems: Applied to Stuxnet." *IEEE Transactions on Dependable and Secure Computing*. MIT Sloan School Cybersecurity Interdisciplinary Systems Laboratory. December 2015. web.mit.edu/smadnick/www/wp/2015-07.pdf.

NRC. 2012. "Terrorism and the Electric Power Delivery System." In *Mitigating the Impact of Attacks on the Power System*. National Research Council. Washington, DC: The National Academies Press. doi:10.17226/12050. www.nap.edu/catalog/12050/terrorism-and-the-electric-power-delivery-system.

NYDPS. 2014. "Reforming the Energy Vision. NYS Department of Public Service Staff Report and Proposal." State of New York Department of Public Service. CASE 14-M-0101. Proceeding on Motion of the Commission in Regard to Reforming the Energy Vision. April 24, 2014.

Ofgem. 2013. "Strategy decisions for the RIIO-ED1 electricity distribution price control. Business plans and proportionate treatment." As Supplementary annex to RIIO-ED1 overview paper. Office of Gas and Electricity Markets. 26b/13. March 4, 2013.

Pérez-Arriaga, I.J. 2016. "The Transmission of the Future: The Impact of Distributed Energy Resources on the Network." *IEEE Power and Energy Magazine* 14(4): 41-53. July 2016.

Rogers, M.S. 2016. "Statement before the Senate Armed Services Committee." April 5, 2016. www.armed-services.senate.gov/imo/media/doc/Rogers_04-05-16.pdf.

Scottmadden Management Consultants. 2015. *Energy Industry Cybersecurity Report*. July 2015.

SEDC. 2015. *Mapping Demand Response in Europe Today*. September 30, 2015. www.smartenergydemand.eu/wp-content/uploads/2015/09/Mapping-Demand-Response-in-Europe-Today-2015.pdf.

Shelar, D. and S. Amin. 2016. "Security Assessment of Electricity Distribution Networks under DER Node Compromises." IEEE Transactions on Control of Network Systems.

Smith, E., et al. 2016. "Going Beyond Cybersecurity Compliance: What Power and Utility Companies Really Need to Consider." *IEEE Power and Energy Magazine* 14(5): 48-56. September/October 2016.

Stockton, P. 2016. "Superstorm Sandy: Implications for Designing a Post-Cyber Attack Power Restoration System." The Johns Hopkins University Applied Physics Laboratory. NSAD-R-15-075.

Taft J.D. and A. Becker-Dippman. 2015. *Grid Architecture*. Richland, WA: Pacific Northwest National Laboratory. January 2015. www.pnnl.gov/main/publications/external/technical_reports/PNNL-24044.pdf.

Ruester S., I. Pérez-Arriaga, S. Schwenen, C. Batlle, and J.M. Glachant. 2013. *From Distribution Networks to Smart Distribution Systems: Rethinking the Regulation of European Electricity DSOs*. THINK Project, Topic 12. Florence, Italy: European University Institute.

Ton D.T. and M.A. Smith. 2012. "The U.S. Department of Energy's Microgrid Initiative." *The Electricity Journal* 25(8): 84-94. October 2012.

PART 2: A FRAMEWORK FOR AN EFFICIENT AND EVOLVING POWER SYSTEM

04

A Comprehensive and Efficient System of Prices and Regulated Charges for Electricity Services

4.1 The Importance of Getting Prices and Charges Right

Distributed technologies allow the agents in a power system to meet their own energy needs and to deliver electricity services to the system. Distributed energy resources (DERs) bring new options for service provision and enable demand for electricity services to become increasingly price-responsive. Well-designed prices and charges become ever more important in this environment, as agents are increasingly capable of adapting their behaviors to power system conditions at specific locations, during particular times, and in relation to what specific services are being consumed or provided.

Acknowledging the importance of designing economic signals that reflect power system costs and serve as efficient signals for distributed decision-makers is not new (Bonbright 1961; Schweppe 1978; Schweppe et al. 1988). Even in the absence of DERs, there are clear benefits to using efficient economic signals, since they result in more efficient response of demand connected at all voltage levels and also in enhanced efficiency of operations and investment in the bulk power system. The growing integration of DERs increases the importance of well-designed economic signals and the ramifications of poorly designed signals.

DERs are a new class of potential competitors for the provision of electricity services traditionally delivered by centralized generators and network assets. Economic signals that accurately reflect the costs and values of utilizing varying resources for service provision will enable the creation of a level playing field on which centralized and distributed energy resources can compete and complement one another in an efficient resource portfolio (Pérez-Arriaga et al. 2013). These economic signals serve to coordinate planning and operational decisions related to all resources and make it possible to achieve efficient outcomes — particularly if the resources are numerous and distributed. Ideally, these signals will elicit efficient responses from all resources, at all times, no matter where they are located in the network.

Centralized and distributed resources are sited in different locations and have different sizes and temporal patterns for both production and consumption. The only way that these two categories of resources can jointly and efficiently operate, expand, compete, and collaborate, is in the presence of a comprehensive system of economic signals — market-determined prices and regulated charges[1] — with adequate granularity to capture important variations in the value of the specific services across time and location. This study advocates the need for a comprehensive system of prices (for those services provided in markets) and regulated charges (for remuneration of network activities and any policy costs included in electricity rates); such a system can act as the nervous system of the power sector, coordinating the actions of many disparate providers and consumers of electricity services by communicating prices and charges that reflect time- and location-specific conditions.

Ideally, these prices and regulated charges will reflect all the operating conditions and investment needs of the system, and the markets for electricity services will allow the participation of all system users. For example, to fully realize the possible benefits of distributed storage or price-responsive demand,[2] energy prices that sufficiently reflect temporal and spatial variations in the costs of meeting electricity demand or providing electricity generation must be communicated to distribution system users, aggregators (including retailers), and/or automated energy management systems. These economic signals must also encourage the placement and operation of DERs when and where they can prove more cost-effective than centralized providers of services.

This study advocates the need for a comprehensive system of prices (for those services provided in markets) and regulated charges (for remuneration of the network activities and any policy costs included in electricity rates); such a system can act as the nervous system of the power sector, coordinating the actions of many disparate providers and consumers of electricity services by communicating prices and charges that reflect time- and location-specific conditions.

To date, most power systems have employed simplified prices and charges to allocate power system costs to most electricity customers; these are proving inadequate in the face of increasing penetration of DERs and opportunities for flexible demand (see **Section 4.3**). Such simple methods exhibit limited (if any) temporal or spatial differentiation and result in tariffs that typically bundle costs of all of the services that customers receive in a nontransparent manner. As a result, some electricity users are making inefficient investments in DERs and are over- or undercompensated for the services that they provide to the power system. At the same time, many more opportunities to deliver value are being left untapped because of inadequate compensation. The result is a power system that costs more than is necessary — a loss in overall efficiency — and is unable to reveal and appropriately compensate the value that DERs and price-responsive demand can provide, as will be shown in the following sections.

1 Tax incentives and other policy incentives can be part of the ensemble of economic signals that the power system agents are subject to, beyond the basic system of prices and charges.

2 In this report, "price-responsive demand" means the efficient spontaneous response of demand to the existing price. See **Chapter 2** for precise definitions of "price-responsive demand" and "demand response."

This chapter is critical to understanding the message of this study, and it allows two levels of reading. For those who are very familiar with the power sector (specifically, with how the price of electricity and the tariffs are determined) and others who want to get the main messages of the chapter in concise format, it should suffice to read **Section 4.2**, which provides the general principles underlying the design of prices for electricity services and charges for network and policy costs; perhaps do a cursory reading of **Section 4.3**, which summarizes electricity pricing and cost allocation practices current in multiple jurisdictions of the United States, Europe, and elsewhere, and highlights important areas for reform; and then jump to **Section 4.6**, which summarizes the main findings behind the recommendations for the most effective and easiest to implement improvements in electricity tariffs.

Those looking for greater detail on how electricity prices and charges can be computed with an increased level of granularity should also read **Section 4.4**, which first presents an efficiency-driven approach to the pricing of electricity services and the design of network and policy charges, then considers how practical challenges to the implementation of such an "efficient ideal" can be addressed via implementation of proxies with manageable complexity.

Finally, those who want to be convinced of the increases in efficiency gained by avoiding common errors in tariff design and progressively improving the granularity in time and location conveyed by prices and charges should read **Section 4.5**, which provides examples of the benefits[3] of employing prices and charges that more and more closely approximate the efficient ones.

3 This study also comments on the costs and complexity associated with increasing granularity in time and location, but to a smaller extent.

4.2 Basic Principles for the Design of Prices and Regulated Charges

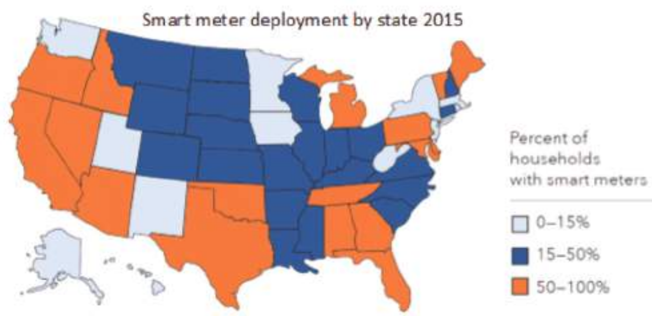
As grid users introduce DERs, their network utilization patterns become increasingly diverse, making it no longer meaningful to continue using existing customer classifications when designing electricity tariffs. The present customary assumption that "all consumers connected in low-voltage level behave the same" is already or will soon become untenable. Fortunately, advanced meters allow each individual network utilization profile to be recorded hourly or even at shorter time intervals. Such advanced metering infrastructure is expected to be widespread among electricity customers in coming years (see **Figure 4.1**, which shows the current level of smart meter deployment in the United States and the expected deployment in Europe for 2020).

Furthermore, prices and charges for electricity services should be nondiscriminatory, and thus agnostic to the particular activities for which the network is used, or "technology neutral." Accordingly, any cost-reflective component of prices and regulated charges should be based exclusively on the individual injections and withdrawals at the network connection point, regardless of the specific technology producing those injections or withdrawals.[4] Such an approach differs from a "value of resource" approach that provides incentives targeted toward specific resources, and is instead similar to a "value of service" approach in that the costs and benefits — or overall value to the power system — of any resource can be fully captured via prices and charges for services such as those described in **Chapter 2** (NARUC 2016). In addition, cost-reflective prices and regulated charges should be symmetrical, with a marginal injection

4 From the point of view of the power system, it does not make any difference whether a change in the power injected or withdrawn at a given connection point and moment in time has been caused by turning off an appliance, discharging a battery, or injecting more power from a photovoltaic panel or micro cogeneration unit. The impact on the overall system does not depend on the technology involved, so prices and charges shouldn't either. Moreover, it is a hopeless task to try to intrude behind the meter and apply different rates depending on the nature of the device used. There may be reasons to favor or disfavor particular technologies besides power system efficiency, but efficiency will be sacrificed if this is done. The allocation of residual costs is a different topic (addressed later in this chapter) where the most efficient cost allocation criterion may distinguish between types of consumers, for instance according to their price elasticity.

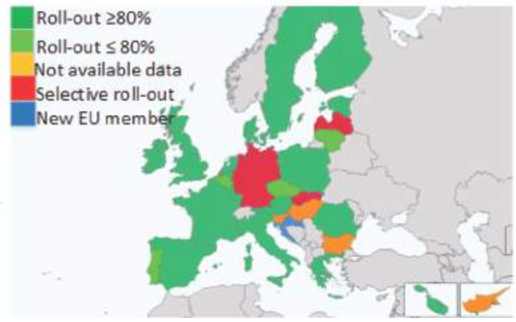
Figure 4.1: Smart Meter Deployment

4.1a: Smart Meters Installed in the United States as of 2015



Source: JRC Smart Electricity Systems and Interoperability (2016)

4.1b: Expected Level of Smart Meter Deployment in the European Union in 2020



Source: Institute for Electric Innovation (2016)

at a given time and place compensated at the same rate that is charged for a marginal withdrawal at the same time and place.[5] Designing prices and charges in a way that is technology-neutral and captures relevant differences in the costs and benefits offered by injections and withdrawals effectively obviates the need for defining a DER-specific customer class and subclasses. Any subsidies or penalties — or other kinds of explicit support or burden on specific technologies or resources — can be legitimate and justified in many cases (for instance, to internalize some externalities), but they should be applied in addition to and separately from the economic signals that will be developed in this chapter.

To establish a level playing field for all resources, cost-reflective electricity prices and regulated charges should be based only on what is metered at the point of connection to the power system — that is, the injections and withdrawals of electric power at a given time and place, rather than the specific devices behind the meter. In addition, cost-reflective prices and regulated charges should be symmetrical, with injection at a given time and place compensated at the same rate as that charged for withdrawal at the same time and place.

Prices and charges should accomplish two key objectives: (1) They should send efficient economic signals to the agents in the system, and (2) they should recover the regulated costs.

There is a long history of literature in applied economics devoted to the topic of tariff design in regulated sectors and the power sector specifically (see, for instance, Bonbright 1961; Schweppe et al. 1988; and Pérez-Arriaga and Smeers 2003). For those power systems that have been restructured and liberalized, the "price of electricity" incorporates the charges to recover the regulated costs of services not provided by markets (mostly network services), the prices and charges for services provided via markets (such as energy and some operating

5 If these prices and charges are not symmetrical, incentives can be significantly distorted. For example, if exports of power are compensated at a lower rate than withdrawals or consumption of power at the same location and time, then a generator that is individually metered and exporting power will earn less than a behind-the-meter generator at the same location, despite having the exact same value to the power system. Similarly, a megawatt-hour of demand reduction behind the meter would see greater economic value than a megawatt-hour of injection from a generator at the same location. These kinds of distortions should be minimized in order to put all resources on a level playing field and avoid inefficient regulatory arbitrage opportunities.

A comprehensive system of prices and charges consists of four core elements: (1) a price for electric energy; (2) prices or charges for other energy-related services, such as operating reserves or firm capacity; (3) charges for network-related services; and (4) charges to recover policy costs.

reserves), as well as other regulated taxes and items that regulatory authorities decide to include in the electricity tariff. Among the diverse criteria advocated by experts and embodied in most tariff designs, two principles are dominant: allocative efficiency and sufficiency to recover the regulated costs.[6] Only the second one is routinely met in practice.

A comprehensive system of prices and charges that provides economic signals and enables cost recovery consists of the following core elements: (1) a price for electric energy (active and reactive, in principle); (2) prices or charges for other energy-related services, such as operating reserves or firm capacity; (3) charges for network-related services; and (4) charges to recover policy costs (such as taxes and costs incurred supporting energy efficiency or renewable targets).

Following the first objective, efficient economic signals should try to capture and reflect the marginal or incremental costs of the production and utilization of electricity services. Such signals serve as the key tools with which to coordinate all the planning and operational decisions made by the diverse range of power sector agents to achieve efficient outcomes. For services provided competitively, the corresponding markets generally provide the required prices. For other services, regulated charges must be designed to send efficient signals reflecting each user's marginal or incremental contribution to the regulated costs (such as network capacity).

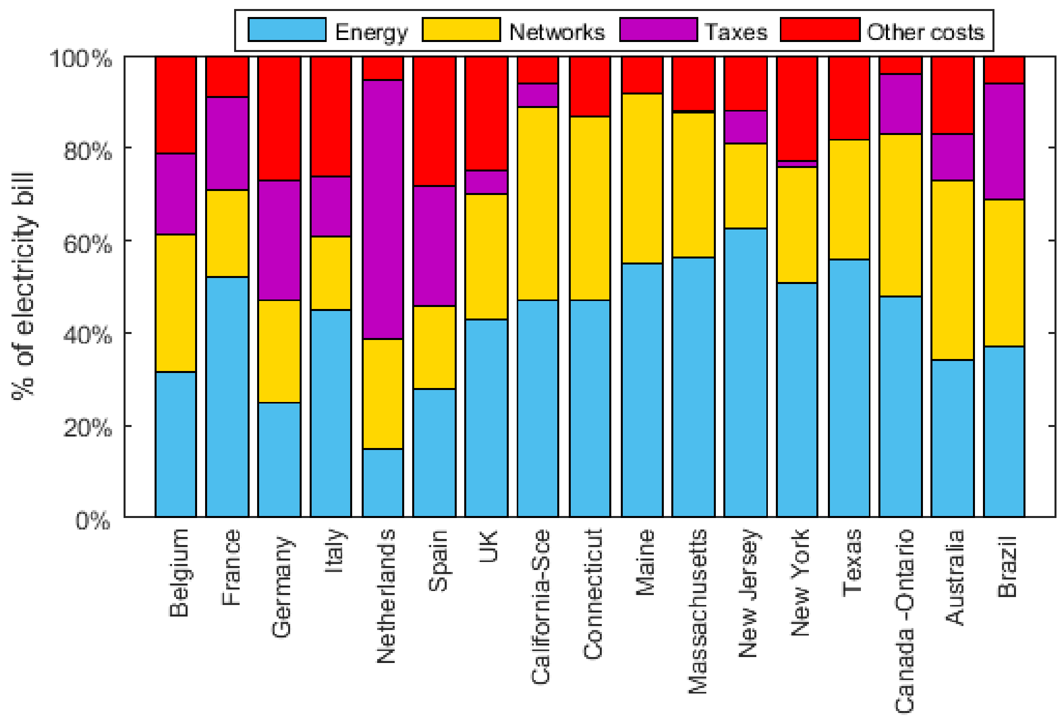
The second objective requires that prices and charges also enable the economic sustainability of regulated services via recovery of regulated costs (such as distribution network costs and policy costs). As will be shown later in this chapter, while prices and charges that provide economic signals by reflecting marginal or incremental costs contribute to recovery of regulated network costs,[7] such prices and charges alone are unlikely to be sufficient for full cost recovery. Regulated costs not recovered via cost-reflective prices and charges — the "residual costs" — should be recovered in a minimally distortive manner. In addition, electricity bills have often been a convenient tool used by politicians to allocate costs derived from energy policy objectives such as energy efficiency, climate change mitigation, or income redistribution. These taxes and a large component of policy-related costs represent a significant part of electricity bills in some jurisdiction (**Figure 4.2**). Any such costs not affected by changes in electricity consumption add to the amount of "residual costs," and must also be recovered in minimally distortive manner.

If not carefully designed, charges to recover the residual costs may distort the economic signals received by power system users, leading to decisions that increase the overall cost of power systems and/or shift policy and network costs onto other users. For example, inefficient installation and use of behind-the-meter generation may be incentivized if reducing volumetric electricity consumption allows users to reduce their payment of policy costs or regulated network charges that are allocated via volumetric rates but are not a direct function of volumetric consumption (see **Boxes 4.2** and **4.3** for more details). This study proposes methods to allocate residual costs and in a manner that minimizes economic distortions.

6 Besides allocative efficiency and economic sustainability, other commonly mentioned criteria are: transparency in making public the procedure that is followed to compute prices and charges; as much simplicity as possible, to facilitate understanding and acceptance, without compromising other more important criteria; consistency with the adopted regulatory framework, specifically the level of liberalization and unbundling of the different activities; and stability to minimize deviation from the status quo as a basic regulatory principle meant to reduce the risk perceived by the agents in the power sector and to enhance the social acceptability of the electricity tariffs.

7 For example, the differences among locational marginal prices (LMPs) of energy in the nodes of the network, due to the effect of losses and congestion, contribute to network cost recovery because they generate revenues when charging for consumption and paying generation at their individual LMPs.

Figure 4.2: Breakdown of Residential Electricity Bills in Different Jurisdictions in 2014–2015



Source: MIT Energy Initiative elaboration based on publicly available data

Care must be taken to minimize distortions from charges that are designed to collect taxes, recover the costs of public policies (e.g., efficiency programs, subsidies for renewable energy, cross-subsidies between different categories of consumers, etc.), and recover residual network costs (i.e., those network costs that are not recovered via cost-reflective charges).

On the basis of fundamental principles of economics and power systems, this study proposes a framework for the design of prices and charges according to what can be labeled an “efficient ideal” approach (i.e., an approach based on achieving economically efficient outcomes). The implementation of such an approach under a range of specific technical, regulatory, policy, and social contexts may present numerous challenges; therefore, alternative approximate or quasi-ideal schemes will also be considered.

This study proposes an "efficient ideal" approach (i.e., an approach based on achieving economically efficient outcomes) for the design of prices and charges. However, the use of "proxies" or "second-best" alternative schemes may help address a range of specific technical, regulatory, policy, and social implementation challenges.

In addition, the economic benefits resulting from increasing the temporal or spatial granularity of prices and charges should also be compared to any associated implementation costs. The appropriate level of spatial and temporal granularity of the economic signals in each jurisdiction should thus reflect pertinent trade-offs between both benefits and costs in each context. For example, while greater spatial and temporal granularity of economic signals can yield system-wide efficiency improvements, costs associated with computational complexity, market operation and administration, information and communication requirements, and behavioral response[8] may outweigh the benefits gained from a given degree of granularity.

Both economic benefits and relevant implementation costs should be considered when identifying the appropriate level of spatial and temporal granularity of economic signals in different jurisdictions.

Furthermore, the end of the "top-down paradigm" (**Chapter 3**) may have implications for the allocation of network charges. Presently, distribution network costs are paid only by users directly connected to the distribution network, with no allocation of distribution network costs to generators or large consumers who, while directly connected to the transmission network, utilize the distribution network to either transmit generated power or perhaps consume power generated in the distribution network. This top-down approach to network cost allocation made sense when network users were only either "upstream" centralized generators or "downstream" purely end consumers who paid the full supply cost. This rule has to be reconsidered when DERs become more common, since generation and load are then located all over the network. In such a case, the present cost allocation rules could create a tilted playing field for some of the players as centralized and distributed resources compete to provide the same services.

When computing network charges in the presence of many DERs, the current quasi-arbitrary separation between transmission and distribution networks has to be reconsidered.

Finally, widespread deployment of some DERs may have benefits or costs outside the power sector, such as land-use impacts, reductions in air pollutants, or mitigation of greenhouse gases. These positive or negative externalities should be internalized in the system of prices and charges as much as possible (also for centralized generation) and, in any case, accounted for in any comprehensive economic evaluation leading to policy decisions.

4.3 Current Electricity Tariff Practices

All agents connected to the power grid — whether producers, consumers, or any combination thereof, and under any regulatory framework (e.g., market-based or traditional cost-of-service) — are subject to some economic regime that determines the remuneration and/or charges that correspond to their respective situations. This system of economic signals has evolved with time, mostly driven by the changes that have taken place in the regulation of the power sector and the evolution of the information and

8 A behaviorally informed analysis shows that when transaction costs and decision biases are taken into account, the most cost-reflective policies are not necessarily the most efficient (see Schneider and Sunstein 2016).

communication technologies available for metering, data management, and communication. Progress has also been made in understanding the complex economics of the power sector.

Determining the economic signals for each agent when all the activities involved in supplying electricity are performed by a single vertically integrated company and the rest of the agents are pure consumers, mandated to purchase the electricity from that same company, is a relatively straightforward task. The company receives remuneration based on the efficient cost of electricity supply as estimated by the regulator. The end consumers cover this cost via regulated electricity tariffs. The prevalent structure of these tariffs is very simple, with no time or spatial differentiation, since it is designed for price-inelastic consumers with similar and stable consumption patterns. However, it is possible to maintain this traditional regulatory framework while exposing customers to more differentiated (granular) economic signals that would allow them to respond by modifying their demand patterns and installing storage or generation behind the meter, as any other customer could do in a more market-based regulatory environment. Improved economic signals could also enable the provision of electricity services by DERs to the incumbent company. Large customers of traditionally regulated companies have long been subject to much more granular signals in the price of electric energy, charges for peak demand (either for demand coincident with the system-wide peak or for individual peak loads), deviations with respect to pre-established consumption patterns, and interruptibility contracts related to firm capacity services. In some jurisdictions more advanced tariff schemes have been implemented for small and medium customers as well: time-of-use (TOU) pricing, critical peak pricing (CPP) and, very rarely, real-time pricing for the energy component of the tariff.

In power sectors that have been restructured and liberalized at the wholesale level, the activities of wholesale generation, transmission, system operation, and distribution are unbundled. Therefore, the prices and charges for the services provided among these activities and to end consumers have to be established either by specific markets that reveal the prices or by regulated charges. Retail liberalization brings another layer of complexity, since this enables end consumers to choose their suppliers. This forces the regulated electricity tariff to be reduced to the "access tariff," i.e., the regulated component of the tariff that is meant to recover the regulated cost components: transmission and distribution network costs and policy costs to which all consumers should contribute. Costs attributed to the competitive components of electricity supply — electricity production and the commercial activity of retailing — are then freely established by the independent retailers and incorporated in the "integral tariffs" (i.e., those that include all components) that retailers offer to the customers. In most jurisdictions there is an "integral default tariff" designed by the regulator as a default option for customers who do not want to shop around for a better option. Again, in most jurisdictions, both the default and the competitive integral tariffs are very simple, with low granularity levels.

This study confronts the ultimate layer of complexity, which occurs when any agent connected to the power grid at any voltage level cannot be simply classified as either a consumer or a generator, and each agent has the capability of providing — by itself or in aggregation with other agents — all kinds of electricity services to any of the other agents in the power system (these may include transmission or distribution system operators, producers, consumers, or any combination thereof). In such circumstances, which may come to characterize the power sector in the next decade or beyond, the need for more sophisticated economic signals becomes apparent. Regulation must anticipate this scenario (which is already incipiently present in some power systems) with a sound approach to the design of prices and charges.

4.3.1 A cursory review of current electricity tariff practices[9]

While complex wholesale and "ancillary services" markets have evolved in many jurisdictions that capture the marginal price of electricity services in varying degrees of detail, the electricity tariffs that are seen by end consumers and DER owners have been uniformly less complex to date. This is particularly true for residential

9 The profound changes in tariff design that the presence of DERs requires — and that are made possible by the availability of advanced meters — suggest it is not worthwhile to spend much time dwelling on the analysis of present electricity tariffs.

and small commercial and industrial (C&I) customers. These tariffs, which form the system of prices and charges that incentivize distributed actors and energy consumers to make countless operational and investment decisions, typically lack sufficient granularity or accuracy to incentivize efficient response across three key dimensions: time, location, and type of service.

In most countries to date, residential and small C&I customers receive monthly or bimonthly electricity bills that consist of a small fixed component (a per-customer charge in dollars per month or per year) and a volumetric component (in dollars per kilowatt-hour consumed). The volumetric component may include a separate generation or supply rate that captures the cost of electrical energy and an access (or delivery) rate that recovers transmission and distribution network costs, which are often bundled into a single charge, and policy costs. The generation or energy rate is typically an average per-kilowatt-hour (per-kWh) rate reflecting the average cost of energy plus reserves and firm capacity (including any hedging performed for that customer). The access or delivery rate is an average per-kWh rate that divides the total costs associated with transmission, distribution, and policy costs by the total kWh of electricity consumption. These rates reveal little to nothing about variations in generation or network costs arising from differences in location, scarcity in generation or network capacity at different times, or other system operating conditions.

In some cases, crude signals about network capacity utilization (in dollars per kilowatt) are sent by charging customers in proportion to their contracted capacity (in those few jurisdictions where contracted capacity exists for residential and small C&I consumers). Under such rates, if customers exceed their contracted capacity, their service is temporarily discontinued or some penalty is applied (if the excess power can be metered). Specific demand charges are also employed in some jurisdictions, reflecting the user's own peak consumption during each month or other period (again, if it can be metered), even though this individual peak in consumption is unlikely to be coincident with the system-wide or local peaks in power withdrawal (or injection) that drive network and generation capacity costs.

Time-of-use (TOU) tariffs have been implemented in multiple service areas, relying upon interval metering that distinguishes between coarsely defined use times, such as daytime versus nighttime (or peak demand time versus off-peak demand time), with distinct prices applied during different times. These tariffs provide a basic level of temporal granularity and have encouraged efficient investments, such as off-peak electric heating with thermal accumulators. The obvious shortcoming of TOU tariffs is that the time intervals in which the different prices and charges apply are established long in advance, are fixed for long periods of time (e.g., a season or year), and may not correspond to the actual conditions in real time. This renders the temporal incentives provided by these tariffs inefficient in most cases. In addition, many TOU rates continue to bundle generation and network costs into the (time-varying) volumetric rate, which distorts incentives for network usage.

Critical peak pricing (CPP) has been successfully applied in some power systems, and it requires some means of one-way communication with the customers. On a limited number of occasions per year, the system operator can inform customers with some prescribed anticipation that a critical period will occur during which the value of the tariff will be particularly high. If designed well, these programs can effectively signal to the power system agents their marginal contribution to system costs during the few hours of the year that drive the largest share of system costs. CPP tariffs, however, fail to capture more regular, hour-to-hour, and day-to-day variations in the price of electricity services, and may lack sufficient locational granularity.

Large C&I customers typically receive more granular price signals through retail rates that account not only for volumetric energy consumption but also for peak demand (either coincident or not) as well as network connection and use-of-system costs. They often also face tariffs that feature variations in costs across different times of use and customer locations. Large C&I customers are equipped with interval meters that can measure their electricity utilization in a detailed manner; therefore, these users can more readily alter consumption and distributed generation profiles in response to price signals. Current pricing practices for large C&I customers provide valuable guidelines for constructing a system of retail-level economic signals that can enable efficient utilization of DERs in electricity service provision. Indeed, large C&I customers receive more granular prices precisely because they have historically had more options for responding to prices, which increases the costs of ill-designed tariffs. The same rationale is now becoming true for smaller customers as well, as opportunities for DER adoption and price-responsive or flexible demand proliferate.

Concurrently, the advent of advanced meters that measure electricity injections and withdrawals on hourly or shorter time frames at each network connection point enable evolution and improvement in the design of prices and charges for a broader class of end-users of electricity services. Existing information and communication technology (ICT) capabilities allow the computation of more granular prices and charges, accounting for the actual system operating conditions and applying prices and charges to the individual injections and withdrawals at any given time and connection point. Some power systems, such as that in use in Spain, have started to apply wholesale day-ahead hourly energy prices to residential customers in the integral default tariff.

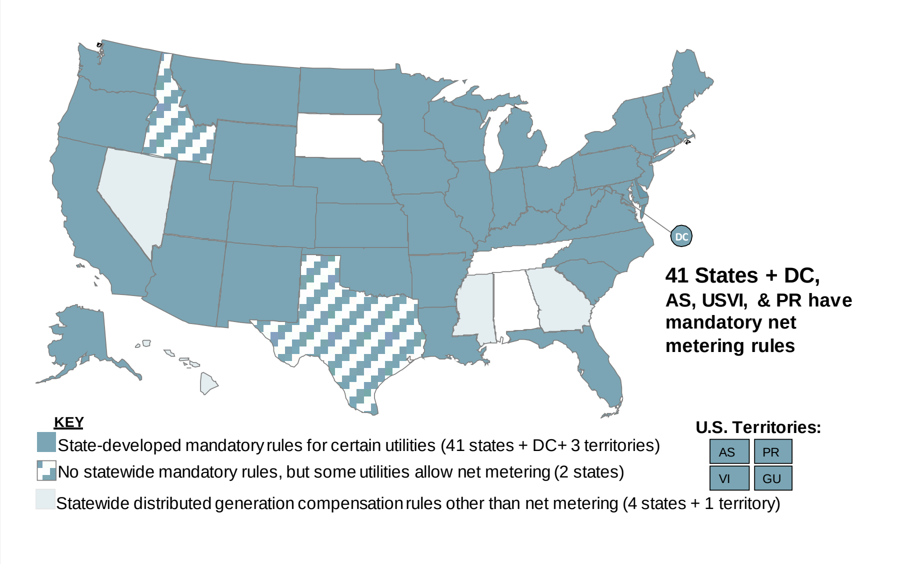
Box 4.1: Net Energy Metering and Volumetric Rates

The evolution of net energy metering (NEM)[10] policies and their impact on power systems provides an instructive example of the importance of designing incentives that anticipate — and provide flexibility to adapt to — developments that can rapidly evolve to have significant unforeseen impacts. While net metering began as an apparently harmless form of support to distributed generation, the rapid growth of DER capacity incentivized by NEM regulation can have adverse outcomes that must be rectified before they more significantly impact power systems.

NEM is widely used in the United States and also in some European countries. **Figure 4.3** shows the net-metering policies that were in place in 41 states of the United States as of July 2016. In seven European Union countries, electricity customers can net their demand with electricity produced from solar panels any time during that same year (European Commission 2015).

10 NEM is an approach used for the compensation of distributed generation and other DERs. NEM charges system users for net energy consumed (that is, energy consumed minus energy generated during the time period) during each netting period, typically one or two months (note that, the longer the period, the more distortional the netting is).

Figure 4.3: Net Metering Policies in Place in the United States in July 2016



Source: www.dsireusa.org

Under NEM, payment for any costs recovered via volumetric charges in the electricity tariff — including energy, network, and policy costs — is reduced or entirely avoided by distributed generation (DG) users, leading to cross-subsidization of DG network users by customers without DG, since only energy costs should be netted by DG production. NEM plus volumetric tariffs asymmetrically value behind-the-meter generation above electricity at the same location that does not come from behind the meter. Also note that DG production typically peaks at a different time than demand, and therefore customers with DG typically use the network more than customers without DG. In addition, injections of DG power may stress the network and lead to increased distribution network costs (MIT 2015). The full recovery of network and policy costs requires an increase in volumetric rates, further encouraging the deployment of DG.

How can this criticism of NEM be compatible with our basic principle that "cost-reflective prices and charges should be based only on what is metered at the connection point," i.e., the net of the production and consumption of all the devices behind the meter? The problem with NEM, as it is used in many power systems today, is twofold. On the one hand the netting is done over a long period of time, as required by the standard meters still in use today in most power systems, so that the injection of power into the grid from a solar panel at noon on a sunny day when the energy price is low is netted with the consumption at peak demand time on a cold evening when the energy price can be much higher. On the other hand, when the volumetric component of the electricity tariff includes network and policy costs, these costs are avoided when the metered injection is netted with the metered consumption.

Netting the internal generation and demand of all the devices behind the connection point to the network during a short time interval (one hour or less) is what we propose. However, the actual implementation of NEM — over long periods of time and accompanied by volumetric tariffs that include network and policy costs — introduces serious distortions in the tariff system.

4.4 Design of Prices and Charges: From First Principles to Implementation

As discussed at the beginning of this chapter, the main objectives of a comprehensive system of prices and charges are to promote efficient outcomes (e.g., to minimize the cost of desired electricity services) and to recover regulated network and policy costs. In this section, the fundamental principles of power system economics and engineering are applied to provide the framework for an "efficient ideal" approach to designing prices and charges — that is, an approach based on achieving the most economically efficient outcome, while guaranteeing the recovery of the regulated costs.

As indicated in the introduction, reading this section can be omitted by those who are just interested in the principles (**Section 4.2**) and the recommendations (**Section 4.6**) as well as by those with a good knowledge of the design of prices and charges in a market environment, although such readers may wish to at least skim the sections on firm capacity (**Section 4.4.3**), allocation of network charges on the basis of responsibility in new network investments (within **Section 4.4.4**), and the treatment of residual regulated costs (both residual network and policy costs) (first part of **Section 4.4.5**).

In practice, the level of granularity in space, time, and the disaggregation of services reflected in these economic signals will depend on several factors, including: trade-offs among the efficiency gains associated with increased granularity; the availability of communication and computational technologies and implementation costs necessary to increase granularity; and the acceptance of different prices and charges by various agents. Therefore, alternative simplified approaches must also be considered. Where possible, we present a continuum of options and underscore the key trade-offs relevant to evaluating alternative designs for prices and charges (see **Section 4.6** on implementation).

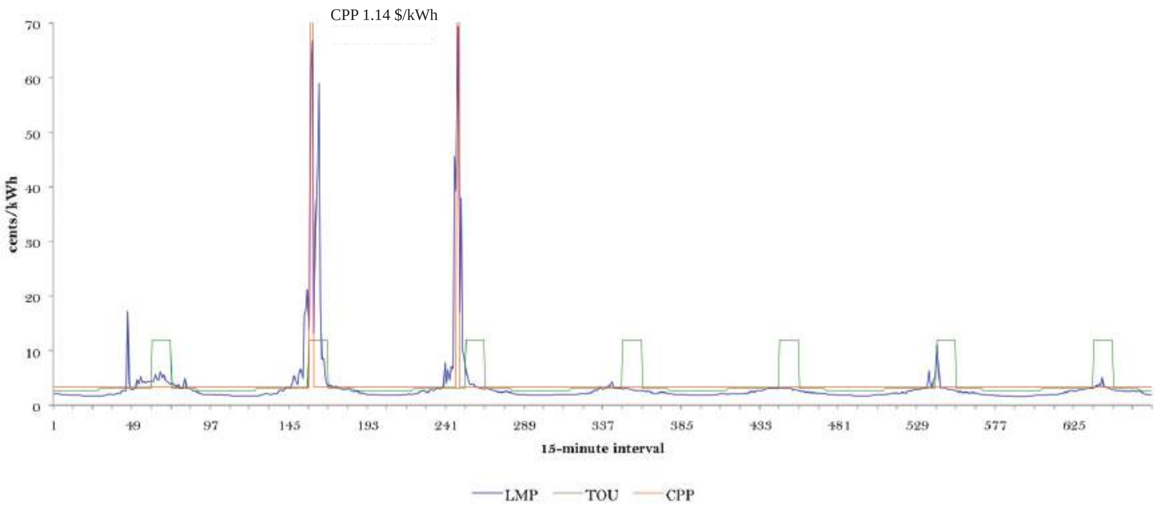
All economic signals at all voltage levels are relevant and will be considered. However, our focus will be on the prices and charges that apply to electricity customers and DERs connected in the distribution network and the manner by which the more advanced economic signals typically available at the wholesale level can filter down to all voltage levels in distribution. As explained in **Chapter 3**, in practice one can expect that — at least for some time — intermediary agents or aggregators will manage the response of residential and small I&C customers. Such aggregators can receive the signals of prices and charges for all the relevant electricity services corresponding to all their customers and make use of diverse estimates or forecasts of demand, weather, plant availability, state of charge of storage, network flows, and other information in order to decide how to participate in the several markets for services. This participation may include making use of the responsive demand and DERs of its customers, sending to each one simple signals to obtain efficient responses at all times, and maintaining prespecified levels of comfort or meeting other customer needs. A likely scenario is that competing aggregators will offer retail customers lower flat prices in return for being able to manage and automate a defined degree of flexibility in the timing of DER demand and operation.

The following sections will cover the economic signals related to key electricity services (e.g., those services responsible for the largest share of power system costs), making use of the breakdown of prices and charges that was presented in **Section 4.2**. First, we will discuss the prices for electric energy, operating reserves, and firm capacity. Next, we will discuss methods for determining forward-looking cost reflective network charges and policy costs charges and finally how to allocate the residual regulated costs (both network and policy residual costs).

4.4.1 The price of electric energy

The economically ideal price for electric energy given current power system conditions is embodied in the "nodal" or locational marginal price of electric energy at each point of connection and at each moment in time, calculated on the basis of the costs of supply and the demand in response to these prices (Schweppe et al. 1988). Locational marginal pricing is the ideal that power systems should try to achieve, while taking into consideration the implementation costs and the incremental benefits of progressively increasing the temporal and spatial granularity of energy prices.

Figure 4.4: ERCOT Wholesale LMP, Austin Energy Time-of-Use Rate, and Illustrative Critical Peak Pricing During One Week in July 2015



Accurately computing locational marginal prices (LMPs) at the connection point of each agent anywhere in the grid is a long way from current practices in tariff design (see **Section 4.3**). Yet, any sensible simplifications are possible, and our objective in this chapter is to understand the implications of each option and to make a choice that brings more benefits than costs in each particular context.

There are three major possibilities for simplifying the ideal reference LMP: (1) to make use only of the active (also known as "real") power component of the LMP, ignoring reactive power in energy prices[11]; (2) to use average values of LMP over a period of time, which could range from as little as five minutes to as much as one or more years; and (3) to use average values of the LMP over a geographical portion of the network, which could typically range from the individual connection point of each agent (i.e., no simplification) to the entire considered power system, with intermediate options (e.g., applying average LMP values at the head of the low-voltage feeder, distribution zones with homogeneous network characteristics, individual nodes of the transmission network, or zones of the transmission system where the nodal LMPs typically have similar values at all times). While simplifications in any of these three dimensions are unavoidable in practice, this section will illustrate that some level of granularity is necessary to capture significant differences in the marginal cost or value of energy.

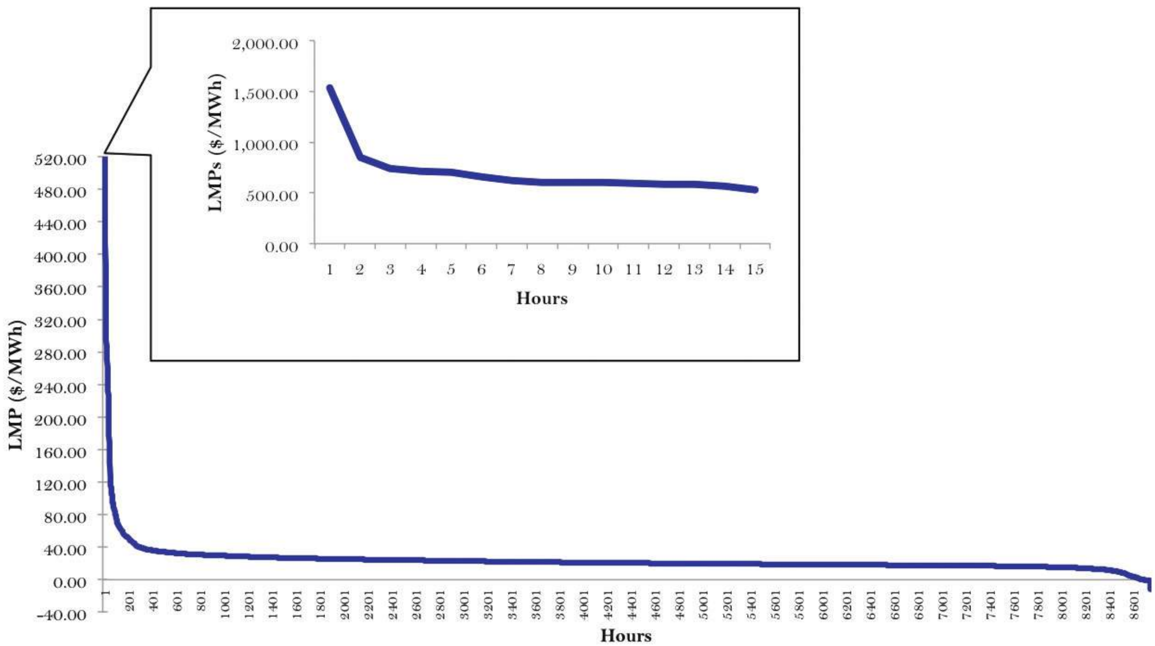
4.4.1.1 The temporal dimension

First, the marginal cost of electrical energy varies with time, and it can do so significantly. These variations arise from changes in load patterns and generation costs. The changes in the cost of energy over the course of a day or other time period are incompletely reflected in or approximated by existing rate structures like flat annual or multiannual rates, time-of-use (TOU) rates, and critical peak pricing (CPP) tariffs.[12] As an illustrative example, **Figure 4.4** shows one week of zonal settlement prices in the Austin Energy load zone of the Electric Reliability Council of Texas (ERCOT) system. The settlement price

11 The instantaneous power (product of the voltage and the current at any given time) can be split mathematically into a component that has a positive average value (so called "active power") and another one ("reactive power") that withdraws and then delivers power in equal amounts, so that the average value is zero. This reactive power component goes back and forth between the supplier and the receiver, only heating the wires and other network components or contributing to the violation of some network constraint, with no useful purpose other than being an inseparable part of the actual complete power.

12 Under TOU rates, customers are charged different static, predetermined rates in each of multiple time blocks (e.g., different days and hours of each day). Once set, the TOU prices are not updated or allowed to fluctuate for some fixed period. TOU rates are often coupled with a peak/off-peak demand charge. Under CPP or variable peak pricing, customers are charged according to whether or not the retailer or distribution utility has identified a particular hour or set of hours as "critical," with loads ranging from low, to standard, to high, to critical (for a particularly high load). Customers receive day-ahead or few-hours-ahead notice of the anticipated occurrence of a critical period during which they will be charged a CPP rate. The CPP rate is set in advance (i.e., it does not vary in real time).

Figure 4.5: Relative Frequency (in Time) of ERCOT Wholesale LMPs Over One Year (2015)

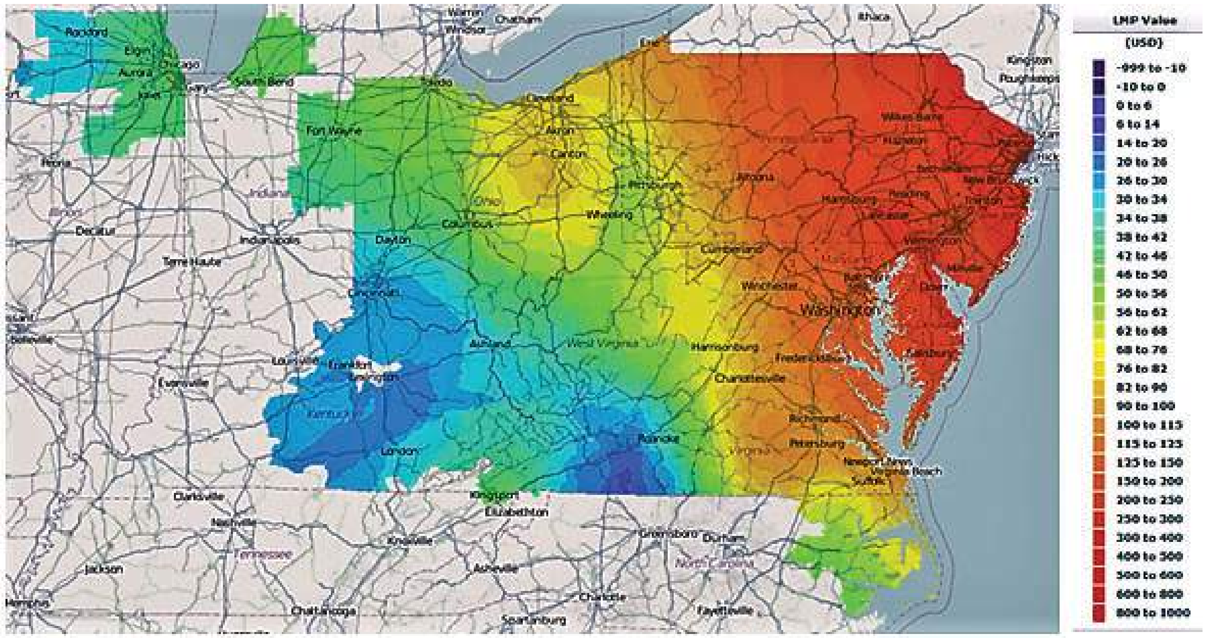


is a load-weighted average of the LMPs at all nodes within the load zone and also an average of the prices computed every five minutes into 15-minute intervals (it thus already includes some degree of simplification along both temporal and spatial dimensions). **Figure 4.4** also shows a TOU rate offered by Austin Energy, as well as an illustrative CPP that is consistent with the LMPs in that the designation of critical event hours coincides reasonably well with the actual occurrence of peak load hours. Although the TOU and CPP rates are derived from predicted LMP values (through estimates of peak load hours, contingencies, and other events that will affect LMPs), the economic signals received by the customer from TOU and CPP rates can differ significantly from the actual short-term marginal prices.

The mismatch between temporal simplifications for energy prices, such as flat or TOU rates and the true marginal cost of energy, is particularly apparent at times of either generation scarcity or network congestion. During scarcity events or other periods when wholesale energy prices spike (sometimes an indication of a network constraint, signaling the need for network investments), significant cost savings could be achieved via relatively small reductions in energy use. These cost-saving opportunities are not realized, however, with flat energy prices or most TOU rates. On the other hand, CPP schemes can reflect most of the cost variation and capture most of the available response to such price variations, if the identified "critical" peaks coincide with the actual periods of generation or network capacity scarcity. **Figure 4.5**, which depicts the relative frequency of the value of the hourly LMPs in the Austin Energy load zone of the ERCOT system, indicates that, while instances of price spikes can drive LMPs to very high values, such spikes are rare and thus idiosyncratic. Therefore, spikes cannot easily be estimated a priori or approximated by TOU rates. This is also an indication that even hourly energy prices reflecting day-ahead market clearing prices — e.g., computed 12 to 36 hours in advance — may fail to reflect the actual operating conditions of the power system.[13] This issue is treated in detail in **Chapter 7**.

13 Most so-called "real-time" pricing programs do not reflect the real-time settlement prices, but rather convey day-ahead clearing prices to end-users.

Figure 4.6: Wholesale LMP Variation Across More Than 11,000 PJM Nodes on July 19, 2013, at 4:05 p.m.



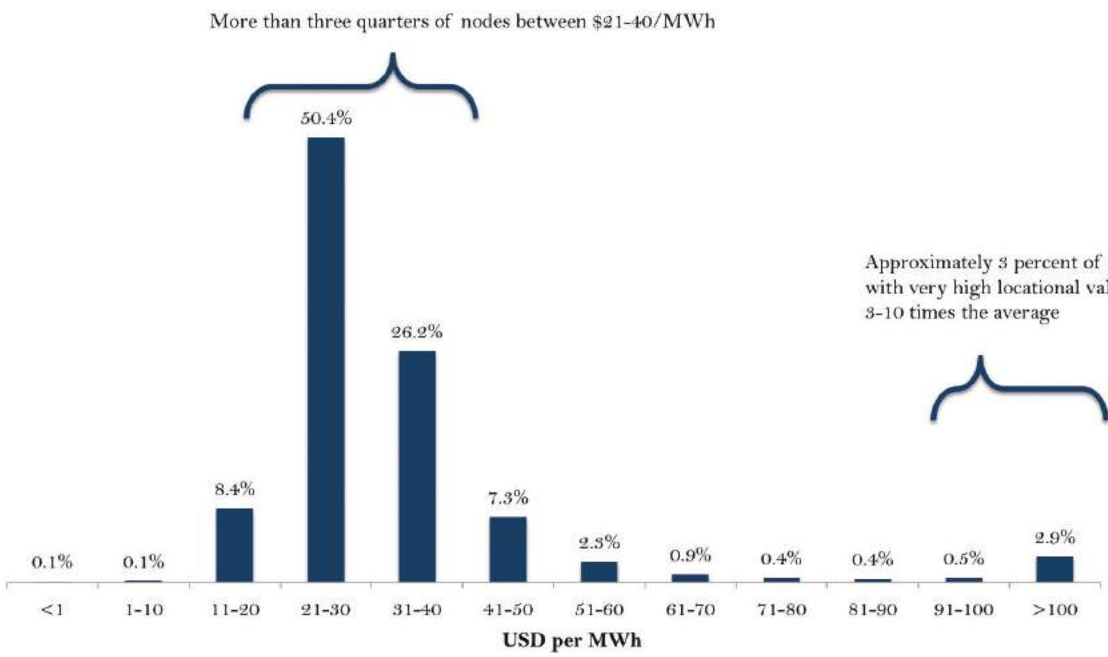
4.4.1.2 The spatial dimension

The marginal cost of electrical energy also differs by location within a network. These differences arise from the presence of losses within transmission and distribution lines, and from the occurrence of active network constraints that limit power flow in sections of the network.[14]

At any given time, the differences in LMPs among the nodes of a transmission network can be very significant, particularly when there are active network constraints. At the transmission level, the differences in LMP values due to losses may be in the realm of 5 percent to 10 percent, depending on the size of the power system and the specific operating conditions. However, as **Figure 4.6** shows for the PJM power system, vast differences in LMP (due to both losses and congestions) may exist at any given time. In this case, prices in the eastern portion of the PJM system are an order of magnitude higher than in the western portion, illustrating the importance of spatial granularity in energy prices, particularly at times of binding network constraints.

14 These constraints are commonly termed “congestion,” although voltage problems may constrain the flow in the network without any particular network element being “congested.”

Figure 4.7: Frequency Distribution of 2015 PJM Average Nodal LMP Over One Year



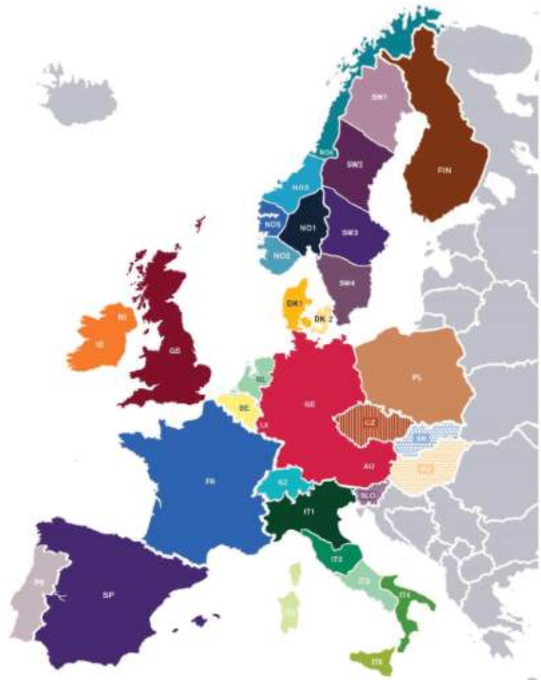
The situation depicted in Figure 4.6 is not an unusual one. **Figure 4.7** shows the distribution of the annual average LMP value at each locational node of the PJM system in 2015. This distribution provides clear evidence of the need for a locational characterization of the energy prices in the PJM system. Similar differences in the marginal price of energy at different locations are observed in other power systems.

Other power systems have adopted spatially simplified approaches to LMPs, such as zonal prices (with each zone comprising many transmission nodes that are assumed to share the same energy price) or the use of a single, uniform energy price throughout an entire transmission region. In the European Union, a uniform zonal price for energy is computed for each one of the bidding zones shown in **Figure 4.8**. These zonal prices are used instead of LMPs and are cleared via coordination among different power exchanges (as distinct from system operators), neglecting network constraints within each bidding zone. After prices are cleared, internal transmission constraints are subsequently solved through diverse and uncoordinated market mechanisms. This approach improves the tractability of price computation by limiting consideration of network constraints to those arising from a very simplified representation of cross-border flows, which has facilitated the integration of multiple national markets over a large geographical area into a common trading platform on the timescales required for market operation, reducing horizontal market concentration and enhancing market liquidity considerably in the European Union.[15]

LMPs, and to a lesser extent zonal prices, are effective locational signals for operation and investment decisions at the wholesale level. **Chapter 7** elaborates on the advantages and difficulties of adopting LMPs in large wholesale electricity markets such as the ones encountered in the United States and in the European Union. It should be highlighted that, despite all the effort bestowed on the computation of nodal or zonal prices in these large markets, the spatial granularity is lost for LMPs — and sometimes also for zonal prices — when the economic signal is passed through to the agents connected to the diverse voltage levels in the distribution networks. Note that, when constraints occur within a zone (rather than between zones), zonal average prices may hide the fact that high prices on one side of an internal constraint are offsetting low prices on the other side of the constraint in the same zone. While there might be zones with few internal constraints, presenting temporal variance only at the zonal level may dull the impact of binding constraints on price-responsive demand and DERs that could contribute to an efficient response.

15 The following information should help in establishing comparisons between electricity markets in the United States and the European Union. All power systems within the European Union operate under common market rules on a common trading platform — the Internal European Market coordinates decisions across an area covering a population of more than 500 million people in the EU-28 and an aggregate consumption of 3,220 terawatt-hours (TWh) in 2013 (Eurostat 2016) — but each country typically has one system operator (though some have several). The United States encompasses multiple power systems with diverse regulatory frameworks, ranging from traditional vertically integrated electricity companies under cost-of-service regulation to liberalized markets operated by independent system operators. The continental US power sector (excluding Alaska) is, in practice, physically separated into three systems: the Eastern Interconnection, the Western Interconnection, and Texas. The largest of US electricity markets, PJM, serves an area of 61 million people with annual energy consumption of 793 TWh (FERC 2016). Germany, the largest national power system in Europe (actually comprising three system operators), has an annual energy consumption of roughly 575 TWh. The US power systems of California (CAISO), New York (NYISO), or New England (ISONE) have "electrical sizes" of the same order of magnitude as the power systems of large EU countries such as France, Italy, Spain, or Poland.

Figure 4.8. Bidding Zones in European Market Coupling



Source: Ofgem 2014

From the viewpoint of "getting the prices right," much of the effort bestowed on the computation of nodal, and even zonal, energy prices in wholesale electricity markets is lost when the economic signal is passed through to the agents connected to diverse voltage levels in the distribution networks.

4.4.1.3 Distribution LMPs

At the transmission level, system operators and market platforms have vast experience in computing LMPs. LMPs have been employed in the United States and also in other countries for more than a decade (Chile started using a simplified version of LMPs in 1981 and Argentina in 1992). On the contrary, as far as we know, no power system uses distribution LMPs today, although serious research has been done on the topic (Caramanis et al. 2016; Ntakou and Caramanis 2014). This does not mean that distribution LMPs could not be useful. Rather, advanced meters, inexpensive ICT solutions and DERs that could use distribution LMPs effectively have become available only recently. Therefore, the question is: How important is it to account for spatial variations in distribution? Let's examine losses and congestion separately.

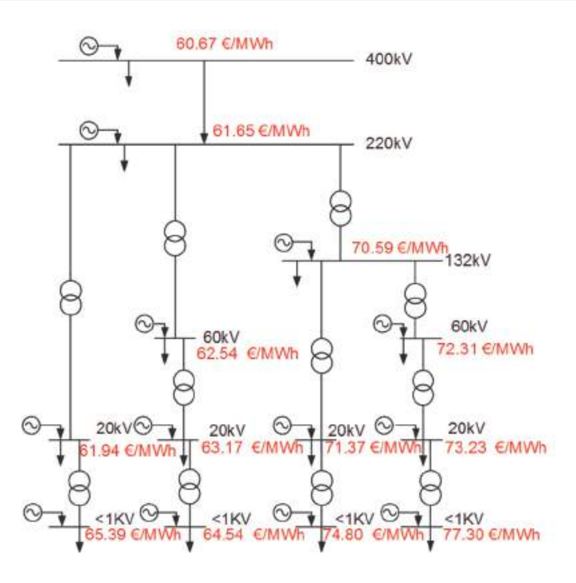
Using actual loss data from the Spanish power system in an aggregated fashion, it was possible to estimate the "marginal loss effect" for each voltage level, i.e., the factors that, multiplied by the wholesale market price, allow us to estimate the average LMPs at the high-, medium- and low-voltage levels of the distribution network at a given moment in time and also as an average over a year or other time period. Over the course of the year, LMPs in distribution average 11 percent to 17 percent higher than the wholesale LMPs for medium and low distribution voltages respectively, although at particular moments, when demand levels are high, the effect of marginal distribution losses on LMPs can be as high as 40 percent for low-voltage levels. Other references also report marginal loss factors that indicate relevant differences in the value of energy consumed within distribution networks that is not captured in LMPs calculated for wholesale nodes (see **Box 4.2**). These numbers change with different penetration levels of distributed generation in the network. In any case, these variations create a different — and typically higher, for moderate penetrations — local economic value for distributed generation than for generation connected at the transmission level.

Box 4.2: Marginal Effect of Losses on Locational Marginal Prices in Distribution Networks

Using actual losses and flow data from the Spanish power system in an aggregated fashion, it is possible to estimate the "marginal loss effect" for each voltage level — e.g., by using an electric equivalent network of the whole system and a quadratic losses approximation (see **Appendix A** for a detailed description of the ROM model used for this computation). The streamlined network in **Figure 4.9** is equivalent to the complete network in terms of the share of losses at the different voltage levels. This network is used to estimate the differences in LMPs between voltage levels due to marginal technical losses. Average yearly differences of as high as 27 percent are shown to exist between the high-voltage transmission and a low-voltage distribution node (the one with the average marginal price of 77.30 euros in **Figure 4.9a**), which can make all the difference in the viability of investing in a given DER or in deciding whether to operate it or not at a given time. Note that in the operating conditions shown in Figure 4.9a, lower voltage levels present higher LMPs. This is due to the higher resistance of distribution lines at low-voltage levels. When considering the quadratic losses representation, yearly average differences in distribution-level LMPs with respect to the wholesale LMP are 11 percent and 17 percent when averaged for all nodes of medium- and low-voltage distribution levels, respectively. Similarly, Ausgrid (2011) reports marginal loss factors during peak load periods as high as 33 percent to 43 percent in Australia for the years 2000–2008 in specific distribution networks. For 2016, average loss factors for certain low-voltage distribution networks vary from 5 percent to 16 percent (AEMO 2016). Higher marginal price levels at low-voltage levels occur due to the marginal losses effect, as most of the demand is being located at low-voltage levels and generation at high-voltage levels. But, this effect may change with the high presence of distributed resources and reverse power flows from low-voltage to high-voltage levels. For instance, **Figure 4.9b** depicts the one-hour snapshot for the same Spanish case, but with high solar photovoltaics (PV) at low voltages (in certain nodes higher than the nodal demand); in the represented hour, PV generation is even curtailed. This situation leads to zero locational prices at certain parts of the grid (including some low-voltage nodes) but prices higher than 30 euros per megawatt-hour in other nodes, including those at high-voltage levels (since the curtailed generation is not enough to supply an increment of demand at those nodes with prices higher than zero while accounting for associated losses and the need for must-run units). This example shows that the value of distributed resources (in terms of energy loss reduction) reaches a limit at high penetration levels.

Figure 4.9: Distribution Locational Marginal Prices when Aggregated Distribution Losses are Considered in the Spanish Power System

4.9a: Yearly average prices for the current Spanish system



4.9b: Marginal prices for 9 a.m. during a day in April when the Spanish system had a penetration level of solar photovoltaics connected at low voltage (under 1 kV) providing 35 percent of total generation capacity

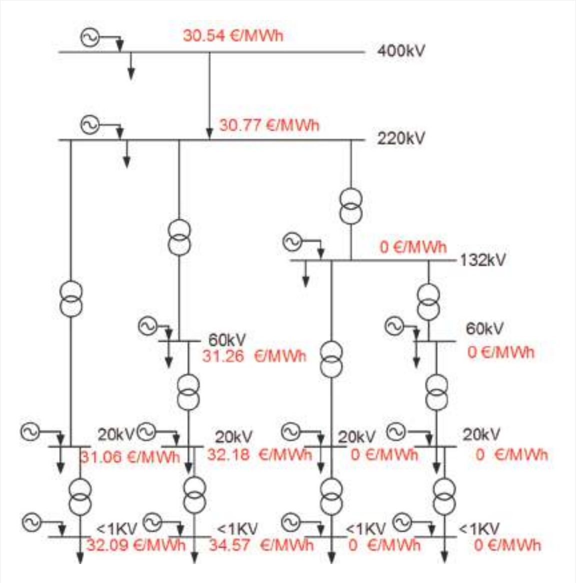
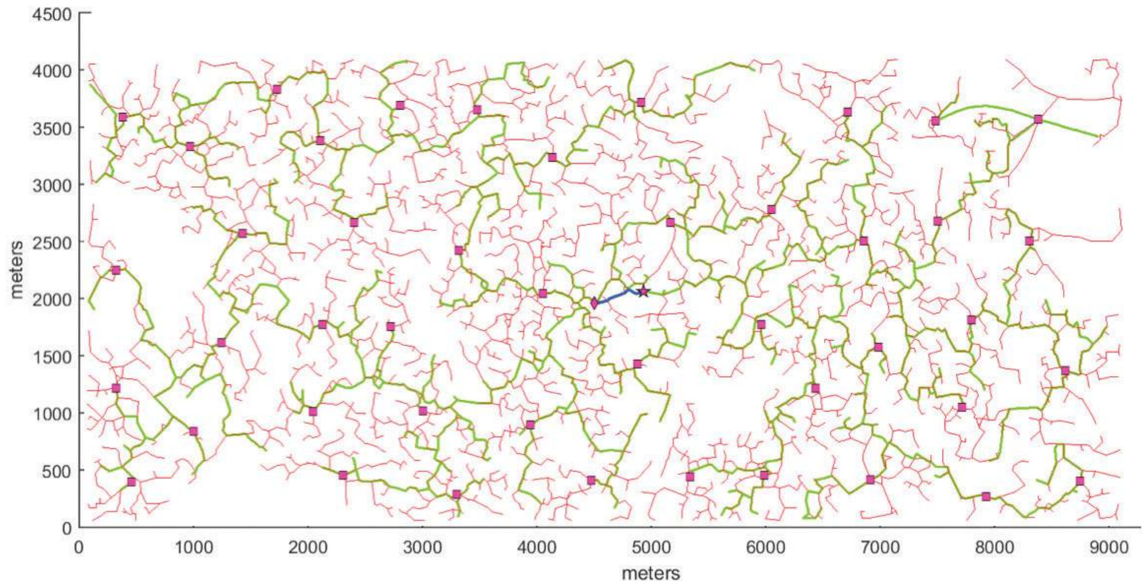


Figure 4.10: Layout of Urban Distribution Network Used in Price-Computation Simulations



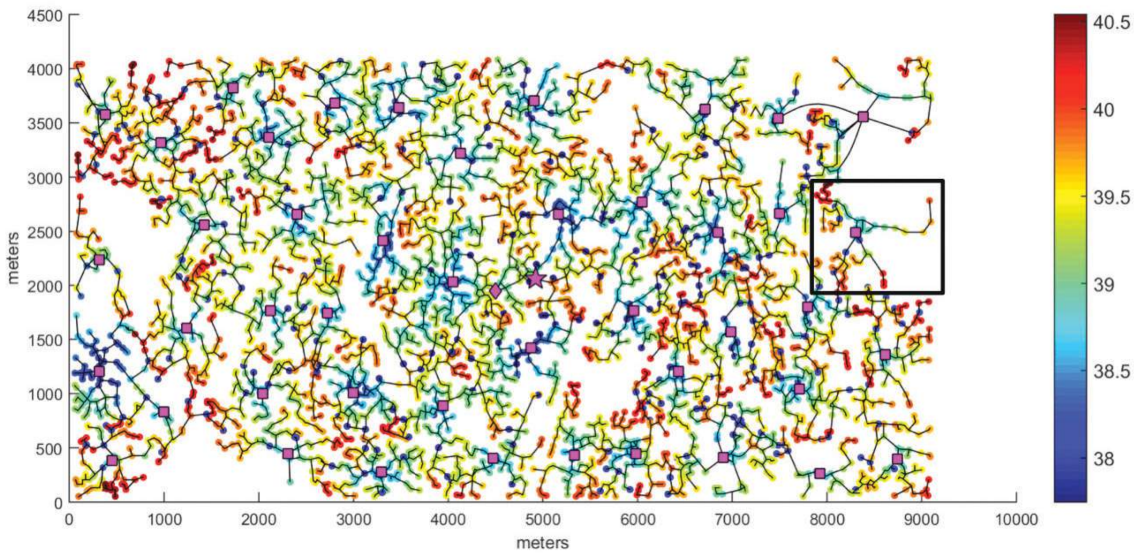
Next, with a more detailed model of distribution system operations, we can characterize the behavior of LMPs in distribution networks under a diverse range of conditions. **Figure 4.10** shows the layout of an urban distribution network (designed according to a street map of a portion of Austin, TX) that we use to illustrate the computation of distribution LMPs.[16]

Figure 4.11a illustrates the spatial distribution of LMPs in the network shown in **Figure 4.10** in the absence of any network congestion; that is, with differences in LMPs arising solely from network losses. **Figure 4.11b** provides a close-up view of the low-voltage (LV) laterals connected at a single medium-voltage (MV)/LV transformer, illustrating the price variation across nodes along MV and LV lines. As expected, the maximum differences in nodal prices along the MV and LV feeders relative to the transformer occur at the ends of the LV laterals — i.e., farthest from the transformer. In Figure 4.11b, the increase in nodal price as we move farther along the LV lines (away from the transformer) reflects the marginal cost of distribution line losses. At the relatively light loading of the network used in this example, the range of spatial differences in prices along radial LV lines reaches just over 5 percent.

16 Note that all connections between medium-voltage and low-voltage lines occur at MV/LV transformers (pink squares). What may appear, in the figure, to be connections between MV and LV lines without a transformer are in fact shared rights-of-way by MV and LV lines that, when plotted in the diagram, give an appearance of line discontinuity.

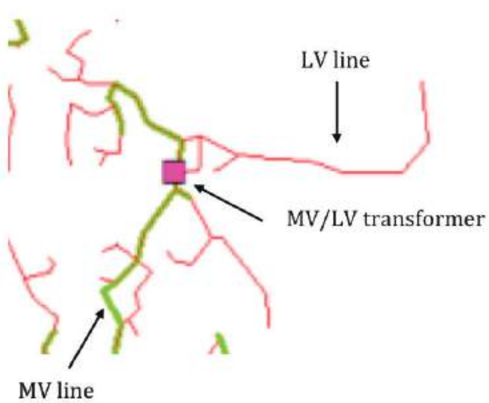
Figure 4.11: Spatial Variation in Distribution-Level Active Power LMPs Caused by Network Losses in the Network of Figure 4.10

4.11a: Distribution-level active power nodal prices (in $/MWh) across the entire considered network during a single hour



4.11b: Distribution-level active power nodal prices along a small fraction of the network connected at a single MV/LV transformer

Topology of zoomed-in network section



Active power nodal prices of zoomed-in network section

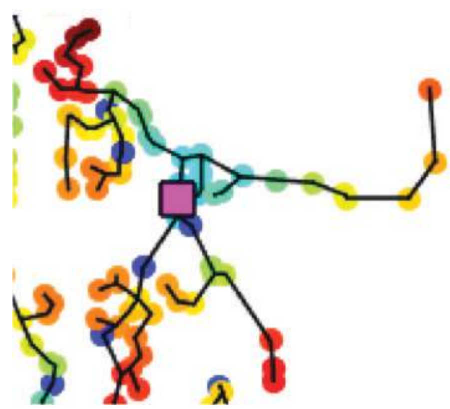
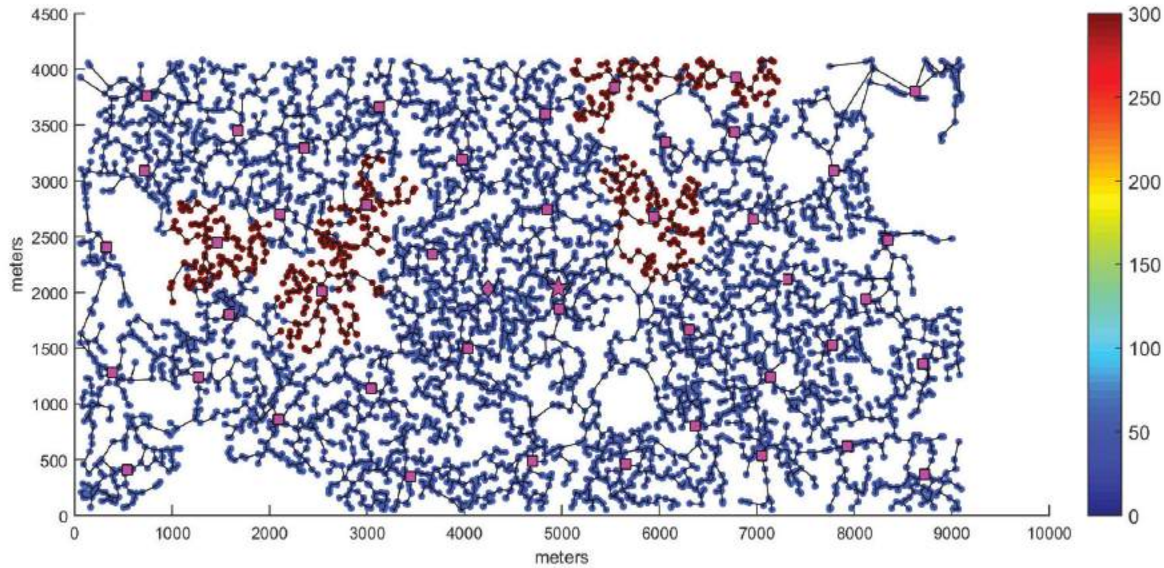


Figure 4.12: Spatial Variation in Distribution-Level Active Power LMPs Caused by Network Congestion in the Network of Figure 4.10



Most LMPs are in the range of $35 to $40 per megawatt-hour (MWh), while LMPs at congested nodes rise to the cost of nonserved energy of $300/MWh since load must be curtailed to abide by constraints on transformer capacity.

Next, the same network is examined in the presence of congestion. Currently, congestion is generally uncommon in distribution, as the only practical solution for a violated physical network constraint — once transformer tap changers, voltage regulators, and network reconfiguration have been employed — is load curtailment on the concerned feeder. Network capacity planning margins are therefore typically generous, and with low load growth, congestion in existing networks may be sufficiently infrequent that prices will not capture congestion-related signals in many distribution networks. In the future, however, if price-responsive DERs become commonplace and/or demand becomes more price-responsive, networks may be more closely adapted to the actual needs of network users,[17] resulting in more frequent congestion, caused by such factors as binding network constraints that limit the response the demand and DERs to prices and charges.

Under such hypothetical future conditions, prices would serve a critical role in communicating scarcity of network capacity to the network users. **Figure 4.12** depicts the spatial variation of active power LMPs for the same distribution network as in **4.10** and **4.11**, but in the presence of congestion (reflected by the load on multiple transformers exceeding the rated transformer capacities). The price at congested nodes (those downstream of a congestion) rises to the cost of nonserved energy (CNSE), or cost of voluntary demand curtailment, which in this example has been set to $300 per megawatt-hour (MWh).

17 The level of adaptation of the network to current system conditions also depends on the existing mandatory technical procedures to be used when planning and designing upgrades. These procedures may have to be adapted to permit utilities to manage distribution networks with binding constraints.

4.4.1.4 Accounting for reactive power

All power systems that currently use LMPs at the transmission level compute these prices through a linearized DC load flow representation of the network, which provides computational robustness but does not account for the production and consumption of reactive power. However, under some conditions, both active and reactive power must be considered to correctly account for losses, operating limits, and voltage differences.

When voltage problems occur in some area of the transmission network, ad hoc arrangements are agreed upon with local generators to provide the active and reactive power needed to keep the voltage within prescribed limits. Currently, congestion within the distribution system is rare and often voltage-related. If congestion becomes more prevalent in future distribution systems because they are operating closer to their limits (to accommodate power withdrawals or injections), it may be efficient to employ price signals to regulate the production and consumption of reactive power — since reactive power can help alleviate congestions. For this to be feasible, agents would need to be capable and willing to supply and/or consume reactive power in response to prices, and this response would have to take place under sufficiently competitive conditions. If such conditions are met, LMPs for both active power (P) and reactive power (Q), could, in theory, be calculated for all nodes in transmission and distribution networks, indicating the marginal cost of production or consumption of active or reactive power. However, it is likely that reactive power prices — due to the local nature of reactive power (i.e., only a few agents can participate) and the high sensitivity of voltage to the response of agents — would be very volatile and lend themselves to the exercise of market power.[18]

Reactive power in distribution networks can be provided and consumed by various devices, including synchronous generators, capacitors, and inverters. The value of reactive power LMPs — as of active power LMPs — is expected to be zero or very small when no constraints are binding (i.e., when price variations arise primarily from losses, which can be far more significant at lower voltage levels than at higher voltage levels). When a constraint becomes binding, the price of reactive power will depend upon the marginal cost to produce the required amount of reactive power to avoid constraint violation. Inverters and synchronous machines can transfer a limited amount of power. This finite capability to transfer some combination of active and reactive power may lead to the need to restrict active power transfer in order to enable reactive power transfer, giving rise to an opportunity cost. It is important to note that in distribution networks there is a strong coupling between active power and voltage. Hence, if the only option to deal with a voltage problem in an area of the network is to curtail active power consumption, the marginal value of providing reactive power will be high, since doing so will save active power curtailment costs. Additionally, more frequent inverter switching operations can also impose wear-and-tear costs.

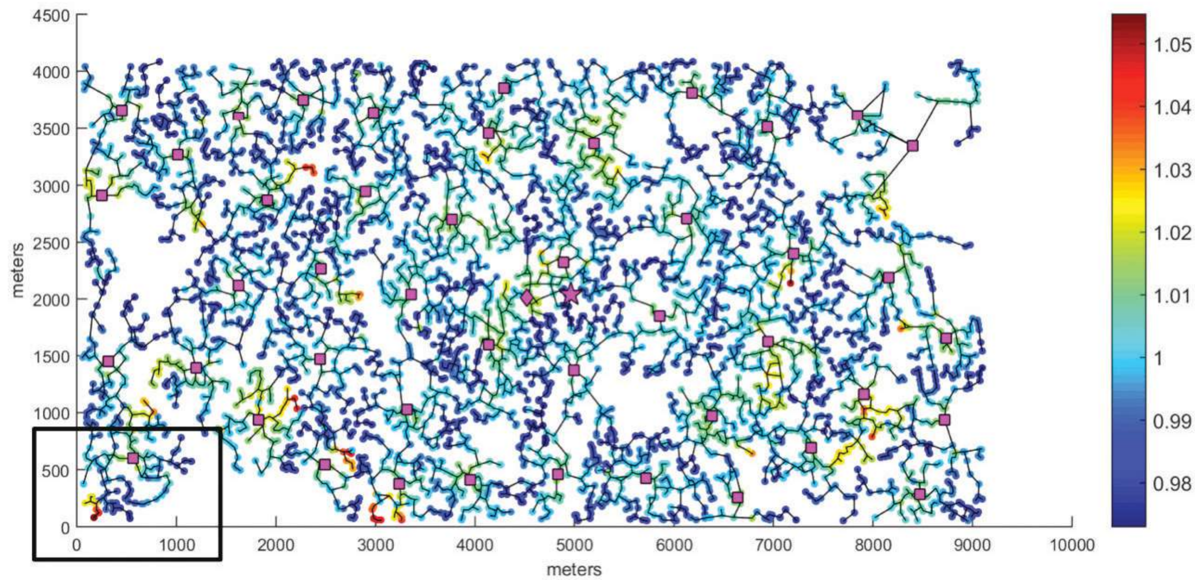
Reactive power LMPs can signal local voltage problems such as feeder overvoltages in sections of the network with high penetration of distributed generation, such as rooftop PV. **Figures 4.13a** and **4.13b** display nodal voltages (expressed per unit[19]) and reactive power LMPs, respectively, in the same urban distribution network that was used in previous examples. In this case, a low-voltage feeder in the southwest region of the network, shown in **4.13c**, is experiencing a rise in voltage due to a high penetration of distributed PV. Consequently, the reactive power LMPs are most negative at the nodes experiencing the greatest voltage rise (**4.13d**),

18 In the future, many of the devices that have inverters, such as solar panels and batteries, could provide voltage control and reactive power compensation in distribution systems (see Kroposki 2016 and Moghe et al. 2016). These devices could obtain some compensation when committing their capabilities to the system and mitigating price volatility.

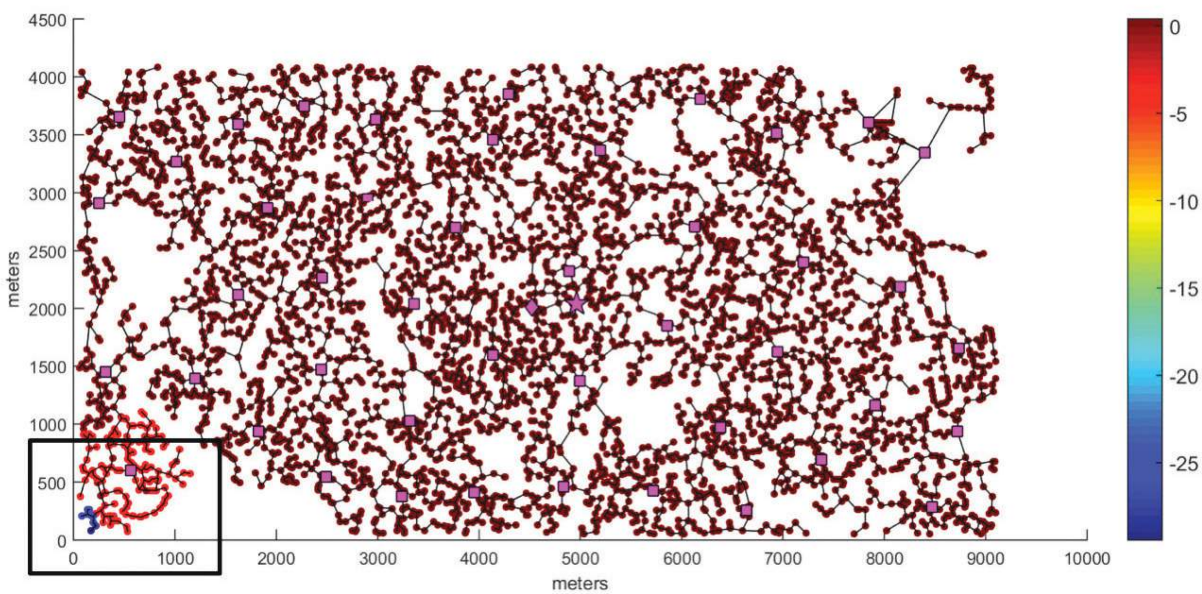
19 Voltages at all nodes are expressed "per unit," or normalized to a base voltage quantity, which is defined to be the voltage at the transmission/distribution substation (the pink star in **Figure 4.10**) in this system.

Figure 4.13: Spatial Variation in Distribution-Level Reactive Power LMPs in the Network of Figure 4.10, with 40 Percent PV Penetration (Measured as Installed Capacity as a Percent of Peak Load)

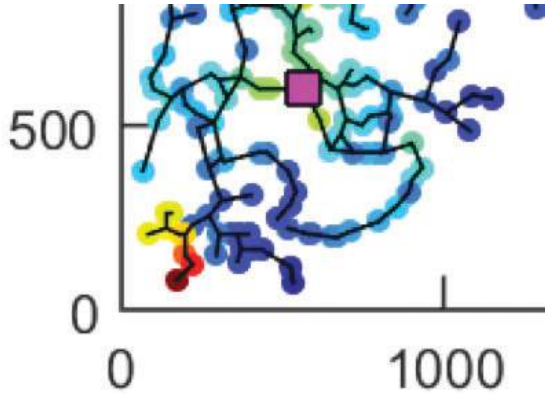
4.13a: Spatial variation of node voltage (per unit)



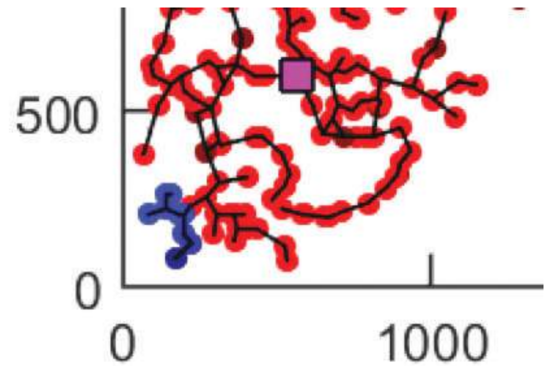
4.13b: Spatial variation of distribution level reactive power LMPs (dollars per Mvarh)



4.13c: Node voltage (per unit) in a zoomed-in section of the above network



4.13d: Reactive power nodal price (dollars per Mvarh) in a zoomed-in section of the above network.



4.13c and 4.13d illustrate the concurrence of the highest per-unit nodal voltage and the most negative reactive power nodal price.

since it is at those nodes that reactive power absorption can mitigate the voltage rise caused by greater PV output. This information can be used in several ways, such as: (1) the reactive power nodal price can be sent to network users with the expectation that they will dynamically react to it and alleviate the voltage constraint; (2) the sizeable magnitude of reactive power LMPs can be used to flag an area of the distribution network for the installation of a capacitor or inductor bank by the distribution utility; (3) reactive power nodal prices may help identify candidate nodes at which to run an auction to solve a voltage or reactive power compensation problem through a long-term contract; or (4) reactive power LMPs may signal the need for an adjustment of static settings of inverters in the relevant area of the distribution network (if the grid-code allows the utility or distributed resource to do so).

In summary, reactive power LMPs may be useful for flagging opportunities to install compensation devices, change equipment settings, or run long-term auctions for compensation services. The utilization of reactive power prices as market signals applied to nodal reactive power injection and consumption in day-ahead and real-time local markets in which market agents bid prices and quantities, may lead to volatile results and small gains from implementation. Further research in a representative set of networks is needed to evaluate this assertion.

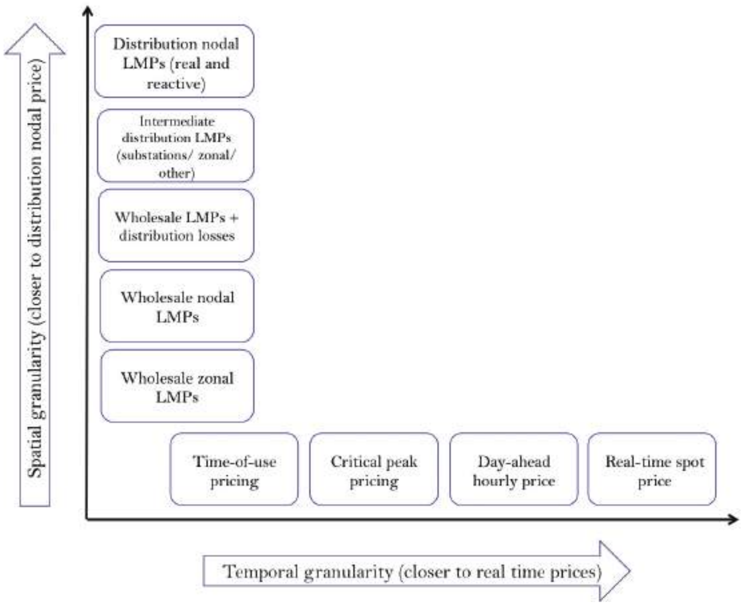
4.4.1.5 The choice of granularity level

While the cost of providing energy to users throughout the power system can vary significantly by time period and location, as shown in **Section 4.3**, most existing rate structures seen by end-users reflect few, if any, of these temporal and spatial variations in cost, and they are exclusively based on active power (i.e. excluding reactive power).[20] The question that remains is: how much additional benefit can be gained by more granular prices and charges? Obviously, the answer will depend upon the trade-off between the costs and the benefits of greater granularity.

Figure 4.14 illustrates several possibilities for enhancing the temporal and spatial granularity of the energy price, where an implementation choice would consist of some combination of an option along the time dimension and another along the spatial dimension, or variations thereof. Options along the time dimension reflect different degrees of price or rate variation, or differences in the degree to which prices or rates reflect real-time operating conditions. These range from rates that do not change with time over the course of a day to real-time prices

20 In the European Union, customers with power factors (the ratio of active power to apparent power) below a prescribed limit are penalized, restricting the flexibility of customers to absorb or inject reactive power.

Figure 4.14: Choices in the Adoption of Temporal and Spatial Granularity in Energy Pricing



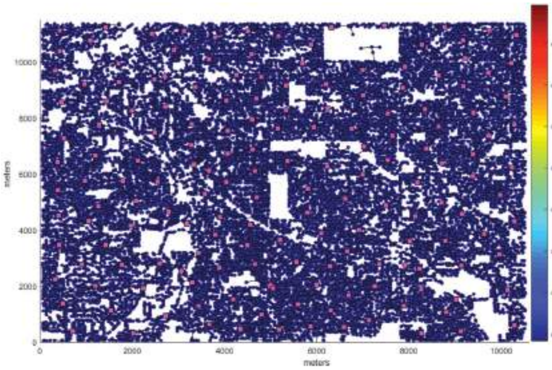
(for some time interval such as 15 minutes or one hour) for end-users. Starting from a flat, uniform tariff for an entire year or more, the first improvement (see **Section 4.3**) is a TOU rate that assigns rates to blocks of time based on expected energy prices. Over the course of a day, a user may see roughly three or four rates corresponding to different times-of-use. A step toward greater granularity is the use of CPP, which enables the utility to alert customers of "critical" events, such as an hour of particularly high load, during which the utility can charge a significant CPP adder. The next option is the use of day-ahead hourly wholesale prices (with or without locational differentiation) or, finally, the use of closer-to-real-time spot prices (which may be hourly or sub-hourly), as discussed in **Chapter 7**.[21] Along the spatial dimension, starting again with a tariff with no locational component, the first step is to include zonal wholesale LMPs, moving next to nodal wholesale LMPs, then into distribution-level granularity by adding loss factors to the transmission LMPs to account for distribution losses. Finally, distribution LMPs can be computed, perhaps for varying aggregations of distribution nodes (such as only throughout the HV network, at each HV/MV substation, or at each MV/LV transformer), or, as nodal prices for every distribution node. **Figure 4.14** also indicates that making use of the reactive component of the energy price might be seriously considered when using LMPs more deeply throughout the distribution network, where voltage-related problems may become more frequent, harder to solve, and can be addressed via reactive power pricing. Evaluating whether or not prices for reactive power need to be extended to distribution nodes requires consideration of their role and significance in both signaling efficient operation and investment, and (to a lesser extent) contributing to network cost recovery.

Figure 4.15 illustrates the spread of energy prices during a single hour (active power prices) at varying levels of spatial granularity for an urban distribution network that is not experiencing congestion (yielding nodal price variations that result solely from losses). The network is comprised of a transmission substation (pink star), HV/MV substations (pink diamonds indicated by arrows in

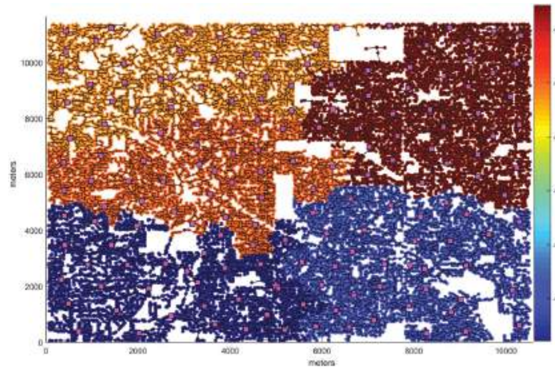
21 Note that temporal granularity is understood here as updating closer to real time, regardless of whether the length of the considered time interval is hourly, half hourly, every five minutes, etc. (although shorter time intervals are typically used closer to real time). However, this is not always the case. For example, intra-day and balancing markets in many European markets have the same granularity (hourly). Also note that using a market as a reference does not necessarily imply using the same time intervals of that market when doing settlements. For instance, the real-time market price in the United States has been used to make hourly, daily, or monthly settlements.

Figure 4.15: Capturing the Cost of Losses with Varying Levels of Energy Price Granularity

4.15a: Wholesale LMPs

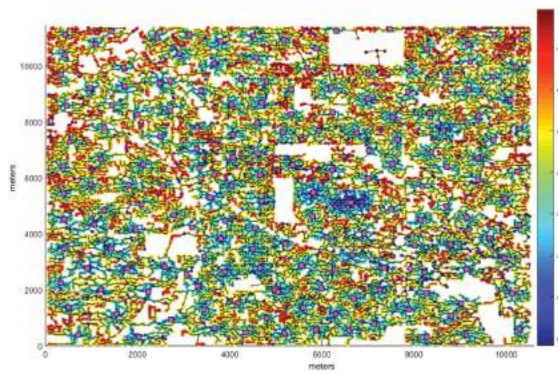


4.15b: Zonal LMPs for distribution users served by each HV/MV substation (substations indicated by arrows)



4.15c: LMPs with voltage-level distribution marginal loss approximations

4.15d: Distribution-level active power LMPs



4.15b), and MV/LV transformers (pink squares), along with HV (132 kV), MV (20 kV), and LV (400 V) lines.

4.15a illustrates the case in which the same wholesale LMP is conveyed to all distribution system nodes. **4.15b** shows zonal LMPs computed at each HV/MV substation (indicated by the black arrows). The marginal cost of HV distribution network losses (which are quite low in this example) is added to the wholesale LMP to produce a marginal price at each HV/MV substation that accounts for HV distribution losses. **4.15c** adds to the wholesale LMP a voltage-level-specific loss factor that estimates the marginal losses at each voltage level. That is, an average marginal loss factor is added to the LMP for all LV users to account for their impacts on marginal losses throughout the network, and, similarly, an average marginal loss factor is added to the LMP for all MV users to account for their impacts on marginal losses throughout the network.

Since most users (93 percent) in the modeled network are LV users, most nodes exhibit the LV marginal price, shown in green. The remaining 7 percent of users (MV users) face a slightly lower MV marginal price, shown in blue. Finally, **4.15d** shows the active power LMPs at all MV and LV distribution nodes.[22] The spread of prices increases with each successive step toward full nodal pricing, capturing and revealing more to network users about their impacts on marginal losses.

The benefits of greater granularity will be discussed in **Section 4.5**, once we review how prices and charges for the most relevant electricity services are computed. The primary costs of improved granularity are computation and implementation costs (see **Box 4.3**). Implementing the computation and use of distribution LMPs, if deemed

22 Reactive power marginal prices can be computed for all distribution nodes as well, but they are not shown in the illustration here.

convenient, may require significant time since ICTs must be deployed, and robust algorithms must be verified in sufficiently representative networks. Implementation requirements can be much lower for simplified, or less-granular, approaches. For example, the pass-through either of wholesale prices or of LMPs computed at the transmission-distribution interface node to network users downstream (preferably with the addition of some loss factors approximated for large distribution zones) would have low implementation costs beyond the deployment of advanced meters, while yielding sizeable efficiency gains relative to the status quo.

4.4.2 The price of operating reserves

The cost of operating reserves is typically a small share of the total cost of energy, adding roughly 1.5 percent to 2 percent to the total production cost of energy (Hummon et al. 2013). The exact share of wholesale electricity costs arising from reserve requirements depends upon the reserve margin specified for a particular system, as well as the mix of resources providing energy and reserves. For example, in 2016, the wholesale cost of electricity (all expressed in terms of dollars per megawatt-hour) in PJM consisted primarily of energy costs (62 percent of the wholesale cost), capacity costs (21 percent), and transmission costs (16 percent), while reserves (including regulation, operating reserves, synchronized reserves, and black-start capabilities) accounted for approximately 0.9 percent of the total wholesale cost (Warner-Freeman 2016). These values can change over time due to changes in fuel costs, market rules, etc. In 2008, the breakdown of the PJM wholesale cost of electricity was: 84percent energy, 10percent capacity, 4percent transmission, and 2percent reserves. In 2014, the ISO New England ancillary services market (which includes operating reserves) was roughly 2.6 percent of the size of the energy market, and, in 2015, the ancillary services market was 3.7 percent of the size of the energy market (ISO New England 2015).

Despite the comparatively small size of the reserves market, it might represent a significant economic opportunity for some price-responsive demand or DERs. Moreover, the participation of price-responsive demand and DERs might be critical during periods of system stress due to lack of operating reserves. Therefore, it should be an objective in the design of wholesale electricity markets to: (a) make it possible for DERs to participate in the provision of operating reserves on equitable terms with centralized facilities (by, for instance, being responsible for their deviations from a committed schedule); and (b) facilitate the management of the variability and uncertainty of intermittent resources such that they are not unduly penalized, and power system requirements for operating

Box 4.3: Computing Bulk and Distribution LMPs

Since differences in node voltage magnitudes and resistance-to-reactance ratios are of greater significance in MV and LV distribution than in transmission and subtransmission, a full AC optimal power flow (OPF) or a suitable convex relaxation that captures relevant physical features must be utilized to compute LMPs within the MV and LV distribution networks. To compute LMPs in the transmission and HV distribution networks, aggregating load and generation along MV and LV feeders into net power flow to or from the point of connection of each HV/MV transformer provides sufficient representation of the MV and LV distribution systems. A DC OPF can be used to compute LMPs at nodes throughout the meshed transmission and subtransmission/HV distribution networks. Carrying out these two stages iteratively until the convergence of meshed LMPs and radial LMPs enables the computation of nodal prices at each HV/MV substation bus and each MV/LV transformer bus that encapsulates the full set of relevant bulk and distribution system operating conditions. As will be described in **Chapter 5**, executing the computation of nodal prices for market operation raises important considerations for the definition of the roles of industry actors and for the delineation of regulatory jurisdictional boundaries.

reserves are kept as low as possible. These topics are discussed in detail in **Chapter 7**.

From the viewpoint of DERs, a requirement for the fully efficient pricing of operating reserves is that economic signals conveying the costs of reserve provision and the occurrence of reserves scarcity may reach all power system agents — in particular, those connected to the distribution system. The provision of reserves usually takes place in auctions at time frames ranging from annual to a day ahead, either in concert with the energy auction (as in the United States in the day-ahead time frame) or separately (as in the European Union). DERs should be able to participate in such auctions, as they already can in some markets, in response to actual system needs and as close as possible to real time. The relatively modest cost of reserves is typically averaged and included as an uplift or adder to the end-user tariff. However, in order to elicit the response of demand and DERs to situations of scarcity in operating reserves, it is critical that this scarcity be reflected in the energy price and in as close as possible to real time. The marginal impacts of reserve scarcity on energy price can be substantial and far exceed the importance of the direct payments for reserves. Again, see **Chapter 7** for a detailed discussion of these topics.

4.4.3 The price of firm capacity

The impact of cost-of-capacity mechanisms on the total cost of electricity supply may vary significantly depending on the characteristics of the power system in a particular location. Even within the same system, it may be possible to have different weights in different years, due to the volatility of other parameters, such as the overall energy cost. Furthermore, capacity mechanism costs may increase as a percentage of the overall supply cost in the near future, due to the penetration of renewable technologies. According to ISO New England (2015), "Renewable resources are expected to put downward pressure on energy market prices, but this action is not without consequences: a decrease in electricity prices will put upward pressure on prices in the capacity market"; thus the remuneration to generation resources is expected to slowly move from energy to capacity markets, at least in this specific system.

In Spain, capacity payments amounted to 735 million euros in 2015, or about 5.6 percent of the price of energy in the wholesale market over the entire year (CNMC 2015 and OMIE 2016). The United Kingdom recently introduced a capacity market, and the first auction was held in 2014 for delivery of capacity in 2019 (Ofgem 2015). If this capacity mechanism cost (942 million pounds) is compared to the current wholesale energy price, we once again see a gain of around 5.6 percent.[23] According to ISO New England's annual report, the ratio between the capacity and the energy market cost has been volatile in the interval 2007–2015, ranging from 11 percent to 24 percent. As mentioned in **Section 4.4.2**, capacity costs make up roughly 10 percent to 20 percent of the total wholesale cost of energy in the PJM market.

Chapter 7 discusses in detail the design of capacity markets or capacity remuneration mechanisms and the opportunity for DERs to cost-effectively participate in these markets. These markets can ensure appropriate compensation for DERs that commit to delivering firm capacity. Here we focus on how to send efficient economic signals that can coordinate the behavior of diverse electricity consumers and DERs that do not participate directly in such markets.

It is important that all agents, particularly DERs, be informed about their effect on firm capacity, so that they can respond efficiently in real time and with investments that can reduce the firm capacity required by the system. Since most agents are not yet exposed to real-time price and charge signals, and price caps commonly limit wholesale energy prices during periods of supply scarcity, the signals about firm capacity scarcity[24] are generally obscured.

Thus, if the energy prices applicable to end-users and DERs[25] are not reflective of marginal costs during periods of firm capacity scarcity (either due to wholesale price

23 Note that in capacity mechanisms that rely on penalty schemes (as in the United Kingdom), it is actually impossible to estimate beforehand the exact cost of the mechanism, since resources that are committed but not available may have to pay penalties, which reduces the cost paid by consumers to support the capacity mechanism.

24 As responsive customers become more automated, creating more short-run price elasticity, this choice can and should be reconsidered.

25 These economic signals may be received by users' smart energy boxes or by aggregators that manage users' appliances automatically. What matters is that economic signals with adequate spatial and temporal granularity are computed and available.

caps or simplified energy prices), it is efficient for some additional signal(s) to convey users' impact on generation capacity costs. For example, in the presence of a capacity market, a peak-coincident generation capacity charge can restore an efficient short-run incentive and determine the allocation of marginal capacity costs among users. This charge would reflect the coincidence of an end-user's consumption (or network withdrawals) with the aggregate peak in the zone of the power system for which firm-generation capacity requirements are defined (this is analogous to the manner in which users' peak coincidence contributes to network capacity requirements as described in **Section 4.4.4**). DERs that inject power into the grid during aggregate peak periods should be credited with a symmetrical payment or bill credit reflecting the reduction in firm capacity requirements.

In practice, peak-coincident capacity charges create significant and volatile short-run signals. Thus, if end-users are able to participate in a capacity market, either directly or via aggregators, agents with good forecasts of their likely availability during scarcity periods and preferences for more certain revenues can commit themselves to contribute to firm capacity requirements by participating in a capacity market *a priori*. By bidding firm capacity — a forward commitment to either inject capacity or reduce withdrawals during aggregate peak periods — users can be hedged against short-run scarcity signals for the amount of firm capacity bid into and cleared in the capacity market. Any consumption above the quantity cleared in the capacity market would remain exposed to the full short-run signal, including the peak-coincident generation capacity charge.

4.4.4 Network charges

In **Section 4.4.2** we have discussed locational marginal prices (LMPs), which determine the price for electrical energy at any node of the network and at any instant of time, incorporating all the short-run impacts of generation, demand, and network effects (losses and network congestion). LMPs precisely determine the price of the fundamental energy service. Our next task is to allocate and recover the regulated network costs efficiently. The discussion that follows focuses on the distribution network, although much of what will be presented also applies to transmission.

From a regulatory standpoint, distribution and transmission networks are customarily treated as regulated monopolies, with the corresponding regulatory authority establishing for each year the appropriate remuneration. There are two main priorities when designing network-related economic signals: economic sustainability and allocative efficiency. Economic sustainability means the totality of the funds collected via network charges must equal the regulated network costs, since this is the amount that the regulator has established must be recovered. Allocative efficiency means efficiency in the allocation of the network charges. Here the objective is to maximize the utility of the network users (i.e., global welfare), eliciting their efficient response and the associated impact on the power system while minimizing economic distortions.

The ideal solution would be to achieve both objectives with a single shot: By applying perfectly efficient LMP signals to network users, the network would acquire power at the injection nodes paying the appropriate LMPs for the energy, then transport the power and sell it (minus any incurred losses) at the withdrawal nodes, receiving for the energy delivered the LMP of each node. As LMPs are generally higher at consuming nodes than at generating nodes, this system generates "network rents" or revenues. In an ideal world, LMPs would thus both send economically efficient signals to all network users and generate sufficient revenues to recover regulated network costs.

In the real world, however, the revenues recovered by this method over the life of the network are only a small fraction of the cost of well-adapted network investments (see **Boxes 4.4** and **4.5**). In the absence of economies of scale in network investment, of noneconomically justifiable engineering design requirements, and of planning errors leading to excessive and irreversible investments, it has been proved that LMPs would completely recover the totality of efficiently incurred network costs (Rubio-Odériz 1999). However, none of these assumptions holds true in reality. Moreover, the lumpiness of transmission investments and the risk aversion to power system failures substantially magnifies the impact of the other factors. In practice then, network rents generated by LMPs account for only a small percentage of total network cost recovery (Rubio-Odériz 2000). Furthermore, to date nodal LMPs have not been implemented anywhere at the distribution level and have only been adopted at the transmission level in certain jurisdictions. In this context, the potential contribution of network rents from LMPs to cost recovery is purely hypothetical.

We are therefore back at square one in the search for an efficient network cost allocation method, but we have learned something: LMPs, the perfect short-term marginal energy prices that incorporate all network effects, are not adequate long-term signals and do not secure full recovery of regulated network costs. With only LMPs, the network users lack information about the implications of their network utilization patterns on the needs for future network investments.

Before embarking on the search for the most efficient cost allocation approach, it is important to review the requirements with which such an approach must comply in any realistic situation.

1. Since LMPs are perfect short-term energy prices, any additional economic signal that is associated with the actual pattern of network utilization should be well-justified. Any charge not justified by efficiency objectives (e.g., designed to incentivize efficient short- and long-run behavior) should be applied in the least distortive manner, i.e., avoiding interference with LMPs (or their implementation by proxy) and with any other well-justified economic signal.
2. The correct additional economic signal to be sent to network users should be based on each user's specifically foreseeable contribution to (or responsibility for) any estimated future investments in network assets. This cost-reflective network charge is consistent with the widely accepted principle of allocation of costs to the beneficiaries of the investments. The fundamental reason a network investment is made is to increase the benefit users derive from the network by a larger amount than the cost of the investment (see Meeus 2013).
3. The "residual network costs," i.e., the fraction of network costs that is not allocated on the basis of the cost-reflective method noted above,[26] should be allocated in the least distortive possible manner.
4. Although this may not be considered a priority at the moment, the allocation of the costs of the distribution and transmission networks must gradually take into account the increasingly blurred boundary between these networks, so that ultimately all network users participate in all network costs according to the same principles.
5. Ramsey pricing[27] advocates the use of the inverse elasticity rule to optimize the efficiency (i.e., maximize welfare) of the allocation of the costs of monopolies, or of some prescribed amount of taxes, under some conditions. It happens today that the availability and decreasing cost of distributed generation and storage result in the possibility of grid defection (i.e., radical disconnect from the grid). (For the substantial amount of costs that can be avoided by defection, see **Figure 4.2.**) Grid defection is an extreme form of elasticity to price that has to be taken into account from an efficiency perspective when deciding the regulated costs to be included in the electricity tariff and in the design of the tariff itself (see **Section 4.4.5** for a more detailed discussion).
6. It is a common regulatory practice all over the world that all consumers supplied by the same distribution company, connected at the same voltage level, and under the same type of contract (e.g., flat rate, critical peak pricing, etc.), are subject to exactly the same tariff — regardless of where they are connected. This implies substantial subsidization of customers in high-cost portions of the network (such as rural networks) by those in lower-cost portions of the network (such as dense urban areas). This cross-subsidization is generally accepted without complaint, either because people ignore that this socialization or averaging of

26 Or whatever fraction of the total cost could be recovered via congestion and lost "rents" associated with distribution LMPs if implemented.

27 See **Section 4.4.5.**

the distribution network costs is taking place or because such cross-subsidization is desired to address distributional concerns.[28] If cross-subsidization is desirable, the advanced methods of tariff design that are proposed in this study should make compatible the twin goals of sending efficient economic signals and maintaining the "equal-tariff-for-all" principle, at least approximately. We discuss solutions to these challenges in **Section 4.6**.

7. Even if the average value of the tariff, or the cost of electricity for the household, remains more or less the same once the advanced tariff designs that this study proposes are implemented, the volatility over time of network charges, and of economic signals in general, could be a matter of concern, as many customers may feel that they are unqualified to respond and take advantage of the opportunities that this information may bring them. We also discuss this challenge in Section 4.6.
8. DERs can provide network services, deferring network-related investments that would otherwise be necessary. In order for DERs to have an economic incentive to provide these services, they need the correct economic signals with sufficient granularity in space and time.
9. Finally, it has to be remembered that it is not only the quantity to be paid, but also the format (lump annual sum, $/kW, $/kWh charge, or other format) that matters in tariff design. We have already explained in this chapter (**Box 4.1**) how charging network or policy costs with a volumetric tariff leads to multiple economic distortions that are exacerbated by net metering. Format is critical in adopting a charge that sends an efficient signal without distortion.

A perfect design of network tariffs is one of the few stubbornly hard-to-solve problems in electricity regulation, and this study does not attempt to terminate the debate. The objective of the rest of this section is to provide useful insights and options for improved network tariff design and to warn against common mistakes. While a perfect outcome that simultaneously satisfies all objectives for tariff design may not be forthcoming, progressive improvements can be made to unlock more efficient outcomes and preserve network cost recovery.

28 Socialization usually has limits. For instance, rural customers usually have lower quality service standards. Therefore, although they are subsidized, they do not receive exactly the same service as urban customers.

4.4.4.1 How to make use of distribution location marginal prices if available

We have already said that so far no power system uses LMPs at the distribution level and that, even if LMPs were computed and used as energy prices, the potential contribution of LMPs to network cost recovery is likely to be modest. Why then worry about distribution LMPs in this section on network charges?

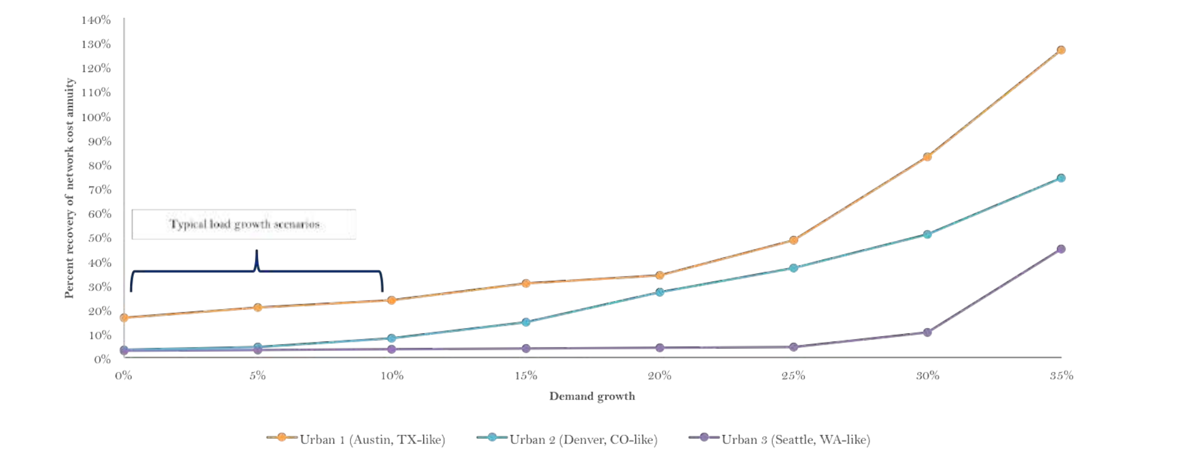
First, distribution-level LMPs, or different approximations of them, may start being used in some systems, since they are the reference or gold standard for efficient short-run signals. If this is the case, particularly in tandem with price-responsive DERs and demand, network capacity may come closer to mirroring actual utilization patterns and more significant differences may emerge between nodal prices in distribution networks, increasing the "network rents" obtained via LMPs. Although such revenues would likely continue to be insufficient to recover most of the network costs, the collected amounts could be substantial and their utilization should be supervised. Depending on the specific separation of activities in each jurisdiction, these funds would be collected by the distribution company, the incumbent retailer, or even independent retailers or aggregators. These funds could then be used to partly offset the network costs that will have to be recovered with network tariffs.

Box 4.4: Distribution Network Cost Recovery from Nodal Prices

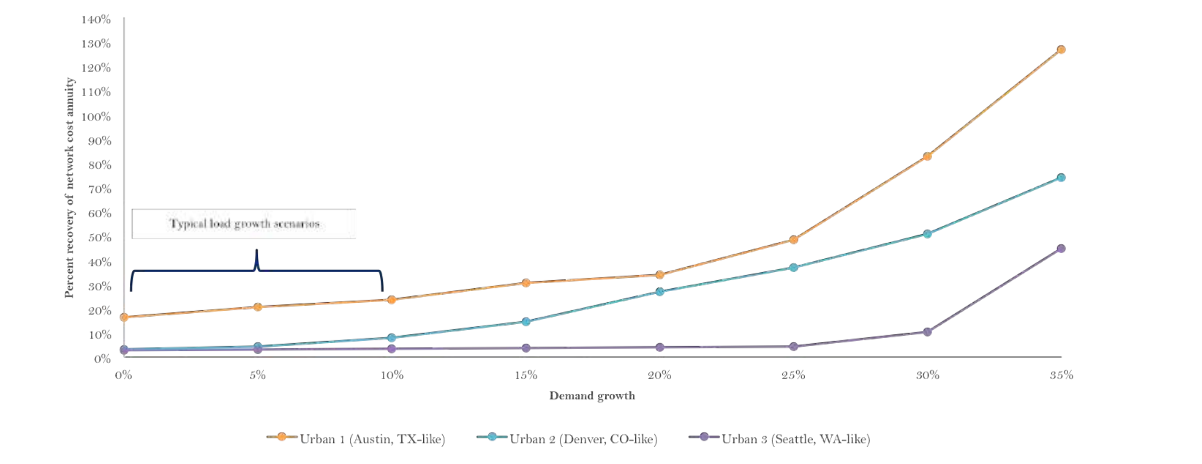
Distribution-level LMPs in a "well-adapted" network — that is, a network built to meet present and expected demand growth during some reasonable time interval — would typically be insufficient to ensure network cost recovery (see the main text for causes of this shortfall). **Figure 4.16** illustrates annual distribution network cost recovery via active and reactive power distribution LMPs in simulated urban (**4.16a**) and rural (**4.16b**) distribution networks with load growth ranging from 5 percent to 40 percent in total user demand.[29] For reference, over the years 2000–2015, electricity demand growth in the United States has averaged about 0.5 percent annually (EIA 2016). The results illustrate that under typical conditions of network design, operation, and growth in network-use, revenue from distribution nodal prices is likely to recover less than 20 percent of the annual cost of the distribution network.

Figure 4.16: Recovery of Distribution Network Costs via Distribution-Level LMPs

4.16a: Network cost recovery — urban



4.16b: Network cost recovery — rural

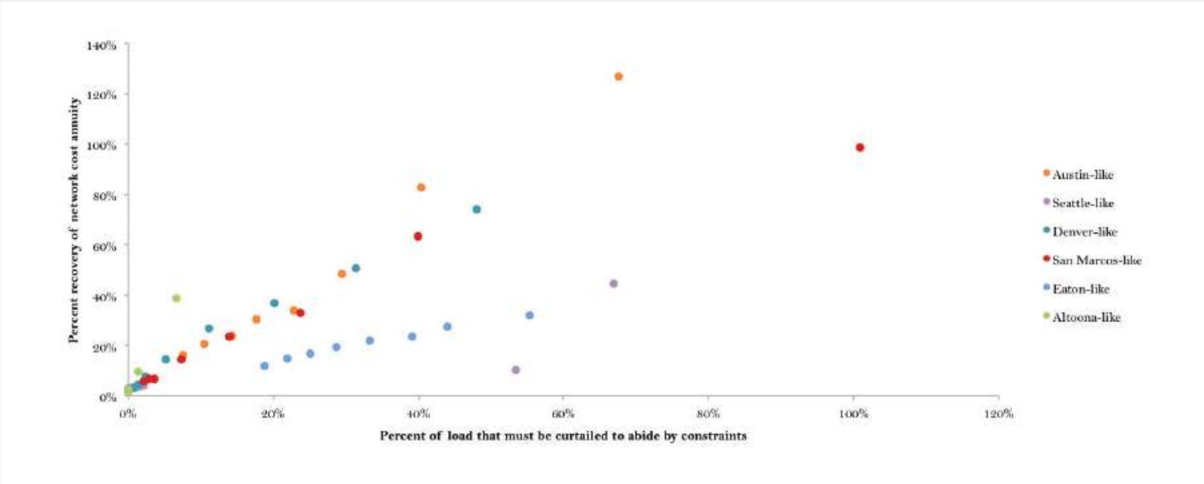


29 Demand growth was simulated in three urban networks and three rural networks in which no network expansion occurred in response to demand growth.

While the above results demonstrate that it would be unlikely for LMPs to provide full distribution network cost recovery under typical network design and operating conditions, multiple factors contribute to the exact magnitude of cost recovery from nodal prices. The prevalence of congestion at various locations throughout the network and the frequency of congestion at any given location over the period of cost recovery are instrumental in determining the amount of recovery. In an "overadapted" network built with significantly greater capacity than peak network use, congestion will occur less frequently; thus, nodal prices will less frequently rise to the cost of nonserved energy. As a result, nodal prices will yield less network cost recovery than in a more tightly built (or "well-adapted") network that experiences more frequent congestion.

Much of the variation in cost recovery trends seen among the networks depicted in **Figures 4.16a** and **4.16b** can be explained by such differences in network adaptedness. For example, the network called Rural 1, which is somewhat overadapted to base demand and experiences relatively infrequent congestion over the full range of examined load-growth scenarios, exhibits steady cost recovery throughout the range of load growth. In contrast, Rural 2 is overadapted until load grows by 30 percent over base demand, at which point congestion increases significantly, and a sharp rise in cost recovery is apparent. **Figure 4.16c** illustrates the relationship between the magnitude and frequency of congestion (as measured by the amount of load curtailment required to abide by network constraints) and the amount of recovery of the network cost annuity for the networks shown in Figures 4.16a and 4.16b.

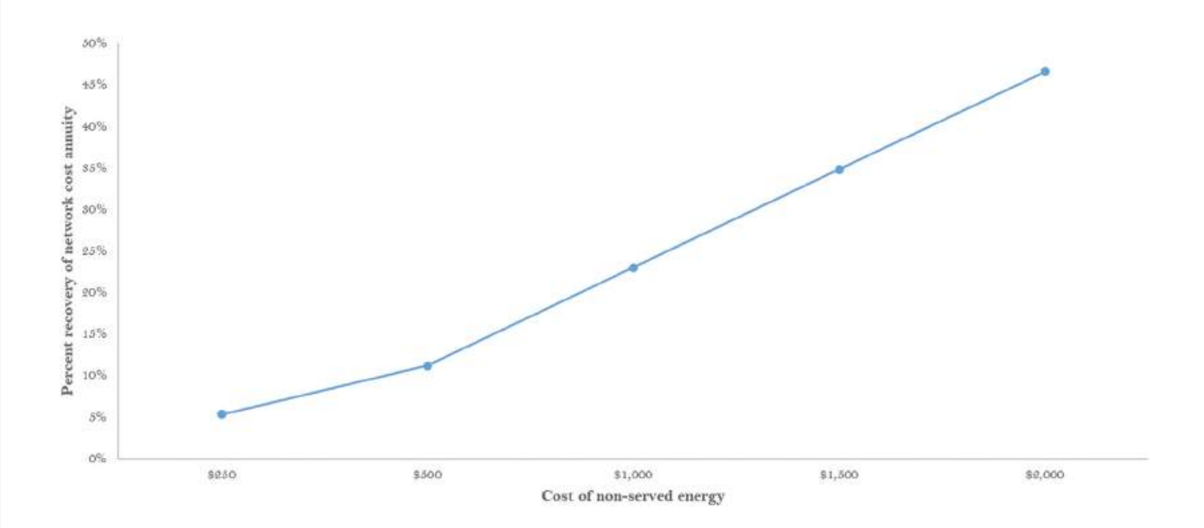
4.16c: Network cost annuity recovery



This chart shows recovery of the network cost annuity from distribution LMPs as a function of load curtailment for each of the six networks in **Figures 4.16a** and **4.16b**, assuming cost of nonserved energy is $300 per megawatt-hour.

Of course, in the presence of congestion, the assumed cost of nonserved energy (CNSE), or the marginal cost of some demand-side curtailment or distributed generation, strongly influences the resulting network cost recovery. For example, **Figure 4.16d** illustrates the influence of the CNSE on the percent recovery of the network cost annuity over a range of CNSEs (from $250/MWh to $2,000/MWh) that yield a consistent dispatch outcome (i.e., constant curtailment level) with a fixed level of load (the base demand level with no load growth) in the Austin-like urban network, Urban 1.

4.16d: Impact of cost of nonserved energy



This chart reflects the impact of the cost of nonserved energy on the percent recovery of the network cost annuity with distribution LMPs (base demand, no load growth, in urban network similar to that of Austin, TX).

4.4.4.2 Assigning responsibility for new network investments

We want to relate the patterns of network user behavior and the network costs[30], and convey the corresponding economic signal to the network users. This signal, if it reflects actual system conditions, will interfere with the LMP that ideally would be employed. Does this make economic sense, given that the LMPs, either at the transmission or distribution levels, are the efficient short-term signals? We maintain that adding an economic signal on top of LMPs (if employed) — for the agents connected to a node where the withdrawal or injection of power at a given moment in time indicate the need for new investment — makes economic sense. The action of the agents has two separate economic consequences: First, to modify the cost of operation (this is captured by the LMP) and, second, to indicate the need for new network investments (which the LMP alone, due to economies of scale in network investment, enhanced by engineering planning rules, discreteness and irreversibility of network investments, fails to capture, in practice by a large amount, as the collected experience until now seems to indicate).

The literature on network cost allocation at the transmission level is enormously abundant (in contrast with that for distribution), since the design of distribution network charges was considered straightforward under the top-down paradigm. We can therefore borrow ideas from transmission to be used at the distribution level. The criteria for cost allocation should be to quantify the impact that the actual behavior of network users has on creating the need for new network investments (by connecting to the network, installing new electricity-using appliances or DERs, or augmenting the ones that they have). **Box 4.5** summarizes several approaches that are inspired by this principle.

4.4.4.3 The format of the peak-coincident network capacity charge

A peak-coincident network charge, however it is calculated, can be an efficient method for guiding long-term utilization of electricity networks and recovering a portion of regulated network costs. A key question, however, is how the corresponding economic signals should be conveyed to the network users. The first challenge is to define what can be understood as a "peak" during which peak-coincident network charges are applied. Do charges apply only in the one single moment of greatest demand? Should the peak periods be specified in advance, or should the charges be applied *ex post* to the actual peak periods, whenever they occur? Should we define the peak at the level of the entire power system, the HV/MV substation, or the MV/LV substation? As with efficient charges for energy, efforts to define efficient signals for network utilization must determine the appropriate degree of both temporal and spatial granularity.

To answer these questions satisfactorily, the proposed approach must be able to send signals that reflect any actual stressful conditions in the system, capturing patterns of contribution to the peaks; therefore it must use an ensemble of the highest peaks and not just the highest single peak. The results obtained could be averaged over a certain area (to be specified for each case) to avoid having very different charges for customers whose locations are very close. What follows is a discussion of a possible implementation approach that meets all of these requirements.

It can be safely assumed that distribution networks are planned and designed to accommodate peak capacity requirements under a range of plausible operating scenarios and anticipated system peaks (accounting for variations in season, weather patterns, load growth, and other factors). Making use of a sufficient number of distinct system peaks (in withdrawals or demand and in injection or generation) throughout the year to allocate the capacity charges ensures that variability in network use patterns is accounted for and minimizes randomness in the computation of network charges. Thus, the distribution utility and/or regulator can allocate the annuitized capacity-related network costs (for peak demand and peak generation) to be collected over multiple periods (months may be reasonable, or portions of months) within a year, based on the probability and magnitude of anticipated peak conditions. For instance, assume that the 30[31] highest peaks in demand of the

30 In particular, the reinforcements that will be needed in the future or have been recently made as a consequence of the behavior of each specific network user.

31 Thirty, or a similar number, reduces the randomness in the application of the charges and captures the most significant peak conditions in the power system.

Box 4.5: Approaches to Determining User Responsibility for New Network Investments

The most direct approach to determining responsibility for network investments would be to try to estimate the marginal (or incremental) impact that the actions of the agents have on the cost of future network reinforcements or replacements. Distribution network capacity expansion models could be used to identify reinforcements, according to some efficient planning criterion employed by the distribution company (distribution companies typically plan the distribution network to meet the estimated future peak conditions of demand and, more recently, also the peaks of production of distributed generation). The cost of these reinforcements could be allocated according to the users' coincidental contribution to the peak power flows in the feeder (both injections and withdrawals). The discreteness of network investments creates difficulties in relating the gradual growth of demand and generation with the discontinuities of the investments. This relationship can be established quantitatively in some approximate way for each situation.

A second approach, from Abdelmotteleb, Gómez, Chaves, and Reneses (2016), is based on the marginal participation method, which consists of first selecting those critical network components with a scarce margin of installed capacity over utilization. Then, the network users that contribute to the flows in the critical network components are identified so they can be charged for future reinforcements.

A third approach, from Pérez-Arriaga and Bharatkumar (2014), starts by determining the weight in the distribution network cost (total or by voltage level, area, or feeder) of the two major cost drivers for incremental network investment: meeting the expected peak demand and meeting the expected peak reverse flow due to distributed generation. Assuming that the costs of future network investments and the weights attributable to peak demand and peak distributed generation remain consistent with estimates derived from a planning model for the existing network, it is possible to evaluate per-unit charges for the contribution of coincidental peak demand and coincidental peak generation.

A fourth approach, from quite a different perspective, is presented in detail in **Chapter 6.3.3**, whereby regular auctions of options on network capacity would be held to elicit information on network users' willingness to pay for future network capacity investments as well as DER alternatives. The results of this auction would then determine network users' responsibility for network cost recovery.

In all these approaches, the strength of the economic signals sent to network users through their network charges should be modulated to account for the capacity margin of the network under consideration (e.g., at the aggregate level or, by voltage level, zone, or feeder) over the existing maximum power flows. The objective is clear: Do not incentivize unnecessary changes to network utilization because that distorts the economic signal and reduces efficiency. For instance, if a distribution feeder has ample capacity to accommodate existing peak network-use as well as significant growth in network utilization, the economic signals conveying users' contributions to projected reinforcements should be weak or nonexistent. Under such circumstances, it would be inefficient to encourage an even lower utilization of the network assets. In contrast, if a network is being utilized at or near its full operational capacity, or is projected to be so soon, then the economic signals conveying the marginal or incremental cost of projected reinforcements should be stronger. Determining the most effective format for the relationship between the strength of the charge and the degree of network utilization is a key aspect of the design, and it will also depend on the response of distributed agents in a diversity of contexts.

previous year occurred during five hours in January, nine in February, three in March, six in July, five in August, and two in December. This effectively allocates the peak demand-driven costs to each month according roughly to the probability of a system peak occurring within that period. Dividing the total recoverable peak demand-related costs across months according to anticipated system peaks provides network users with economic incentives to guide their network utilization behaviors and shift network use away from potential peak periods. Within each month to which some fraction of total peak demand-related costs have been allocated, each network user's share of the total capacity costs can be computed according to that user's *ex post* share of the actual peak or collection of actual highest peaks each month — i.e., according to the ratio of the user's kilowatts of consumption to the system kilowatts of consumption. An entirely parallel procedure must be followed for peak generation-related costs.

Note that the total sum that must be collected from users for network capacity investments in a given month is established *a priori* and does not depend on the actual behavior of the network users during that period. However, the allocation of the monthly charge among network users does depend on each user's contribution to the actual peaks that month, creating efficient incentives for network utilization in the short run. The magnitude of total recoverable capacity costs allocated to each month is based on the anticipated peaks used in distribution network planning because the anticipated peaks (rather than actual peaks) guide network investments. With a record of the historical system peaks plus regulatory validation for the projected peaks used in the course of determining distribution utility remuneration, it is likely that the anticipated system peaks will match the actual system peaks reasonably well, producing an accurate allocation of total capacity-related network costs to several forecasted peak periods. Operations and maintenance costs that vary according to capacity requirements (i.e., on a per-kilowatt basis) can be added to each user's share of capacity-related investment costs on a per-kilowatt basis.[32]

These "peak-coincident network capacity charges" can be either positive or negative, depending upon whether network users consume or inject power during periods of peak demand or production. For example, injection from distributed generation (DG) during periods of peak system consumption may help reduce the aggregate network withdrawal peak, reducing the need for network investments. Injections that occur during the aggregate withdrawal peaks can thus receive a negative network capacity charge — in the form of a payment or credit on a user's bill — because of the value of DG in serving load and reducing power flow and system congestion. Likewise, power injection during periods of system peak injection can lead to a positive capacity charge, while withdrawals during the peak injection period should be compensated symmetrically. This symmetry in peak coincident network charges/payments mirrors the symmetry in LMPs and is critical to ensure all resources are placed on a level playing field and compensated or charged in a nondiscriminatory and nondistortionary manner.

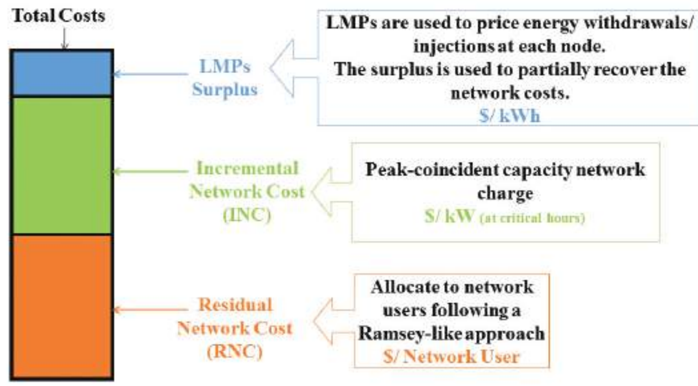
4.4.4.4 What to do with the residual network costs?

It is possible that only a modest fraction of the total revenue requirement for a distribution network can be recovered with the "peak-coincident network capacity charges" and via network rents from LMPs, if they are employed. Therefore, we shall define "residual network cost" as the fraction of the network cost that still remains unrecovered after the application of LMPs (if used) and the peak-coincident network capacity charges. These three components of the network cost have been represented in **Figure 4.17**.

Once the efficient short-term signals (LMPs, if used) and long-term signals (peak-coincident network capacity charges) have been applied, there is no need to send additional economic signals to influence the behavior or investments of network users. Indeed, any other charges based on the relationship between the pattern of user injections and withdrawals would result in economic distortions that would increase the cost of the power system. Other economic reasoning is necessary to find the criterion for allocating these residual network costs in a manner that minimizes economic distortions. This criterion will be common for the residual network costs and the policy costs, and it will be presented in the next section.

32 We illustrate the application of this format for network charges in **Boxes 4.6** and **4.10.**

Figure 4.17: Charges that Recover Regulated Network Costs



4.4.4.5 The end of the "top-down paradigm"

Our proposed approach to the design of network charges follows the classic distribution/transmission separation: (1) the agents connected at the transmission level (this includes the agents connected in the distribution network, conveniently aggregated at the transmission/HV distribution transformation nodes) cover the transmission network costs; and (2) the agents connected to the distribution network pay the distribution network costs. This approach makes perfect sense for current power systems in which the distribution network has been designed to bring power to end consumers, who are the obvious major beneficiaries of the distribution network. If a feeder does not have enough capacity, the consumers connected to that feeder have to be disconnected, resulting in a heavy loss of individual utility (therefore these consumers would benefit much from a grid reinforcement), whereas the large generators, connected to the transmission network, only reduce their sales a small amount.

This situation may change in the future if distribution networks are populated by a multiplicity of DERs, including abundant distributed generation and storage. Under these conditions, the response to a shortage of capacity at the distribution level will be to activate the local resources, and the loss of utility for the agents connected in the congested distribution area will be much smaller. The mindset about who should pay for distribution network costs will thus have to be changed accordingly, paving the way for methods more like what is now used for the transmission network.

Under these hypothetical conditions, a separation of the networks into "meshed" and "radial" would make more sense than the present division between "transmission" and "distribution," which is mostly artificial and follows different criteria in the diverse power systems.[33] Any sound method for determining transmission charges would be applied to the meshed part of the network, and the resulting charges for nodes connecting the meshed and radial networks would be passed to the agents in the radial feeders.[34] Similarly, the cost of each radial feeder would be allocated to the agents connected to it, including the agent at the head of the feeder.[35]

33 Note that in most distribution networks, the high-voltage part has a meshed structure, as in transmission. The medium-voltage network is also meshed, but it is operated in a radial configuration, with some breakers open, but with the possibility of reconfiguration, if needed. The low-voltage network is radial, typically with a tree-like structure.

34 These charges should be based on who is responsible for network investments and would adopt the format of peak-coincident capacity charges. In other words, the radial feeder would be seen, from the meshed network perspective, as a generator (power injection) or a demand (power withdrawal) at the corresponding meshed/radial connection node.

35 The peak-coincident capacity charges that are transferred from the radial to the meshed part of the network at each node will have to be allocated to the generators and demands connected to this meshed network following some criterion of responsibility in the injections or withdrawals at the meshed/radial connection nodes.

4.4.5 Charges to recover policy costs and residual regulated costs

We have already defined network and policy costs and their rising importance in some jurisdictions (**Section 4.2**). Two additional major issues are now addressed, as how to allocate residual regulated costs (the costs that cannot efficiently be recovered via forward-looking cost-reflective charges). First, it is necessary to discuss the merits and risks of including policy costs and residual network costs in electricity tariffs. Second, for those residual (network or policy) costs that the regulator decides to include in the tariff, it is necessary to find a sound criterion for allocating these costs to network users.

It is not a central objective of this study to recommend what policy costs should be included or removed from electricity tariffs in different jurisdictions, although we briefly discuss certain criteria that could be taken into consideration. Most of these costs are included in tariffs for political reasons, and it is the responsibility of policy-makers and regulatory authorities to decide such issues. Moreover, there is no clear consensus among policy-makers and regulators on this issue. For instance, in the European Union it is routinely accepted that subsidies to renewable electricity production (i.e., payments that occur in addition to market remuneration) should be included in the electricity tariff,[36] while in the United States those costs are mostly recovered by federal, state, and/or local taxes. In principle, we accept any policy costs that regulators and policy-makers have decided to include in every specific tariff, but we shall highlight the potential consequences of incorporating a large percentage of these costs into electricity tariffs.

In the discussion that follows, we include both residual network and policy costs together since both share two key features. First, both are costs that we assume regulators have decided to recover via electricity tariffs (later in this section we will discuss whether some such costs could be recovered by other means). Second, for each of these categories there is no clear criterion for assigning costs to particular electricity injections or withdrawals (at any given time or location, or aggregated over a period of time) since the economically efficient signals of prices and charges associated with energy and network services have been identified already. It is therefore clear that the most important criterion for the allocation of these residual network and policy costs is to minimize distortions of the already defined economically efficient signals (prices and charges) for energy and electricity services.

4.4.5.1 Cost-reflective policy cost charges

In cases where changes in electricity consumption relate directly to the cost of public policies, such as renewable energy obligations that require a fixed share of electricity from renewable energy sources, it would be efficient to allocate some portion of these costs in a manner that reflects the impact of electricity use on the marginal cost of these policies. For example, many jurisdictions have established renewable energy obligations or renewable portfolio standards policies, which require utilities or retailers to produce or procure a percentage of their electricity from qualifying renewable sources by specified dates. National level policies also often specify national renewable energy targets or obligations (for instance, the various national commitments made as part of the European Commission's "20-20-20" policy targeting 20% renewable energy across Europe by 2020.

Increases (or decreases) in electricity demand directly increase (or decrease) the marginal cost of compliance with such policies, as explained in Batlle (2011) and further developed in Batlle et al. (2016).[37] For example, in the case of a 20% renewable electricity obligation, increasing total electricity demand by 10 kWh would require an increase of 2 kWh of electricity supplied by qualifying renewable electricity sources. If these resources receive public policy support in the form of direct subsidies or tradable renewable certificates (also known as renewable energy credits), then the

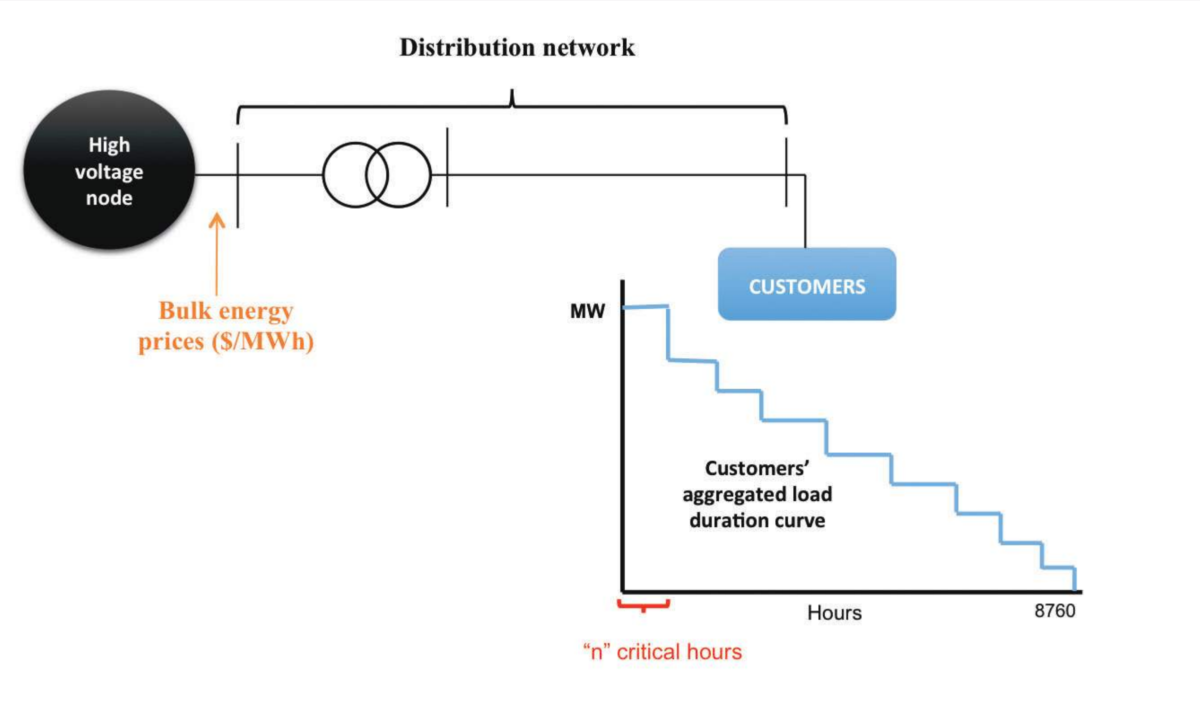
36 The European Commission has so far considered that directly charging electricity consumers for subsidies to renewables is compatible with the European rules on state aids and the environmental-polluter-pays principle (Ecofys 2014). We suggest other possibilities later in this section.

37 Note that if the policy objective is established in terms of a fixed quantity requirement completely decoupled from the evolution of consumption, then there would not be room for a cost-reflective policy cost signal. Changes in electricity consumption would not affect the cost of compliance with this fixed quantity requirement. In such cases, all costs would have to be allocated as residual costs in a minimally-distortive manner, as discussed below.

Box 4.6. Illustrating the Advantages of the Forward-Looking Peak-Coincident Network Capacity Charge

The simplified distribution network system in **Figure 4.18** is used to illustrate the advantages of using the "efficient ideal" peak-coincident network capacity charge. The system consists of a distribution network connected to a high-voltage node, which provides electricity to one consuming node, representing a group of customers in a medium-voltage distribution feeder. The distribution network has a fixed cost (X), and it is connected to a high-voltage substation from which electricity is supplied at the corresponding bulk energy LMP. The peak demand in this distribution network is expected to increase in the following years. Due to the discreteness and economies of scale of network investments, the network would need to be reinforced with a capacity higher than the expected load increase, with an estimated incremental cost that would amount to an annuity (Y). In addition to purchasing power supplied by the grid, the customers have different alternatives to respond to electricity tariffs (e.g., back-up generators, batteries, diverse forms of demand response), which would be employed if the use of such resources reduces customers' total cost of electricity supply.

Figure 4.18: Simplified Power System Configuration Illustrating the Impact of the Design of Network Charges



The effect of three network tariff designs is discussed: (1) an "efficient ideal" approach, consisting of a forward-looking peak-coincident charge computed as the incremental costs (Y) divided by the expected load growth (as proposed earlier in this section), plus a fixed charge per customer computed based on a Ramsey-like approach (see **Section 4.5**) for collecting the remaining residual network costs (Z); (2) a volumetric charge ($/kWh) to recover the entire network cost; and (3) a fixed charge to recover the entire network cost. The forward-looking peak-coincident charge would be applied to the total load at critical hours at which a further increase in load would require the network reinforcements. The residual network cost (Z) would be equal to the total network cost (X) minus the amount recovered (W) with the peak-coincident charge during the critical hours, that is Z=X-W (all of these values are annuities).

The customers are assumed to optimally react to each tariff design by minimizing their total cost of electricity supply from the grid plus the cost of any deployed responses to the received economic signals. When presented with the efficient ideal approach, the customers would reduce their load during critical hours (through any of the available options for responding) if the cost of these alternatives is lower than the bulk energy price plus the forward-looking peak-coincident charge. In this case, the network reinforcements would be avoided if the expected load growth is sufficiently reduced such that it can be supplied with the existing network capacity. Otherwise, if the cost of the response alternatives is higher than the cost of the energy supplied from the grid, including network charges, the customers would buy all the energy from the grid and pay the associated costs. As a result, the efficient ideal approach successfully guides the agents to find the most efficient solution for the system. In contrast, if total network costs (X) are allocated based on a volumetric charge, the customers may find it beneficial to invest, for instance, in a backup generator to partially self-supply electricity and reduce their volumetric charges, with no specific incentive to focus on reducing load or network use during peak demand hours. In this case, the system cost in general would increase, since the tariff may incentivize the deployment of a less efficient generator. Finally, with only a fixed network charge, the customers would decide to fully supply their load from the grid, as they face no marginal incentive to reduce consumption during peak hours. As a result, a fixed charge to recover total network costs could lead to more expensive network reinforcements in future periods.

Efficient network cost recovery may be comprised of three components: surplus from locational marginal prices (if any), a peak-coincident capacity charge, and a fixed charge to recover residual costs.

marginal policy cost associated with increased energy consumption by X kWh is equal to the level of public support for the marginal renewable energy resource times the renewable energy percentage specified by the obligation. Cost-reflective allocation of these policy costs would entail a volumetric charge that should be applied to the energy tariff. This charge would provide demand with an efficient signal indicating the impact of changes in electricity consumption on the marginal costs of such policies. The charge would also appropriately value and compensate energy efficiency and operations of DERs that reduce total energy consumption and lower the cost of compliance with these policies.[38]

4.4.5.2 The allocation criterion

Policy costs are a component of electricity tariffs in every jurisdiction (see **Figure 4.2**), and residual network costs are expected to be a significant fraction of the network costs in most power systems (as we have seen in **Section 4.4.4**). Therefore it is important to find a sound allocation method to recover these costs in the tariffs. As discussed, the best criterion for choosing among several possible allocation methods is the minimization of distortions of efficient prices and charges for the energy and network services. For instance, recovering the residual network and most policy costs with a flat volumetric rate (some amount of money per kWh consumed, regardless of the time or the location) can result in a significant efficiency distortion, because the signal modifies the price of energy at all times and at all network nodes. This approach also invites network users to net out their demand for electricity by installing behind-the-meter generation in an inefficient manner. By canceling electricity demand with embedded generation, those customers thus avoid paying residual costs, which have to be reallocated to other customers. This approach, in the extreme, can also incentivize grid defection.

The opposite approach would be to recover policy costs via a fixed charge (e.g., a lump sum determined annually by the regulator for each customer and then charged, for instance, in equal monthly installments). In principle, this annual lump sum would not distort short-term efficient signals,[39] which is an excellent feature. However, there are obvious equity implications: Should all network users pay the same charge, irrespective of their energy consumption or their peak or contracted power (if any)? Network users that consume more energy are likely to be wealthier than users that consume less energy, which means an equal fixed charge for all agents would disproportionately affect low-income users, which would be socially unacceptable.[40] Therefore, although an annual lump sum that is not directly linked to electricity consumption is an efficient instrument, how to allocate this sum to customers requires further consideration.

Under cases like this where there is no clear criterion for cost allocation, we can resort to the principle underlying Ramsey pricing.[41] Ramsey pricing advocates the use of the inverse elasticity rule to optimize efficiency (i.e., maximize welfare) in allocating the costs of monopolies or of some prescribed amount of taxes, under some conditions. This principle has been implicitly used by some countries in electricity tariff design, whereby costs such as policy and residual network costs (which cannot be assigned on the basis of any cost causality or responsibility principle) have been primarily allocated to residential users (in developed countries) or to large industrial consumers (in developing countries). In some jurisdictions this principle has been applied more explicitly, wherein proxies have been used as estimations of the elasticity of the different types of consumers. Ramsey pricing is discriminatory (Lewis 1941) but more efficient (in maximizing welfare) than other methods of allocating policy and residual network costs. Unfortunately, the intrinsic difficulty in implementing Ramsey pricing is the estimation of the price elasticity of electricity consumption for each network user class. This elasticity should measure the impact of rent reduction (due to the payment of the electricity bill) on household welfare, which, in simpler terms, is an indication of

38 Note that this cost allocation method reflects the marginal cost of compliance with the policy. The average cost of compliance may be higher or lower, depending on changes in the marginal cost of renewable energy over time and as the renewable obligations are progressively met. Total revenues collected via these cost-reflective charges for policy cost do not equal the total costs of compliance with the policy, adjustments to the fixed portion of customer bills (either lump-sum credits or surcharges) can be made to ensure costs are not over- or under-recovered.

39 Unless the regulator makes the mistake of computing the fixed charge for each customer in proportion to the annual energy consumed or the peak demand, for instance.

40 This statement has multiple exceptions, such as scantily used vacation homes owned by wealthy people or very large poor families that live in larger-than-average apartments. Any implementation that is based on this criterion would need ad hoc verification.

41 Frank P. Ramsey found such a result in 1927 in the context of taxation, see regulationbodyofknowledge.org/tariff-design/economics-of-tariff-design/ramsey-pricing/.

wealth. One viable proxy for wealth, which perhaps has the advantage of being perceived as equitable in the design of electricity rates and is also indirectly related to electricity consumption, is the customer's property tax or the property size,[42] and either of these proxies could be used to implement a Ramsey pricing approach. Other similar metrics might also be acceptable.

The presence of a substantial amount of policy and residual network costs in the electricity tariff has several negative consequences and invites a serious reflection on this aspect of tariff design.

4.4.5.3 Costs to be included in the tariffs

There are several negative consequences of including residual costs that are not directly affected by changes in electricity consumption in electricity tariffs. First, there is a very clear risk of efficiency loss, if these costs are allocated in the tariffs in such a manner that they distort cost-reflective prices and network charges. Second, if the adopted tariff design is such that the policy and residual network costs end up increasing the volumetric component of the price of electricity for the end customers, this will have the effect of discouraging the use of electricity with respect to competing alternatives and slowing down the transition to a more electrified future, where electricity appears as the dominant energy vector, generated from carbon-free resources and supplying also the needs for mobility and space conditioning of buildings. Finally, even if our recommendation on how to include the residual network and policy costs in the tariff is followed, we have to be aware of the potential implications that loading the tariff with policy and residual network costs could have on grid defection (see **Box 4.7** and the following discussion), perhaps incentivizing defections that result in a reduction of social welfare that would not have happened if residual network and policy costs were lower.

As indicated before, this report will not make recommendations on which policy costs should or should not be included in the electricity tariff, since this is very jurisdiction-dependent. However, since the drawbacks of including these costs in the tariff have been indicated, it seems pertinent to reflect on this topic.

Every cost item within the residual network and policy costs should be subject to scrutiny. In many jurisdictions, the subsidies to promote low-carbon technologies — renewable generation, in particular — are a significant component of the policy costs in the electricity tariff. Climate change is the major justification for incurring these costs. Certainly, electricity production is an important contributor to greenhouse emissions, both in the United States and in the European Union, but other sectors of the economy are as well. Since electric power appears to be easier to decarbonize than other sectors, a strong effort is in progress to decarbonize electricity faster and more drastically. It therefore seems justifiable to share the burden of electricity decarbonization — i.e., the corresponding policy costs — with customers of other sectors that are spared this extra effort (see Batlle 2011 for a detailed discussion). In the United States a substantial fraction of the cost of renewable support schemes is born by federal, state, or local taxes.

A more radical and rather speculative proposition is to question whether the costs of the electricity network, both for transmission and distribution, could be paid

42 There is a causal relationship between the amount of energy consumed and the need for more renewable production or more energy efficiency measures, since the targets for these measures are frequently expressed as a percentage of total energy consumption. Unless this relationship is clearly established, we recommend avoiding recovering these costs with a charge in the tariff that is proportional to the energy consumed, as this would artificially modify the short-term price of energy. Rather, we recommend the use of a proxy, such as a property tax, to allocate costs. Using this type of proxy would ensure that allocation is related to energy consumption without modifying short-term signals for electricity services.

by the taxpayers instead of the power system agents.[43] This topic would become relevant in the grid defection debate, to be discussed next, if the reduction of costs in distributed generation and storage turns grid defection into a widespread, economically attractive proposition in some power systems.

4.4.5.4 Implications on grid defection

The present trend toward widespread availability and decreasing cost of distributed generation and storage results in the possibility of grid defection (i.e., radical disconnect from the grid),[44] from a physical (for customers that meet some conditions, such as being able to install some embedded generation within the household or business premises) and economic (given the substantial costs to be avoided by defection) standpoint. Grid defection is an extreme form of elasticity to price, and it has to be taken into account — from an efficiency perspective — when deciding the regulated costs to be included in the electricity tariff and the design of the tariff itself.

The implications of tariff design on grid defection have to be carefully examined. The economic viability of grid defection depends critically on the volume of costs that can be avoided by disconnecting from the grid, which in principle include all the costs of supplying the energy and network services plus the policy costs. **Figure 4.19** shows the costs and the savings of grid defection from the customer's viewpoint. The potential grid defector has to compare the savings achieved by defection with the costs of independently supplying electricity with a satisfactory level of reliability. Obviously, personal preferences will influence a customer's final decision.[45]

What a sound tariff design should try to avoid is inefficient grid defection, i.e., one that reduces social welfare. Grid defection has implications for the remaining network users, which have to shoulder whatever costs are not paid by the defectors. Thus, tariff design should try to ensure that no excessive or ill-assigned costs in the tariff make separating from the grid attractive, particularly if better-designed economic signals could prevent defection.

A measure that can be explored — although we are aware of the potential legal obstacles — is to establish that those who defect from the grid should continue paying residual network charges (e.g., the cost of any network assets left stranded by defections) or perhaps receive some credit (if, by disconnecting from the grid, they contribute to the deferral of network investments), and maybe pay some other policy costs (guaranteeing that all customers contribute to recover an appropriate share of the policy costs underwriting established societal objectives) via an exit charge, thus ensuring that other network users do not have to pay for these costs, increasing their tariffs. Some challenging design decisions are necessary to implement exit charges: How much of the estimated cost of any stranded investment should the defector pay? Is this just one lump payment and, if not, for how long should the defected customers have to pay this charge? Would these consumers continue receiving electricity bills from the utility?[46]

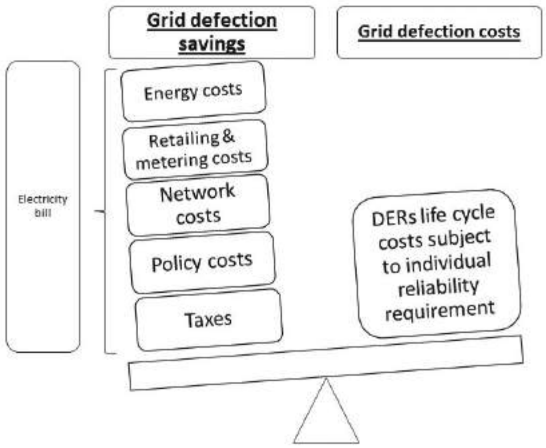
43 We are not proposing to give up the network-related short-term economic signals (LMPs, in detailed or approximated form) or the long-term ones (peak capacity network charges), which are necessary for efficient operation and investment, but to consider whether the majority of the network costs (the residual network charges) should be included in the electricity tariff or not. In the same way that public schools are paid for by all taxpayers, regardless of whether or not they have children, and public libraries are supported by those who never use them, and most roads are paid by general taxes regardless of how much each taxpayer uses them, the existence of an electric network reaching to every building can be considered a necessary basic infrastructure in any developed society, and thus paid for via taxes and not specific electricity rates. Access to connection to the electric network is an added value for real estate, even if its owners decide not to connect to the grid and to self-supply electricity for their own needs. Note also that there is already a high level of averaging or cross-subsidization in the allocation of the distribution network costs in the tariffs of every power system, whereby all customers connected at the same voltage level are subject to the same tariff, regardless of the actual distribution costs in their location.

44 The refusal of a new agent to connect to the network can be seen as another form of grid defection, with different implications.

45 A customer's decision to defect may be driven by other motivations besides strict monetary savings, such as a strong preference for renewable technologies or the satisfaction derived from being energy self-sufficient. These motivations are beyond the scope of the present discussion.

46 An additional difficulty that arises when addressing these implementation issues is the difference in the economic lives of the assets involved on both the electric grid side (lines, transformers, large generators) and the customer side (PV panels, batteries, heat pumps, ICT investments, etc.) — and the expected time horizon for investment recovery.

Figure 4.19: Costs and Benefits of Grid Defection from the Customer Viewpoint



Regulators and policy-makers must carefully monitor for conditions that could lead to a serious threat of inefficient grid defection. If these arise, they must reconsider the costs that are included in the tariff as well as other measures to prevent substantial cross-subsidization among consumers and a potential massive grid defection with unforeseen consequences.

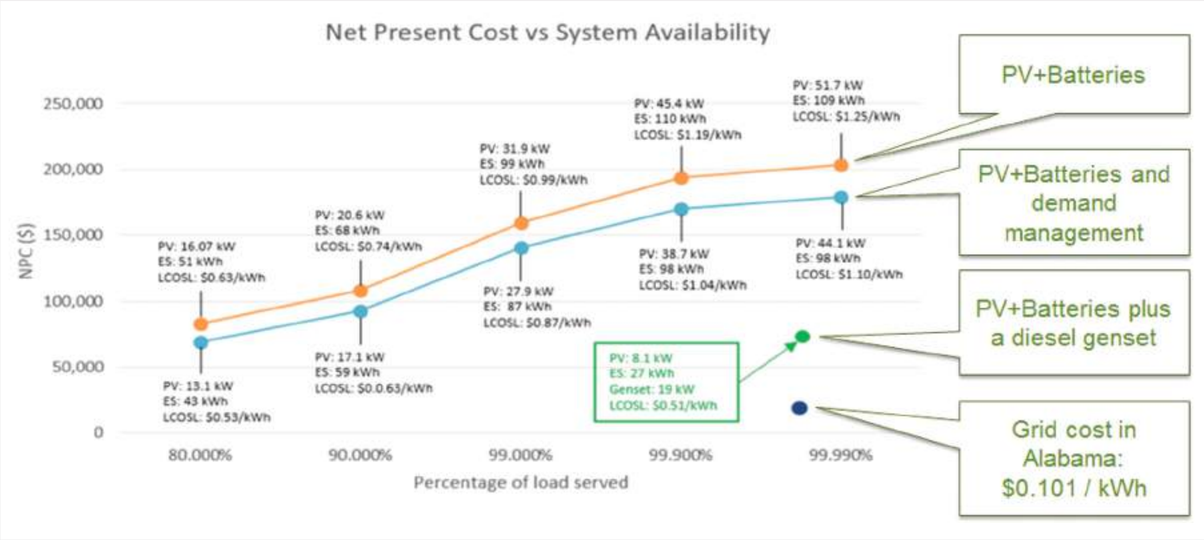
Box 4.7: Grid Defection

With current technologies, system-wide, efficient, and grid-isolated solutions rarely happen in developed countries as a well-developed, interconnected electricity system takes advantage of economies of scale (of electricity generation resources and networks, risk and system management, and the complementarity of energy resources throughout a large territory) to provide a highly reliable electricity supply at a cost that is difficult to improve upon off-grid. This is why electricity systems have become increasingly interconnected.

Some studies, however, claim that customers may find it economically attractive to disconnect from the grid in the near future in some power systems in developed countries (Rocky Mountain Institute, Homer Energy, and Cohnreznick Think Energy 2014; CSIRO 2013). Most of the published studies on grid defection forecast future retail prices (based on current cost components) and compare those prices with the estimated average cost of an isolated system. These studies usually underestimate the costs of the typical, low-reliability, off-grid systems. EPRI (2016) estimates the net present costs of off-grid supply for a US single-family household (**Figure 4.20**) for different estimated levels of reliability (as a percentage of load served). The results show that the isolated system is more than 10 times more expensive than the grid-connected one for the same reliability level, and still about five times more expensive if a very low reliability level (80 percent) is considered acceptable for the off-grid system.

In addition, off-grid systems would waste plenty of renewable resources, which otherwise could be exported to the grid, thus reducing the cost of the system. When a diesel generator is used in addition to the PV and battery, the reliability reaches 99.97 percent load served, similar to the cost of PV plus storage with 80 percent load served. However, the cost is still several times more expensive than purchasing the power from the grid, and, moreover, the amount of carbon and other pollutant emissions dramatically increases, offsetting the benefits of installing renewables.

Figure 4.20: Off-grid System Costs



Source: EPRI 2016

Net present cost and levelized cost are shown per kilowatt-hour for off-grid systems with solar photovoltaics and storage, with and without demand management, and with a diesel generator for different reliability levels.

The costs included in a tariff are critical to determining the economic viability of grid defection. For instance, the Australian Energy Market Commission recognizes the need for adjusting network tariffs as “there is a high risk of grid defection” (AEMC 2016). The situation might evolve into what has been termed the death spiral (Costello and Hemphill 2014; Felder and Athawale 2014; Graffy and Kihm 2014), whereby the incentives for grid defection increase with the number of defecting customers because residual network costs and policy costs are borne by fewer remaining customers.

Box 4.8: Cybersecurity-Related Costs

The computation of appropriate prices between DERs and transmission system operators (TSOs) or distribution system operators (DSOs) requires consideration of cybersecurity. TSOs/DSOs will have the responsibility for monitoring nonutility DER suppliers and assuring malware or cyber attacks are not introduced into the TSO/DSO control or pricing systems. Cyber risks can come from DER control systems, demand response systems, dispatchable energy storage systems, smart meters, and other DERs that communicate with TSO/DSO control systems. As DERs participate in the price formation by making buying and selling bids in different markets, digital connections to utility pricing and accounting systems will need to be sufficiently protected and monitored. Fixed charges can be used to recover the cost of meeting protection standards (NIST 2014).[47]

Cybersecurity is also an important factor in determining appropriate communication system interfaces between DERs and TSOs/DSOs. Data interface and protocol standards, communication architectures, and advanced control algorithms should be established that are consistent, have strong encryption, require multifactor authentication, and introduce security by design progressively in all new equipment. Communications between DERs and TSOs/DSOs need to be monitored for anomalies to avoid unexpected disruptions of power, and DER controls, as well as systems that interface with TSOs/DSOs, should allow approved upgrades to enable enhanced data interface standards and to fix identified cybersecurity flaws. TSOs and DSOs could provide secure communications equipment (at charges agreed upon with regulators), and charges later would need to be allocated to customers.

47 For example, US North American Electric Reliability Corporation critical infrastructure protection standards, European Commission standards, state Public Utility Commission standards, or National Electric Sector Cybersecurity Organization Resource suggested requirements.

4.5 The Potential Benefits of Getting Prices and Charges (Almost) Right

Thus far we have presented key principles and implementation methods for the development of a comprehensive and efficient system of prices and charges. What now remains to be evaluated is the extent to which simplifications in the implementation of a comprehensive system of prices and charges — arising from practical challenges that may hinder a strictly efficient approach — can still yield outcomes that are acceptable from an efficiency standpoint. We present here several quantitative examples that illustrate our assertions and support our recommendations. The numerical results described here depend to a great extent on the assumptions made about specific power system conditions, so while we caution against extending the implications of these experiments too broadly, we can clearly conclude that granularity matters, and that a moderate level of sophistication in the computation and use of prices and charges can capture most of the efficiency lost with spatially uniform and temporally constant economic signals.

Again, as indicated in the introduction to this chapter, those who are just interested in the principles (**Section 4.2**) and the recommendations (**Section 4.6**) can omit reading this section, which consists of a selection of case studies showing the relationship between granularity level and efficiency gains.

A substantial body of literature explores customer response to varying granularity of energy prices, generally suggesting that the benefits to be realized from the responses of customers opting to utilize TOU or CPP rates justify costs of implementation such as advanced metering, communication, and control costs — which have fallen significantly in recent years as advanced metering rollout has expanded (Joskow and Wolfram 2012). Here we are exploring more carefully structured rates, which we have shown in the last section may substantially differ from these simple tariff designs.

The first example examines the impact that increasing the granularity of energy prices in time (from fixed to hourly prices) and of location in the distribution network (from uniform prices to LMPs at each point of connection) has on investment in and operation of DERs that can locate anywhere in the network.

The second and third examples show the responses of representative flexible consumers when they are subject to different types of electricity tariffs in power systems modeled under different sets of assumptions. The responses of the consumers are fed back into an operational model of the power system, eliciting changes that can either improve or reduce overall efficiency. The objective of these experiments is to understand and assess the impact of increasing levels of granularity on efficiency at the customer and the power system levels.

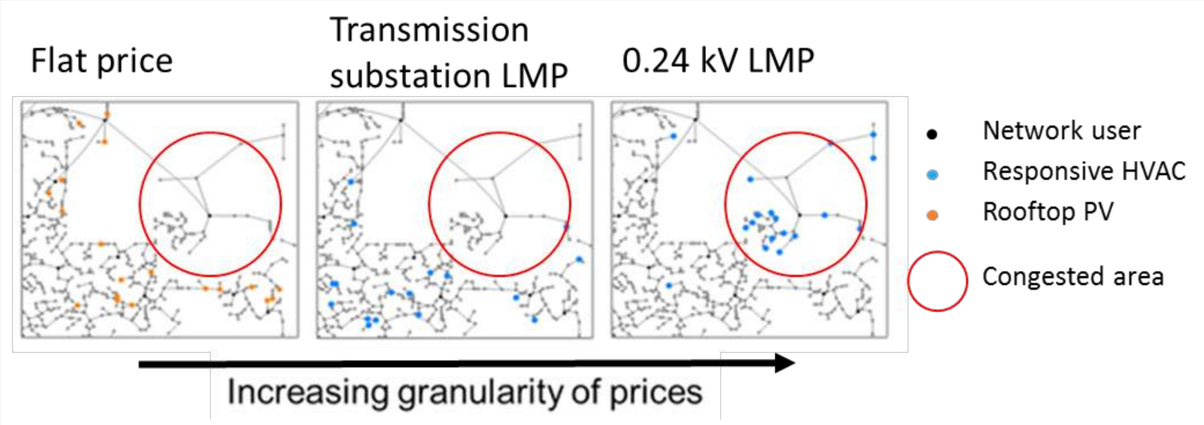
While it seems reasonable to suppose that exposing power system agents to the costs they cause in the system will elicit efficient behaviors that reduce those costs, predicting human behavior is in general very complex. In order to gain any insights into the potential efficiency of different levels of locational and temporal granularity, we assume perfect economic rationality of the end-users — i.e., that each end-user responds to economic signals in a way that maximizes his or her economic profits. This is consistent with the plausible vision of the power system described in **Chapter 3**, whereby aggregators efficiently manage the responses of the end consumers to the system conditions while respecting their preferences, so that their responses are economically efficient without the need for direct human supervision.

Box 4.9: End-user Response to Prices: The Effect of Locational and Spatial Granularity

Using a detailed model of a distribution network, which includes a variety of building types as network users (see the description of the D-SIM model in **Appendix A**), we computed energy prices with different degrees of spatial and temporal granularity and let each network user make optimal (i.e., cost-minimizing) investment and operational decisions. Our results, shown in **Figure 4.21a**, suggest that the type and location of DERs likely to emerge in a distribution network are very sensitive to the structure of the energy price. Furthermore, locational prices (distribution-level LMPs) can effectively indicate which locations will have the highest network benefits, potentially relieving network constraints at a cost that is less than traditional approaches (e.g., reinforcing wires, upgrading transformers, etc.).

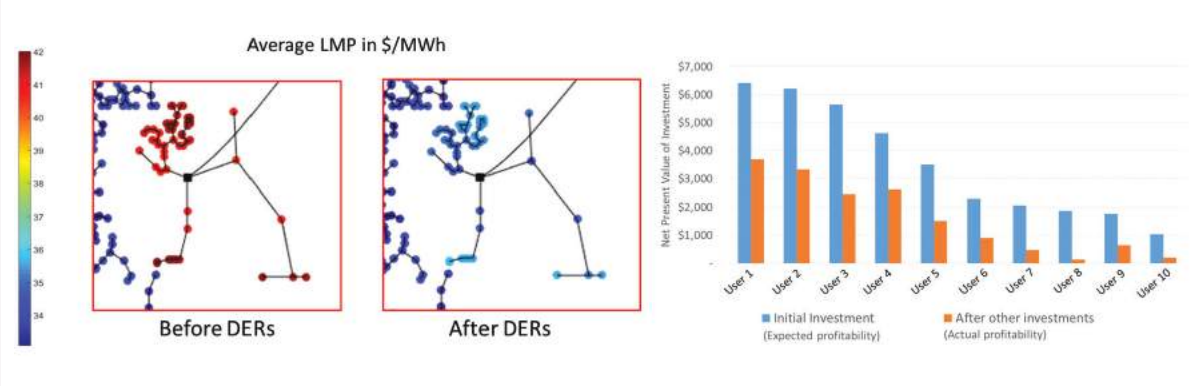
Figure 4.21: The Effect of Increasing Granularity of Prices on DER Deployment

4.21a: DER investments under different energy price structures



4.21a shows that the most attractive DER investments change under different energy price structures. Each panel in the figure corresponds to the same distribution network. Flat prices, for example, lead to investments in solar photovoltaics (PV) with no discernable geographical pattern. In contrast, LMPs calculated down to the meter lead to investments in heating, ventilation, and air conditioning (HVAC) controls that are clustered around the area of congestion.

4.21b: DER Profitability in the Presence of Surrounding DER Investments

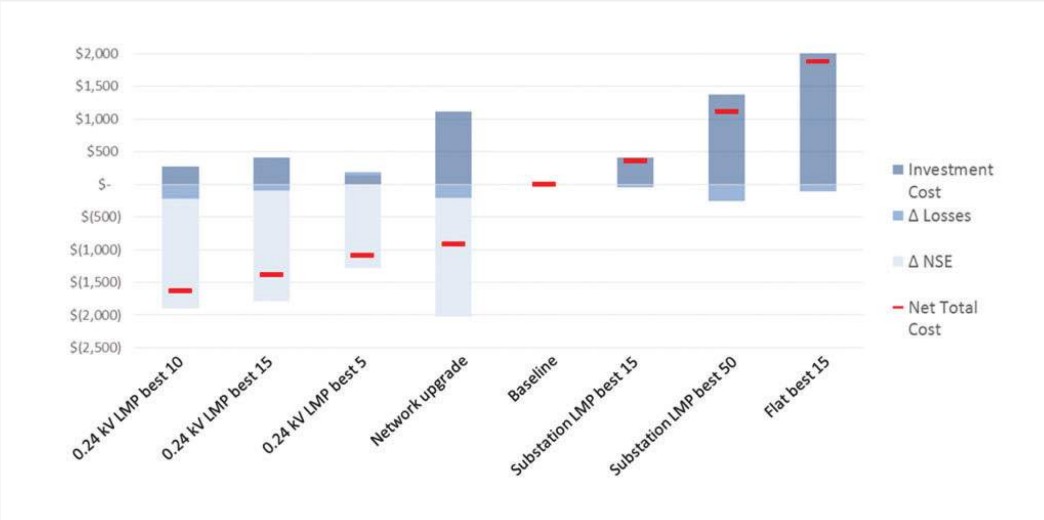


4.21b shows that by alleviating congestion in response to LMPs, early investments in DERs decrease the appeal of future ones. This figure compares the *ex ante* and *ex post* economic evaluation of smart heating, ventilation, and air conditioning systems for 10 network users.

An important observation is that the net present value of a particular DER investment in response to granular LMPs diminishes quickly as other similar investments are made in the same area of the distribution network. As we show in **Figure 4.21b** for this particular case, the expected savings for a single network user from the installation of a price-responsive heating, ventilation, and air conditioning (HVAC) system are considerably smaller if we assume that other neighbors will do the same. Simultaneously, network users who did not invest would still benefit from the exposure to lower LMPs.

As explored further in **Chapter 8**, the ability of DERs to address network constraints as an alternative to traditional "wires" solutions is a key driver of the economic value of DERs. **Figure 4.22** compares the costs of investing in DERs to the costs of network infrastructure investments (in the network modeled above) under a variety of economic signals. As shown in Figure 4.22, the lowest-total-cost solution occurs when the best 10 DER investments, in terms of net present value, are installed under the LMP pricing scenario. Looking back at Figure 4.21a, those investments are HVAC controls, all located in the area of congestion. The traditional utility solution — replacing the transformer — actually results in greater operation benefits in the form of reduced nonserved energy[48] and losses, but carries a higher upfront investment cost, making the total net cost higher than the best DER solutions. DER investments that emerge under flat and substation-level LMPs result in some reduction in losses, but because they are located outside of the congested region (see Figure 4.21a), they do not reduce nonserved energy, and thus produce higher net costs.

Figure 4.22: Net Costs for Congestion Solutions



This figure shows the net costs of different investment decisions related to the alleviation of congestion in the distribution network. In the baseline situation, a certain amount of nonserved energy (NSE) is needed to respect network constraints. Each alternative scenario has a certain upfront cost and is to some extent effective in reducing the losses and NSE with respect to the baseline scenario. The red bars correspond to the net cost-benefit balance with respect to the baseline.

48 Nonserved energy is assumed to come from the continuous forced disconnection of load at a high price. As explained in **Appendix A**, this approximation is necessary to compute LMPs as the dual variables of the nodal power balance constraints. On the building side, the cost of temperature deviations was made very high, making them effectively hard constraints.

Increased granularity in energy prices is a sound way to identify DER investments that can help alleviate network problems — a function carried out in a similar fashion by LMPs in the transmission network. However, given the strong coupling between the expected profitability of investment decisions made by multiple agents, along with the magnitude, irreversibility, and long asset life of such investments, it would be very difficult to prevent inefficient substitution and free-riding. Instead, the necessary coordination could take the form of long-term contracts, cleared through public auctions, between the distribution company and providers of alternatives to network reinforcements, including DERs, perhaps via aggregators. See **Chapter 8** for more examples of the system-wide value that DERs can provide when efficiently located and operated.

Increased granularity in energy prices is a sound way to identify DER investments that can help alleviate network problems. However, since DER investments may experience rapidly diminishing marginal returns, influenced by surrounding DER investments, network users may need to coordinate among one another and with the distribution company to prevent inefficient substitution and free-riding.

Box 4.10: End-user Response to Prices — Impact on Customer Bills, System Costs, and Utility Cost Recovery

The objective of this experiment is to assess the impact of different residential electric tariff designs on customer bills, system costs, and utility cost recovery. Several tariffs are evaluated, including traditional flat rates, flat rates with a noncoincident peak (NCP) demand charge, time-of-use (TOU) rates, critical peak pricing (CPP), and our approximation of an efficient ideal approach (EIA), which, following the recommendations in this chapter, allocates energy costs based on hourly wholesale LMPs[49] with adjustments to reflect distribution network losses;[50] allocates the incremental portion of network costs as a peak-coincident network charge and generation capacity costs as a scarcity-coincident capacity charge; and allocates the remaining residual regulated costs as a fixed per-customer charge.

We model two cases, one reflecting a high-load-growth scenario with higher peak-coincident generation capacity and network capacity charges and a low-load-growth scenario where these charges are more modest.[51] We simulate a residential building in Westchester County, New York,[52] with electric cooling using the Demand Response and Distributed Resources Economics Model (see **Appendix A**). The building's air conditioning (AC) demand is controlled by a home energy management system or "smart

49 Historical 2015 hourly wholesale energy prices for NYISO Zone I are used.

50 We assume average losses in a distribution network are 5 percent and use a quadratic function to estimate marginal losses in each hour as a function of historical 2015 NYISO Zone I aggregate load (e.g., Westchester County, New York).

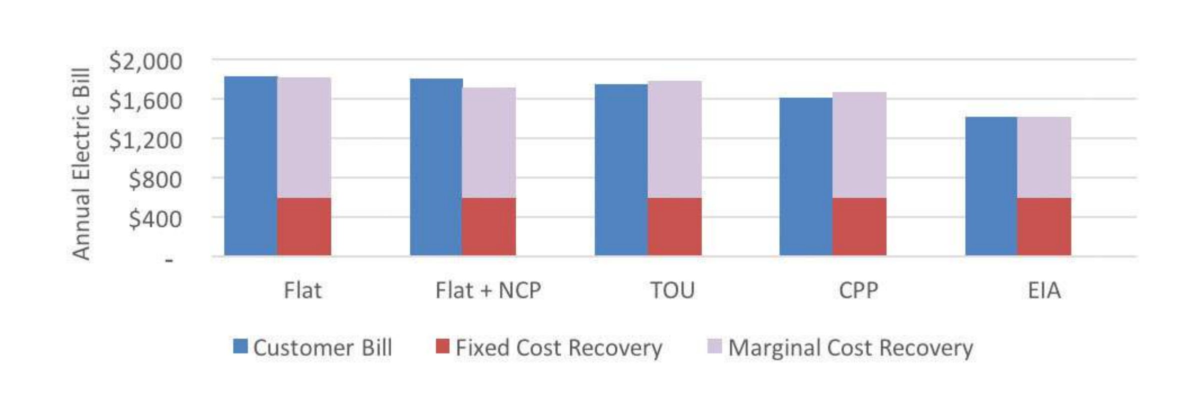
51 Generation capacity costs are assumed to be $70 per kilowatt-year (kW-yr) in the low-growth case, reflecting 2013–2015 average NYISO capacity clearing prices, and $115/kW-yr in the high-growth case, reflecting NYISO load and capacity forecasts for 2016–2026 (NYISO 2016). Distribution network capacity costs range from $25/kW-yr in the low-growth case to $120/kW-yr in the high-growth case, reflecting two values within the range of transmission and distribution deferral benefits estimated in several regulatory avoided cost filings (e.g., Neme and Sedano 2012). Costs are converted to peak-coincident charges by allocating annual costs proportionately to the top 100 hours of aggregate demand at the NYISO zonal level for generation capacity charges and top 100 hours of average residential load profile for the distribution network charge.

52 The household consumes 6,600 kWh annually and is equipped with a 4 kW central air conditioning system.

thermostat," which uses optimization to schedule the AC unit's operation based on an expectation of future hourly energy prices and ambient temperatures, with the goal of minimizing the customer's electricity bill subject to occupancy comfort preferences. We model thermal flexibility using a simplified building energy thermal model derived from Mathieu et al. (2012), with temperatures varying around a target set point (e.g., 70°F/20.5°C). Penalty factors within the model's objective function account for the comfort impact of deviations from the target temperature set point and represent the opportunity cost of shifting cooling demand.[53] The building optimization model is confronted with different tariff structures, and the resulting load profiles are used to calculate customer bills and to assess the impact of load changes on system costs and utility cost recovery.

The results show that the efficient ideal approach provides the lowest customer bills, the greatest reduction in the user's contribution to marginal system costs, and full recovery by the utility of the costs to serve the customer. Simplified alternative tariff designs show reductions in customer bills compared to traditional flat rates but result in either under- or over-recovery of utility costs and thus (implied) transfer payments. These results are summarized in **Figure 4.23** for the high-demand-growth scenario, in which a larger share of total network costs are considered marginal (rather than sunk or residual), and generation capacity costs are higher, reflecting expected capacity investments to accommodate peak demand growth. As Figure 4.23 illustrates, under the TOU and CPP tariffs, the customer's bill falls short of fully recovering the marginal power system costs incurred to serve that customer, including marginal energy costs and marginal or incremental generation and network capacity costs, as well as the share of total residual costs allocated to the customer. In contrast, under the flat tariff (Flat) and flat tariff with noncoincident customer demand charge (Flat + NCP), the customer's bill exceeds the marginal costs incurred to serve the customer and the residual costs allocated to that customer. Only under the efficient ideal approach does the bill accurately match the customer's contribution to marginal system costs and share of residual costs.

Figure 4.23: Summary of Customer Bills and Utility Cost Recovery by Tariff Schedule — High-Demand-Growth Scenario



53 We assume a 0.5°C of temperature deviation with no penalty, reflecting the fact that HVAC equipment naturally fluctuates around a target temperature. Deviations from 0.5°C -1.5°C are penalized at $0.1 per hour, while penalties increase to $0.5 per hour for 1.5°C -2.5°C deviations and rise steeply for each additional degree beyond, reflecting the fact that residential consumers have a demonstrated price elasticity for air conditioning and have shown a willingness to tolerate both precooling and higher temperatures during event hours (EIA 2014; Herter and Okuneva 2013).

Furthermore, customer bills under the efficient ideal approach are 22 percent lower than bills under the flat rate for the high-growth scenario and 6 percent lower for the low-load-growth scenario (where peak prices are not as high). Changes in the user's consumption patterns also reduce marginal system costs (including energy costs and marginal or incremental generation and network capacity costs) by 33 percent under the cost-reflective tariff as compared to flat rates for the high-growth scenario, and by 17 percent under the low-growth scenario. This typical household could save as much as $400 per year on electricity bills solely by employing flexible AC strategies in response to improved tariffs,[54] with system costs incurred to serve this customer falling by an equal magnitude. Greater savings could be achieved by employing other smart appliances, price-responsive demand curtailment, cost-effective energy efficiency measures designed to reduce consumption during high-price periods, and other cost-saving methods. In aggregate, more efficient, cost-reflective tariffs could unlock billions of dollars in savings at a state or national scale by reducing the need for investment in new network or generating capacity, and by reducing consumption during periods with the highest marginal energy costs.

Box 4.11: Aggregated End-user Response to Prices and Charges — Impact on System Operational Costs

The design of prices and charges not only impacts the dispatch of flexible distributed resources, but it can also affect the operational costs of the bulk power system. The outcomes of five approaches to allocating the energy, network, and policy cost components of retail prices are compared using the ROM model (see **Appendix A**). A system similar to Spain as envisioned in 2020–2025 constitutes a reference case, in which a group of flexible customers are assumed to be located within low-voltage networks. The reference system has an annual energy consumption of 254 terawatt-hours, a peak demand of 40 gigawatts, and installed generation capacity of 100 gigawatts. Customers can make use of several energy resources: solar photovoltaics, electric vehicles, demand response, and backup generation units. In the simulation, these customers can lower their electricity bills by dispatching their DERs in response to the locational marginal prices and charges that they receive.

Five cases, together featuring different ways to allocate costs, are compared, and the results presented in **Figure 4.24**.

First, the reference case allocates regulated costs as fixed charges (euros per customer). Therefore, flexible consumers respond only to hourly locational marginal prices (LMPs). In this reference case, there are no peak-coincident network charges, the total annual network costs amount to 9.5 billion euros, and policy costs amount to 15.7 billion euros. Operational costs in this reference case are lowest; however, future incremental network costs will be highest due to the fact that agents do not experience a peak-coincident network charge (and therefore do not change behavior when they are operating in a way that is costly to the network).

The second case represents the proposed efficient ideal approach (EIA), where the application of the peak-coincident network charge corresponding to the estimated costs of required network

54 See Chapter 8.3.1, for additional discussion of flexible AC strategies.

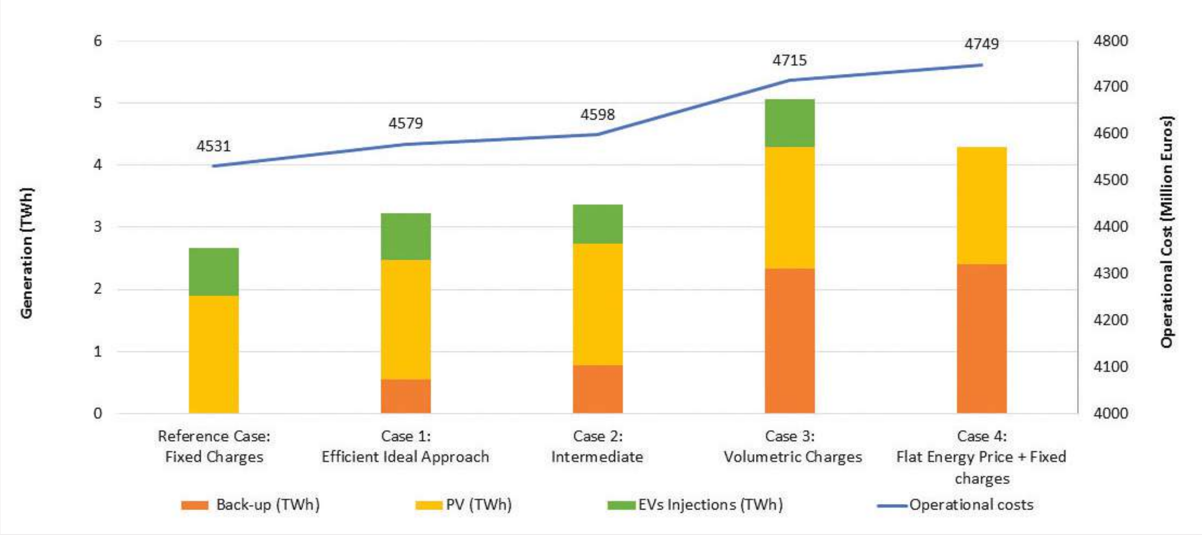
reinforcement happens to recover 20 percent of the total network costs. The residual network costs and the policy costs are recovered via per-customer fixed charges, and export from onsite resources is remunerated at locational marginal prices. Energy prices are computed as hourly LMPs. In this case, due to the fact that agents are exposed to peak-coincident network charges, operational costs are higher than in the reference case (due to increased on-site thermal generation that is dispatched by customers because to avoid the peak-coincident network charge). However, total system costs (not shown in Figure 4.24), would be lower due to avoided network costs that come about as a result of customers' responses to the peak-coincident network charge.

The intermediate case allocates the total network costs to a peak-coincident charge (instead of adopting the cost-reflective approach of the second case), and recovers policy costs with a volumetric charge. Energy prices are computed as hourly LMPs. In this case, operational costs are higher than in the previous two cases due to the distortionary impact of the volumetric charge and the excessive response of on-site thermal generation, without achieving further network cost reduction.

In the fourth case, all regulated costs are allocated on a volumetric basis (i.e., on a euro/kWh of energy basis), and exports to the system are remunerated at the retail price. Again, energy prices are computed as hourly LMPs.

Finally, the fifth case employs a flat average energy price for the whole year, together with volumetric charges to recover network and policy costs.

Figure 4.24: System Operational Costs and Flexible Resources Dispatch for Selected Cases



Hourly price differences incentivize electric vehicles (EVs) to charge when prices are low and inject energy to the network when prices are high, whereas flat energy prices do not incentivize EVs to inject energy since flat prices eliminate the possibility of arbitrage between prices. In addition, in the flat-price and volumetric cases, backup generation is dispatched at almost its maximum capacity, since on-site generation reduces payments of the residual network costs and policy costs embedded in the volumetric tariffs. In contrast, backup generation is not dispatched in the reference case because of its high variable costs.

Flexible customers can net demand with local generation, reducing the energy flow from the rest of the system, as well as energy losses and LMPs for the whole system — thus decreasing system costs and customer payments. Reductions in LMPs (and customer payments) are higher at the low-voltage nodes where flexible consumers are located, since the effects of their behavior permeate through more of the network. As indicated, under volumetric charges, reducing electricity imports from the grid allows flexible consumers to reduce their payments of regulated charges. The volumetric costs avoided through customers' modifications of their demand profiles must be collected from inflexible customers to ensure full recovery of regulated costs.

Well-designed prices and charges can incentivize investments in and operation of DERs in ways that yield lower costs for customers and the system as a whole. Ill-designed prices and charges can incentivize outcomes that yield lower costs for some customers but may not necessarily yield lower costs system-wide.

4.6 Implementation Issues

The material presented in this chapter has illustrated how prices and charges for different electricity services vary in time and by location and has demonstrated how economically rational agents respond to these economic signals in ways that impact the power system conditions that initially produced these very signals. More specifically:

1. There are significant temporal and locational differences in prices for electrical energy — i.e., active and reactive power — at the wholesale level that depend upon the operating conditions and designs of specific power systems. But significant variations in energy prices can occur within the distribution network as well, at all voltage levels. These locational energy prices internalize the costs of energy production as well as the impact of network losses and operating constraints.

2. Temporal differentiation—and in some power systems, locational differentiation as well—is also important for the prices of other relevant electricity services, such as operating reserves and firm capacity. In general, the contribution of these services to total costs in current power systems is significantly lower than that of electric energy (by one order of magnitude for firm capacity and even less for reserves). Charges for end-user responsibility in network investments are highly time- and location-dependent.

3. Demand and DERs can respond to all of these prices, creating value for themselves and increasing the overall efficiency of the power system. Demand and DERs can also sign contracts (individually or via aggregators) with system operators to commit to specific response patterns in relation to the provision of each of these services or any other services that system operators may require.

4. These three levels of differentiation ("granularity") in time, location, and the services that are considered, must be reflected in the design of efficient economic signals intended for the agents connected to the distribution network. Our simulations of the responses of agents to an ensemble of prices and charges featuring different granularity levels have shown significant variation in both operational and investment outcomes. This can translate into substantial gains in global efficiency for the power system and individual agents when granularity is increased.

5. When it comes to implementation of these improvements in the design of economic signals, two objectives dominate: (1) to find, for each one of the major services and the conditions particular to each specific jurisdiction, an adequate level of granularity along all dimensions such that there is a proper balance between gains in efficiency due to increased granularity on the one hand and transaction and deployment costs on the other hand; and (2) addressing distributional concerns such that, if desired,

more efficient economic signals do not result in significant changes in the total costs paid by agents, which typically result from extensive cost averaging and cross-subsidization among customers.

4.6.1 The efficient level of granularity

The fact that significant temporal and locational differences exist in the "efficient ideal" real-time prices and charges of the most significant services does not necessarily imply that the maximum level of granularity must be adopted to achieve a satisfactory level of efficiency. What follows is a discussion of ad hoc simplifications that may reduce implementation costs significantly without a significant efficiency loss. In some cases, the simplifications may be required because of the difficulty of modifying well-established or, on the contrary, very recently adopted regulation. Other simplifications should be avoided at all costs. In most cases, the following sequence of improvements toward "getting the prices and charges right" will capture low-hanging fruit first at reasonable implementation costs:

1. The first recommendation is to fix a flaw in tariff design that has important consequences. Policy costs and residual network costs should be removed from the volumetric ($/kWh) component of the tariff and charged differently, preferably as an annual lump sum (conveniently distributed in monthly installments), with the magnitude of each customer's charge dependent upon some proxy metric of lack of price elasticity or some measure of wealth, such as the property tax or the size of a system user's dwelling, or just via ordinary taxes.[55] The seriousness of the threat of grid defection must be carefully considered in determining which costs are included in the electricity tariff.
2. In most power systems, energy prices vary significantly over time each day (except in those systems with large amounts of hydro storage). In typical power systems with significant proportions of thermal generation, high penetration of intermittent renewables with zero variable costs increases this time variability. Hourly or subhourly energy prices can be easily communicated to end customers via the advanced meters or other affordable information and communication technologies expected to be deployed soon in all developed countries. This degree of temporal granularity in the wholesale energy price should be the second major objective of an enhanced tariff design.
3. Wholesale energy prices can be extended to all voltage levels of the distribution network by means of loss factors with different possible levels of detail or accuracy. Loss factors should change in time depending on the level of power flows in every direction, and they may be considered uniform per voltage level or differentiated by distribution zones or even feeders. At peak times (of demand or of distributed generation) loss factors may depart from unity significantly, creating a difference in value between centralized and distributed generation that could decide the financial viability of some DERs (at least it would impact their efficient response) and therefore must not be ignored
4. Depending upon the characteristics of the power system, peak capacity signals may become critically important to reduce the need for future investments in generation and networks. Once there is an advanced meter available that can record the hourly or subhourly injection or withdrawal of power at every connection point, there is no conceptual difficulty in applying coincidental peak capacity charges for firm capacity and for responsibility in network investment. This notably enhances the level of comprehensiveness of the system of prices and charges without adding significant implementation effort or cost, although the difficulty of regulatory design and acceptance should not be underestimated.
5. It has been demonstrated in US electricity markets that transmission LMPs can be computed precisely and frequently, and that sometimes the differences between LMPs can be significant between nodes of a power system at a given moment in time. These prices have excellent spatial granularity, but are presently only applied to generation, while broad average zonal prices are applied to the rest of the agents. This policy has to be reconsidered and the loss of efficiency evaluated, in particular with regard to the participation of DERs in the wholesale markets and the impact on price-responsive demand.
6. The current situation is very different in the EU wholesale market, where only very coarse zonal prices are used, and most of the granular spatial differentiation is lost. **Chapter 7.2.1.2** discusses in

55 If there is any indication of the actual or potential "contribution to the peaks" of each customer (e.g., in some countries each customer has to contract a certain amount of peak power, which cannot be exceeded at any time without triggering the disconnection of supply), then some fraction of the residual network costs could be allocated in proportion to this amount of contracted power, as another proxy to the inverse of price elasticity and a crude incentive for peak reduction in the absence of better signals. However, this can incentivize the use of batteries or distributed generation to reduce the charges for policy and residual network costs.

detail the EU market strategy, where the level of representation of the transmission network has been simplified in favor of an easier integration of the national markets into a single EU trading platform while maintaining national control of system operation.

7. An interesting question is whether the lack of detailed locational prices at transmission level — as is the case in the European Union, and also up to a point in the United States for demand and DERs — implies that we have to give up the locational signals at the distribution level. Certainly, averaging LMPs over a large zone sacrifices that locational value. But we have to realize that, despite any averaging of LMPs at the transmission level, we may want to signal that DERs have a high locational value in one feeder and none in another (see **Chapter 8** for a detailed treatment of locational values). Therefore, despite the inconsistency, it can make economic sense to apply detailed locational signals at the distribution level, even if it is not done in the transmission network.[56]

8. The next step is to further differentiate by location at the distribution level. Presently, active network constraints at the distribution level are infrequent, since, without adequate demand and DER response, a network constraint violation can only be fixed by load or generation curtailment. A full alternating current representation of power flow, with active and reactive components, along with both active and reactive prices, will be needed to capture the impact of network constraints, which in distribution are frequently related to voltage problems.

This chapter has focused on the development of and justification for a comprehensive system of prices and charges conveyed to price-responsive demand and DERs. However, we want demand response and DERs to contribute more to the provision of system services and the reduction of costs than what is derived from their response to the prices and charges just described. DERs can participate in daily (or longer-term) auctions for the provision of operating reserves, as well as in annual auctions of firm capacity markets. And local network problems that are now typically solved by "centralized" customary measures taken by the distribution system operator — such as investing in line reinforcements, larger transformers, or more voltage regulators — may now be addressed by aggregations of DERs committed to supply services via competitive auctions.[57] It is critical that those commitments provide the same level of certainty as network assets do, in order to ensure the same level of reliability. The next three chapters identify the inefficient barriers that may exist at the distribution and bulk power levels to the efficient participation of DERs in the provision of electricity services, and they offer suggestions for the removal of these barriers.

Granularity matters. The prices and regulated charges for electricity services vary significantly at different times and in different locations in electricity networks. Progressively improving the temporal and locational granularity of prices and charges can deliver increased social welfare; however, these benefits must be balanced against the costs, complexity, and potential equity concerns of implementation.

56 For instance, in line with the material presented in **Chapter 8** on locational value, the efficient decision about whether to deploy storage or solar PV at the residential or utility level in feeders fundamentally under the same transmission node(s) and corresponding LMP(s) does not depend on the value of the LMP(s), but on the locational differentiation of prices and charges at the distribution level.

57 In the mathematical model that optimizes the operation of the power system (detailed in **Chapter 2**), the dual variable of the active network constraint that we try to mitigate is the "system willingness to pay," i.e., the avoided system cost if the constraint is relaxed by one unit. If the optimization model that represents the power system and computes prices includes all the means for mitigating the active constraint (larger transformers, conductor upgrades, new voltage controllers, committed response of an aggregation of DERs or of demand response, etc.) then the optimization model will decide the least costly way to remove the constraint violation and the dual variable of the constraint will be the clearing price of the auction.

4.6.2 Distributional concerns

Tariff design must take into account distributional considerations such as: respect for accepted regulatory practices; avoidance of large changes in the income-effect of consumers' electricity tariffs (i.e., changes in the relative importance of consumers' electricity tariffs in their overall spending), even if the structure of the tariff drastically changes; and accounting for the diverse range of socioeconomic features that characterize power system agents.

In most countries, present retail rates are identical for all network users connected at the same voltage level, or for all consumers within a given distribution service area, regardless of the user's location within the network.[58] While each network user pays a different total amount in the final retail bill — depending upon the user's level of energy consumption and capacity contribution — all users' rates are the same. That is, every kWh of energy consumption or production, or every kW of capacity utilization, is valued equally (sometimes with differentiation based on the user's time of consumption) in all locations. Implementation of network charges differentiated according to location within a distribution system, time of network utilization, or according to other service characteristics that have been ignored in typical rate design, may find opposition in the face of existing widespread policies of socialization of electricity costs or established guidelines for nondiscriminatory rate design.

In addition, and to a first approximation, low-income consumers — who typically also consume less energy — may face higher bills if our first recommendation (reducing the volumetric component of the tariff in favor of an annual lump sum) were implemented, unless the second part of our first recommendation (making the magnitude of each customer's fixed charge dependent upon some proxy metric of lack of price elasticity or some measure of wealth) compensates for the first effect. [59]

Note, however, that distributional concerns and low-income consumer protection can be achieved without giving up the implementation of a more efficient and comprehensive system of prices and charges for broad swaths of network users. For example, lump-sum, means-tested rebates can equalize the average charges experienced in different parts of the network and can reduce total bills for low-income consumers without distorting economic incentives in the prices and charges to which users are exposed. Likewise, the hedging of month-to-month bill volatility can offer individuals with fixed and limited incomes assurances that their bills will not fluctuate wildly. Efficient prices and charges are key to lowering overall costs of electricity services, which benefits customers in general.

58 For the sake of simplicity, in this discussion we ignore the fact that in power systems with liberalized retailing this statement is only valid for network charges, not necessarily for retail prices, as each customer may contract the energy supply from a different retailer.

59 Indeed, this may not always be the case, as volumetric consumption is imperfectly correlated with wealth (Wood et al. 2016).

References

Abdelmotteleb, I., T. Gómez, and J. Reneses. 2016. "Distribution Network Cost Allocation Using a Locational and Temporal Cost Reflective Methodology." *Power Systems Computation Conference (PSCC), 2016*, 1–7. Power Systems Computation Conference. ieeexplore.ieee.org/abstract/document/7540878/.

AEMC, 2016. "Consultation Paper — Local Generation Network Credits." Australian Energy Market Commission. www.aemc.gov.au/getattachment/d163ba4e-0071-49d8-b340-57e0813bcc32/City-of-Sydney.aspx.

AEMO. 2016. "Distribution Loss Factors for the 2016/2017 Financial Year." Australian Energy Market Operator. www.aemo.com.au/Electricity/National-Electricity-Market-NEM/Security-and-reliability/Loss-factor-and-regional-boundaries/-/media/BEAA8BCF75814061B12C1427D21F41A1.ashx.

Ausgrid. 2011. "Distribution Loss Factor Calculation Methodology Paper." www.aer.gov.au/system/files/Ausgrid%20-%20Distribution%20Loss%20Factors%20-%20Methodology%202011-12.doc.

Batlle, C. 2011. "A Method for Allocating Renewable Energy Source Subsidies among Final Energy Consumers." *Energy Policy* 39 (5): 2586–95.

Batlle, C., J.P. Chaves-Avila, P. Rodilla, P. Mastropietro, T. Gomez, and I.J. Perez-Arriaga. 2016. "Regulated and policy charges and electricity bills for a distributed future: efficient price signals for increasingly elastic end-users." MIT Energy Initiative Utility of the Future Working Paper MITEI-WP-2016-09. September 2016. energy.mit.edu/uof.

Biggar, D.R. and M. Hesamzadeh. 2014. *The Economics of Electricity Markets*. Chichester, West Sussex, United Kingdom: Wiley.

Bonbright, J. 1961. *Principles of Public Utility Rates*. New York, NY: Columbia University Press.

Caramanis, M., E. Ntakou, W.W. Hogan, A. Chakrabortty, and J. Schoene. 2016. "Co-Optimization of Power and Reserves in Dynamic T&D Power Markets with Nondispatchable Renewable Generation and Distributed Energy Resources." *Proceedings of the IEEE* 104 (4): 807–36. doi:10.1109/JPROC.2016.2520758.

Costello, K.W., and R.C. Hemphill. 2014. "Electric Utilities' 'Death Spiral': Hyperbole or Reality?" *The Electricity Journal* 27 (10): 7–26. doi:10.1016/j.tej.2014.09.011.

CNMC. 2015. "Report in the Provisional Settlement 1/2015 of the Power Sector: Analysis of Results and Monthly Monitoring of the Annual Estimate of Revenues and Costs of the Power Sector" (in Spanish). Comisión Nacional de los Mercados y la Competencia.

CSIRO. 2013. "Change and Choice: The Future Grid Forum's Analysis of Australia's Potential Electricity Pathways to 2050." Commonwealth Scientific and Industrial Research Organisation. publications.csiro.au/rpr/download?pid=csiro:EP1312486&dsid=DS13.

Ecofys. 2014. "Design Features of Support Schemes for Renewable Electricity." DESNL13116. www.ecofys.com/files/files/ec-fraunhofer-isi-ecofys-2014-design-features-of-support-schemes.pdf.

EIA. 2014. "Price Elasticities for Energy Use in Buildings of the United States." US Energy Information Administration.

EPRI. 2016. "Residential Off-Grid Solar + Storage Systems: A Case Study Comparison of On-Grid and Off-Grid Power for Residential Consumers." Electric Power Research Institute. www.epri.com/abstracts/Pages/ProductAbstract.aspx?ProductId=000000003002009150.

European Commission. 2015. "Best Practices on Renewable Energy Self-Consumption." ec.europa.eu/energy/sites/ener/files/documents/1_EN_autre_document_travail_service_part1_v6.pdf.

Eurostat. 2016. "Electricity Production, Consumption and Market Overview." ec.europa.eu/eurostat/statistics-explained/index.php/Electricity_production,_consumption_and_market_overview.

Felder, F.A. and R. Athawale. 2014. "The Life and Death of the Utility Death Spiral." *The Electricity Journal* 27 (6): 9–16.

FERC. 2016. "PJM at a Glance." Federal Energy Regulatory Commission. www.ferc.gov/market-oversight/mkt-electric/pjm/elec-pjm-glance.pdf.

Fraunhofer ISE. 2015. "Recent Facts about Photovoltaics in Germany." www.ise.fraunhofer.de/en/publications/veroeffentlichungen-pdf-dateien-en/studien-und-konzeptpapiere/recent-facts-about-photovoltaics-in-germany.pdf

Graffy, E. and S. Kihm. 2014. "Does Disruptive Competition Mean a Death Spiral for Electric Utilities?" *Energy Law Journal* 35: 1.

Herter, K. and Y. Okuneva. 2013. "SMUD's 2012 Residential Precooling Study — Load Impact Evaluation." Prepared for the Sacramento Municipal Utility District.

Hummon, M., P. Denholm, J. Jorgenson, D. Palchak, B. Kirby, and O. Ma. 2013. "Fundamental Drivers of the Cost and Price of Operating Reserves." US National Renewable Energy Laboratory. NREL/TP-6A20-58491. www.nrel.gov/docs/fy13osti/58491.pdf.

Institute for Electric Innovation. 2016. "Electric Company Smart Meter Deployments: Foundation for a Smart Grid." The Edison Foundation. www.edisonfoundation.net/iei/publications/Documents/Final%20Electric%20Company%20Smart%20Meter%20Deployments-%20Foundation%20for%20A%20Smart%20Energy%20Grid.pdf.

ISO New England. 2015. "The Importance of a Performance-Based Capacity Market to Ensure Reliability as the Grid Adapts to a Renewable Energy Future." Revised discussion paper. October 2015.

Joskow, P. and C. Wolfram. 2012. "Dynamic Pricing of Electricity." *American Economic Review: Papers and Proceedings*, 102 (3): 381-385.

JRC Smart Electricity Systems and Interoperability. 2016. "Smart Metering Deployment in the European Union." October 31, 2016. ses.jrc.ec.europa.eu/smart-metering-deployment-european-union.

Kroposki, B. 2016. "Can Solar Save the Grid?" *IEEE Spectrum* 53 (11): 42–47. doi:10.1109/MSPEC.2016.7607026.

Lewis, W.A. 1941. "The Two-Part Tariff." *Economica, New Series* 8 (31): 249–70.

Low, S.H. 2014. "Convex Relaxation of Optimal Power Flow — Part I: Formulations and Equivalence." *IEEE Transactions on Control of Network Systems* 1 (1): 15–27. doi:10.1109/TCNS.2014.2323634.

Mathieu, J.L., M. Dyson, and D.S. Callaway. 2012. "Using Residential Electric Loads for Fast Demand Response: The Potential Resource and Revenues, the Costs, and Policy Recommendations." *Proceedings of the ACEEE Summer Study on Buildings*, 2012: 189–203.

Meeus, L., I. Azevedo, X. He, L. Olmos, N.-H. von der Fehr, J.-M. Glachant. 2013. "Cost Benefit Analysis in the Context of the Energy Infrastructure Package." Florence School of Regulation. Florence, Italy: European University Institute. doi:10.2870/60065.

MIT. 2015. "The Future of Solar Energy." Cambridge, MA: Massachusetts Institute of Technology. mitei.mit.edu/system/files/MIT%20Future%20of%20Solar%20Energy%20Study_compressed.pdf.

Moghe, R., D. Tholomier, D. Divan, J. Schatz, and D. Lewis. 2016. "Grid Edge Control: A New Approach for Volt-Var Optimization." *2016 IEEE/PES Transmission and Distribution Conference and Exposition*, 1–5. doi:10.1109/TDC.2016.7519895.

Molzahn, D.K., J.T. Holzer, B.C. Lesieutre, and C.L. DeMarco. 2013. "Implementation of a Large-Scale Optimal Power Flow Solver Based on Semidefinite Programming." *IEEE Transactions on Power Systems* 28 (4): 3987–98. doi:10.1109/TPWRS.2013.2258044.

NARUC. 2016." Distributed Energy Resources Rate Design and Compensation." National Association of Regulatory Utility Commissioners. http://www.naruc.org/rate-design/.

Neme, C. and R. Sedano. 2012. "US Experience with Efficiency as a Transmission and Distribution System Resource." Montpelier, VT: Regulatory Assistance Project. February 2012.

NIST. 2014. "Interagency Report on Guidelines for Smart Grid Cybersecurity." NISTIR 7628: 1-3. National Institute of Standards and Technology. September 2014.

Ntakou, E., and M. Caramanis. 2014. "Price Discovery in Dynamic Power Markets with Low-Voltage Distribution-Network Participants." 2014 *IEEE PES T&D Conference and Exposition*, April 2014. IEEE, 1–5. doi:10.1109/TDC.2014.6863212.

NYISO. 2016. "2016 Load and Capacity Data 'Gold Book.'" Rensselaer, NY: New York Independent System Operator. April 2016.

Ofgem. 2014. "Bidding Zones Literature Review." London, United Kingdom: Office of Gas and Electricity Markets. www.ofgem.gov.uk/sites/default/files/docs/2014/10/fta_bidding_zone_configuration_literature_review_1.pdf.

Ofgem. 2015. "Wholesale Energy Markets in 2015." London, United Kingdom: Office of Gas and Electricity Markets. September 9, 2015.

OMIE. 2016. *Price Report 2015*. Operador del Mercado Ibérico de Energia (Spanish market operator).

Pérez-Arriaga, I., and Y. Smeers, 2003. "Guidelines on Tariff Setting." In *Transport Pricing of Electricity Networks*, Francois Leveque, ed. Dordrecht, Netherlands, and Norwell, MA: Kluwer Academic Publishers. ISBN: 1402074557. August 2003.

Pérez-Arriaga, I., S. Ruester, S. Schwenen, C. Batlle, and J.-M. Glachant. 2013. *From Distribution Networks to Smart Distribution Systems: Rethinking the Regulation of European Electricity DSOs.* THINK Project, Topic 12. Florence, Italy: European University Institute.

Pérez-Arriaga, I., and A. Bharatkumar. 2014. "A Framework for Redesigning Distribution Network Use of System Charges Under High Penetration of Distributed Energy Resources: New Principles for New Problems." Cambridge, MA: MIT Center for Environmental and Energy Policy Research, Massachusetts Institute of Technology. ceepr.mit.edu/publications/reprints/614.

Rocky Mountain Institute, Homer Energy, and Cohnreznick Think Energy. 2014. "The Economics of Grid Defection." homerenergy.com/pdf/RMI_Grid_Defection_Report.pdf.

Rubio-Oderiz, F.J. 1999. "Methodology of Transmission Network Cost Allocation under a Competitive Regulatory Framework." Doctoral dissertation (in Spanish). Madrid, Spain: Institute for Research in Technology, Universidad Pontificia Comillas de Madrid.

Rubio-Oderiz, F.J., and I.J. Perez-Arriaga. 2000. "Marginal Pricing of Transmission Services: A Comparative Analysis of Network Cost Allocation Methods." *IEEE Transactions on Power Systems* 15 (1): 448–54.

Schneider, I., and C.R. Sunstein. 2016. "Behavioral Considerations for Effective Time-Varying Electricity Prices." SSRN Scholarly Paper ID 2841049. Rochester, NY: Social Science Research Network. papers.ssrn.com/abstract=2841049.

Schweppe, F.C. 1978. "Power Systems '2000': Hierarchical Control Strategies." *IEEE Spectrum* 15 (7): 42–47. doi:10.1109/MSPEC.1978.6367813.

Schweppe, F.C., M.C. Caramanis, R.D. Tabors, and R.E. Bohn. 1988. *Spot Pricing of Electricity*. Dordrecht, Netherlands, and Norwell, MA: Kluwer Academic Publishers.

Warner-Freeman, J. 2016. "PJM Markets Report." PJM. www.pjm.com/~/media/committees-groups/committees/mc/20161024-webinar/20161024-item-08a-markets-report.ashx.

Wood, L., R. Hemphill, J. Howat, R. Cavanagh, S. Borenstein, J. Deason, and L. Schwartz. 2016. "Recovery of Utility Fixed Costs: Utility, Consumer, Environmental and Economist Perspectives." Berkeley, CA: Lawrence Berkeley National Laboratory. www.publications.lbl.gov/islandora/object/ir%3A1005742.

PART 2: A FRAMEWORK FOR AN EFFICIENT AND EVOLVING POWER SYSTEM

05

The Future of the Regulated Network Utility Business Model

5.1 Introduction: Challenges in Electricity Distribution Regulation, Old and New

Electricity distribution is considered a natural monopoly activity because it is more affordable to serve a given territory with just one network utility. This in turn necessitates economic regulation. In particular, regulators must establish the allowed revenues or remuneration of distribution utilities,[1] which gives rise to several challenges.[2] First, regulators face varying degrees of incomplete and asymmetric information since they cannot directly observe the utility's costs or service quality opportunities or the level of managerial effort expended (Joskow 2014; Laffont and Tirole 1993).[3] This information asymmetry creates an opportunity for strategic behavior, as firms can increase profits if they convince regulators that they face higher costs than they really do, thus securing greater remuneration (Jamasb, Nillesen, and Pollitt 2003, 2004; Cossent and Gómez 2013). In addition, regulators must assess the prudence and efficiency of long-lived, capital-intensive utility investments, which forces regulators to manage

1 In this chapter, we begin from the current situation and assume that regulated distribution utilities are responsible for network planning, construction, maintenance, and operation. Going forward, a thorough reconsideration of electricity industry structure and assignment of core responsibilities for market platform operation, system operation and planning, and network provision is needed. This may result in different structural arrangements, and these considerations are discussed thoroughly in **Chapter 6**. If responsibility for network planning and operation, as well as network construction and maintenance, are ultimately divided among multiple entities (e.g., some form of independent distribution system operator), that may entail changes to the proposals herein, although in general, the regulatory measures recommended in this chapter apply equally well to any profit-motivated firm responsible for optimizing the design, operation, and maintenance of distribution systems.

2 The authors note that establishing the allowed revenues of the utility, known variously as the rate-making or remuneration process, is only one of the regulator's core responsibilities. A number of important regulatory challenges fall outside the scope of this paper, which focuses only on improving the core rate-making or remuneration process. For example, regulators will likely need to establish a suite of performance incentives as well as the core incentives established by the process herein (Cossent 2013). In addition, after the total volume of allowed revenues is established, rates or tariffs for network users must be designed to ensure allowed cost recovery, establish efficient signals for network use, and address other regulatory concerns, such as equity. For more on this topic, see **Chapter 4** (also Pérez-Arriaga and Bharatkumar 2014).

3 The level of information asymmetry varies across jurisdictions. Regulators try to mitigate the information asymmetry by requesting ample technical and economic information from the companies; this is the case in the United Kingdom, Spain, and other jurisdictions.

uncertainty about future technological change and demand for network services (Cossent 2013; Ofgem 2013a). To reduce costs for consumers, regulators are simultaneously concerned with two forms of efficiency: productive efficiency, which has to do with firms finding the least costly way to produce a given output or service (in both the short and long term),[4] and allocative efficiency, which is concerned with ensuring consumers pay prices that reflect incurred production costs, minimizing the economic "rents" extracted by producers. In practice, regulators must carefully balance inherent trade-offs in the design of regulatory mechanisms, incentivizing productive efficiency by rewarding utilities for pursuing cost savings on the one hand and maximizing allocative efficiency by minimizing rents collected from ratepayers on the other hand (Joskow 2014).[5] Finally, regulators must achieve this balance of efficiency incentives all while preserving incentives for quality of service and loss reduction (Gómez 2013; Cossent 2013).

On top of these persistent challenges, the proliferation of distributed energy resources (DERs) and information and communications technologies (ICTs) is now actively transforming the delivery of electricity services and the use and management of distribution systems in many jurisdictions. Distributed generation and storage introduce bidirectional power flows and, at significant penetration levels, entail profound changes to the real-time operation and the needs for investment in distribution systems (Cossent, Gómez, and Frías 2009; Cossent et al. 2011; Denholm et al. 2013; Olmos et al. 2013; Pudjianto et al. 2014; Strbac et al. 2012; Vergara et al. 2014). Widespread electric vehicle adoption could likewise necessitate new network investments and may enable new "vehicle to grid" services (Fernández et al. 2011; Gómez et al. 2011; Momber, Gómez, and Söder 2013). Advanced metering, dynamic market-based prices, time-varying rates, and energy management systems have the potential to make electricity loads more responsive to economic and operational signals than ever before (Conchado and Linares 2012; Hurley, Peterson, and Whited 2013; Schisler, Sick, and Brief 2008). Efficient price signals, new information and control systems, and/or novel market actors are necessary to manage and coordinate each of these DERs and their associated services (see also **Chapters 4** and **6**).

These emerging technologies constitute a set of important new users of distribution systems, potential new competitors for the delivery of electricity services to end users (Kind 2013), and possible suppliers of services to distribution companies seeking to harness DER capabilities to avoid network investments or improve system performance (Poudineh and Jamasb 2014; Trebolle et al. 2010).

Distribution utilities may need to make substantial investments to accommodate increased penetration of DERs. In many jurisdictions, these new investments will coincide with significant expenditures necessary to install and manage advanced meters and modernize aging distribution systems to take advantage of new ICT capabilities and active system management techniques (Cossent et al. 2011; Eurelectric 2013; MA DPU 2014; NYDPS 2015a). At the same time, the pace of change and impact of distributed energy resources on distribution systems is highly uncertain. For example, while solar photovoltaic (PV) systems generated less than 1 percent of Germany's electricity in 2008, solar supplied 7.5 percent of German electricity consumption and roughly 20 percent of installed capacity in 2015 (Wirth 2016). While rapid solar adoption in Germany was principally

4 Long-term productive efficiency (e.g., improvements in the productive frontier for a given firm) is sometimes also referred to as dynamic efficiency, in contrast to static efficiency (choosing the least-cost production given currently available options). Here we consider both static and dynamic productive efficiency, and we encourage the design of incentives for both important forms of cost-saving effort: e.g., moving firms to the productive frontier via the exercise of managerial effort and moving the productive possibility frontier forward over time via innovation.

5 Here we focus on the implications of establishing the total allowed remuneration of distribution utilities on allocative efficiency. Allocative efficiency also depends centrally on how these allowed revenues are recovered from network users via regulated charges or tariffs. Establishing a system of prices and regulated charges for utility services that reflects the marginal cost of those services is thus also critical to maximizing allocative efficiency. This topic is the subject of **Chapter 4**.

driven by policy support, it is indicative of the rate at which DER penetration can increase, whether driven by policy, improved economics, or a combination thereof.

The rapid adoption or significant penetration of DERs exacerbates fundamental regulatory challenges and strains conventional approaches to the remuneration of distribution utilities. Unless proactive reforms are made to update the regulation of distribution utilities, outdated network regulation may become a key barrier to the efficient evolution of power systems. What follows is thus a discussion of the most important emerging challenges facing the regulation of distribution utilities, identifying the key areas where regulatory reforms and new tools are needed for a more efficient power system.[6]

The rapid adoption or significant penetration of DERs exacerbates fundamental regulatory challenges and strains conventional approaches to the remuneration of distribution utilities. Unless proactive reforms are made to update the regulation of distribution utilities, outdated network regulation may become a key barrier to the efficient evolution of power systems.

5.1.1 New uses of distribution networks and drivers of network costs

First, new demands on networks, new alternatives to traditional network investments, and increased competition for end-user service delivery will pressure distribution utilities to focus on delivering improved outputs at a competitive cost.

Under conventional cost of service regulation,[7] it will likely be difficult for utilities to respond to new demands while taking full advantage of the capabilities provided by DERs and smart grid technologies. Cost of service regulation tends to focus on reviewing the prudence of inputs and identifying departures from established practices, which makes it challenging for utilities to respond to evolving demands for outcomes or focus on delivering improved performance, such as enhanced resiliency or access for distributed resources to sell services to system operators or wholesale markets (Malkin and Centolella 2013). In addition, traditional cost of service regulation generally requires utilities to meet minimum performance levels but provides little incentive or reward for utilities that deliver higher-quality service or new outcomes and services.[8]

At the same time, cost of service regulation provides weak incentives for utilities to take full advantage of the cost-saving opportunities made available by DERs, smart grid technologies, and active system management techniques. Utilities only profit from any realized savings until the next "rate case," when regulators reset rates to align with the cost of providing service. Utilities are thus encouraged to focus primarily on short-term cost savings, sacrificing the opportunities that could be unlocked if utilities were incentivized to invest with a longer-term view. In addition, this approach requires regulatory review of expenditures associated with thousands of individual distribution system assets, which has always posed an expensive challenge for regulatory commissions with limited staff and resources (Gómez 2013). The changing nature of cost drivers and the emergence of novel cost-saving opportunities will further aggravate this challenge, making it difficult for regulators to identify and disallow all but the most obviously imprudent or wasteful investments, further weakening incentives for firms to manage productive efficiency.

6 We do not try to present here an exhaustive inventory of the regulatory challenges facing the efficient integration of DERs into distribution systems operation and planning (nor is the list of solutions proposed later in this chapter exhaustive). Other substantial regulatory challenges exist that are not discussed here, such as the reform of existing technical codes and procedures for connecting DERs to the grid, the lack of distribution-level markets for DERs to participate in distribution planning and operation, and others.

7 Cost of service regulation focuses primarily on improving allocative efficiency by ensuring utilities earn revenues that reflect their costs of service. This is in contrast to the strategy of an unregulated monopoly, which is to maximize profits by increasing prices above the cost of production and extracting economic "rents" or excess profits from consumers. Under cost of service regulation, regulators focus on reviewing and identifying a firm's cost of production and then establishing allowed revenues that match this cost of production, including a reasonable return on capital prudently invested. This form of regulation is thus sometimes referred to as "rate of return" regulation.

8 Cost of service regulation can in principle be complemented with performance-based incentives for key quality-of-service metrics (Lowry and Woolf 2016), but in many jurisdictions this is not the case today.

Finally, cost of service regulation is typically backward-looking in nature, focused on reviewing the prudency of incurred investments after the fact. For utilities in an actively evolving marketplace, this introduces substantial regulatory risk that can impede utility efforts to innovate and take advantage of new technologies and capabilities. In reviewing the prudence of utility investments, regulators typically rely on the incremental development of established best practices with the implicit assumption that the past is an appropriate guide for the future. Traditional regulation thus frequently requires utilities to justify novel investments and departures from established practices by proving that such changes will result in a net reduction in utility costs (Malkin and Centolella 2013). If a utility adopts a novel technology that fails to perform as expected, regulators may disallow cost recovery. As a result, utilities are often slow to adopt innovative technologies and practices and may instead go through a protracted cycle of internal testing and performance validation, small-scale pilot projects, data collection and assessment, and presenting results to regulators before finally, after many years, adopting improved technologies or practices system-wide. Cost of service regulation can thus present a major barrier to the evolution of distribution utilities in light of both changing customer needs and new DERs and smart grid technology capabilities.

Regulatory frameworks that employ a forward-looking, multi-year revenue trajectory or "revenue cap" (such as RPI-X and similar frameworks) are also challenged by the evolving nature of the electricity marketplace. The emergence of new cost drivers and changing customer needs make it increasingly difficult to establish an effective forward-looking, or *ex ante*, revenue (or price) cap. Regulators often employ statistical frontier benchmarking and yardstick approaches to assist them in establishing *ex ante* estimates of efficient network costs (Jamasb and Pollitt 2003; ACC 2012a, 2012b). Yet as network uses, cost drivers, and utility best practices rapidly evolve, benchmarking based on past utility performance or cost will no longer provide an accurate estimate of the forward-looking efficient frontier. At the same time, different patterns of DER adoption in different regions can introduce much more heterogeneity between distribution network costs, further challenging statistical benchmarking approaches (Cossent 2013). For example, the availability of solar, wind, biomass/biogas, and combined heat and power resources differs substantially from location to location and is likely to lead to the divergent evolution of distribution networks in different regions. Without improved, forward-looking tools to assist in accurately estimating efficient network costs, regulators may set *ex ante* revenues that are poorly aligned with realized costs, leading either to excess utility profits and reduced allocative efficiency (if revenues are too generous) or to increased risk that firms will not be able to finance necessary investments adequately (if revenues are too low).

In addition, as the sector evolves, distribution utilities are likely to develop more intimate and immediate knowledge about new cost drivers and opportunities than the regulator, heightening information asymmetries and creating new opportunities for strategic behavior. Tools to overcome the regulator's information disadvantages will thus become even more critical for effective incentive regulation.

In summary, regulators must be equipped with forward-looking tools to identify the impacts of new DER-related network uses on distribution costs and overcome information asymmetries. In addition, regulators need remuneration mechanisms that both incentivize utilities to accommodate cost-effective DERs and to take advantage of new DERs or smart grid opportunities to improve cost and performance outputs. As always, regulators must also ensure firms can raise necessary debt and equity to finance needed investments and, to the extent feasible, limit firms' rents.

5.1.2 Increased uncertainty about the evolution of network needs, cost drivers, and opportunities

Second, changes in the delivery of electricity services and the growth of DERs will increase uncertainty about the evolution of distribution cost drivers and network uses. Since network uses and available technologies may evolve quite rapidly, network costs may deviate substantially from estimates, leading to forecast and benchmark errors. Costs may rise or fall unexpectedly due to new network uses (e.g., the rapid penetration of newly subsidized or newly cost-competitive distributed generation), an example of forecast error. Alternatively, the regulator may fail to anticipate the emergence of new cost-saving technologies or practices that shift the efficient frontier, leading to benchmark error.

Cost of service approaches can address this heightened uncertainty through more frequent *ex post* reviews or rate cases as cost drivers and network uses evolve. However, managing uncertainty in this manner comes at a cost: an even greater reduction in incentives for cost savings. Assured of cost recovery in this manner, utilities will be unlikely to pursue the cost-saving opportunities presented by new DERs and smart grid capabilities.

Heightened uncertainty presents a more fundamental challenge to efforts to establish a forward-looking, multi-year revenue trajectory. More frequent *ex post* reviews and adjustments to remuneration levels can address benchmark and forecast errors, but once again at a cost. Frequent *ex post* "reopeners" of the regulatory contract can create significant regulatory uncertainty and thus may raise the cost of capital for distribution utilities as well as undermine efficiency incentives.

Regulators therefore need new tools to manage uncertainty and ensure adequate cost recovery and firm participation while preserving regulatory certainty and incentives for cost reduction and productive efficiency.

5.1.3 Heightened trade-offs between capital and operational expenditures

Third, the emergence of DERs and ICT-enabled smart grid capabilities will heighten trade-offs between capital expenditures (CAPEX), such as investments in new distribution lines or substations, and operational expenditures (OPEX), such as contracts with DERs to avoid or defer network investments.[9] Traditional regulatory approaches to capitalizing expenditures can skew incentives and result in utilities favoring capital rather than operational expenditures, distorting these important trade-offs.

For example, distribution utilities can achieve important cost savings by adopting an active system management approach, especially as DER penetration increases (Cossent et al. 2009, 2011; Eurelectric 2013; Olmos et al. 2009; Poudineh and Jamasb 2014; Pudjianto et al. 2013; Strbac et al. 2010; Trebolle et al. 2010). Setting up ICT and advanced grid management infrastructure that allows distribution utilities to manage distribution network configurations more actively and make use of DERs for daily grid operations will entail substantial upfront CAPEX. However, such investments would increasingly enable distribution utilities to contract with or procure system operation services from DER owners or aggregators, including deferral of network upgrades, network constraint management loss reduction, or reliability improvement (Meeus et al. 2010; Meeus and Saguan 2011; Poudineh and Jamasb 2014; Trebolle et al. 2010). These contractual arrangements or new markets for distribution system services constitute a new category of operational expenditures that would increase utility OPEX while reducing CAPEX. Alternatively, CAPEX related to new smart grid capabilities can improve the workforce and reduce truck rolls, leading to OPEX savings. Utilities may also face new decisions about whether to contract for information and communication and analytical services or invest in their own capabilities. In short, the most efficient trade-off between CAPEX and OPEX is likely to change significantly over time. Conventional regulatory approaches will need to be

9 These operational expenditures are sometimes referred to as "non-wires" alternatives, in contrast to traditional investments in distribution "wires" and other network assets.

updated to fully exploit new opportunities to effectively balance increasingly important trade-offs between these two expenditure categories.

Under traditional cost of service regulation, utilities earn a regulated return only on capital investments. Allowed returns are calculated based on the utility's "rate base" or regulated asset base, which includes the cumulative, non-depreciated share of capitalized expenditures. Under cost of service regulation, utilities can thus be discouraged from reducing CAPEX, as this may impact their rate base and allowed returns. At the same time, the intrinsically poor incentives for cost saving under cost of service approaches make it unlikely that firms will fully exploit the most efficient trade-offs between capital and operational expenditures.

While multi-year revenue trajectories or revenue caps will reward firms for efficiently reducing total costs, this form of regulation can also distort incentives between savings achieved via reductions in CAPEX versus OPEX. While incentive regulation will reward the utility equally for saving a dollar of CAPEX or a dollar of OPEX, if only CAPEX is capitalized into the utility's regulated asset base, then that dollar in reduced CAPEX will also involve a reduction in the regulated asset base and thus a reduction in the allowed return on equity and a corresponding decline in net profit for shareholders. This decline in net profit will offset some portion of the efficiency-related income, distorting trade-offs between OPEX and CAPEX and potentially encouraging over-investment (Ofgem 2009, 2013b).

Regardless of which regulatory approach is used, regulators therefore need mechanisms to equalize incentives for CAPEX and OPEX savings and to ensure utilities fully exploit these opportunities.

5.1.4 Need for increased levels of innovation

To keep pace with a rapidly evolving landscape, electricity distribution utilities will also need to focus more on harnessing long-term innovation—by participating in applied research and development efforts, investing in demonstration projects, engaging in the technological learning that emerges from those projects, and disseminating knowledge and best practices between network utilities. Uncertainty about how networks will evolve implies that the technological solutions that will lead to the greatest levels of productive efficiency in the medium to long term (i.e., in periods that extend beyond the regulatory period) are also uncertain. The technologies and systems that will be most efficient for facilitating active network management in distribution networks with high penetrations of DERs are simply not known today. Therefore, there is a need for greater investment in demonstration projects and accelerated knowledge-sharing or "spillovers" among utilities. This will involve undertaking experimental projects where the potential cost savings are inherently uncertain, may only be realized in the medium to long term, if at all, and may not fully accrue to the utility incurring the costs due to spillover. However, despite the need for increased levels of long-term innovation, spending on research, development, and demonstration (RD&D) by network utilities has been declining (Meeus and Saguan 2011). Today's regulatory frameworks, including both cost of service and incentive regulation, do not adequately incentivize these types of risky projects and the technological learning that emerges from them.

Under cost of service regulation, utilities are only incentivized to engage in innovation to the degree that short-term cost savings can be retained by the utility until the next regulatory review or rate case (Bailey 1974; Malkin and Centollela 2013). Since this period of "regulatory lag" generally lasts for up to a few years at most, utilities are only rewarded for very low-risk measures that can generate savings quickly—hardly the kind of long-term innovation required in an evolving electricity landscape. In some jurisdictions, regulators allow longer-term RD&D costs to be capitalized into the utility's rate base, enabling companies to earn a rate of return on these costs, which provides an additional incentive to engage in innovation. However, in practice regulators do not blindly accept all costs (Joskow 1989), and many innovative projects are unlikely to be approved during the regulatory process due to their inherent riskiness. Indeed, the majority of regulatory authorities in cost of service jurisdictions are risk-averse, contributing to low levels of RD&D expenditures among electric utilities in those jurisdictions (NSF 2011). Nonetheless, in the United States and elsewhere, a small number of jurisdictions operating under cost of service regulation are deviating from this trend by encouraging network utilities to propose innovative projects that will be included in their rate bases (CPUC 2015; NYDPS 2015b).

When RPI-X was first proposed (Beesley and Littlechild 1983), the assumption was that the multi-year revenue trajectory established under this form of regulation would prove superior to cost of service regulation in promoting both short- and long-term productive efficiency (Clemenz 1991; Armstrong et al. 1994; Dogan 2001; Littlechild 2006). Yet, while the mechanism is effective at rewarding short-run efficiency improvements, the actual results of incentive regulation on promoting long-term innovation are more ambiguous (Kahn et al. 1999). Under a revenue cap, the company is exposed to more risk than it is with cost of service regulation: The company bears the cost of a failed RD&D investment. Therefore, the company is more inclined to engage in "process innovations" that are very likely to lead to cost savings within the regulatory period rather than higher-risk but potentially higher-reward innovations that may take longer to bear fruit. Moreover, under this form of regulation, companies benefit if they outperform the revenue trajectory set by the regulator. Companies are thus incentivized to minimize costs in the short term, which may incentivize reductions in RD&D expenditures as well. The result is that the short-term productive efficiency that can be stimulated by incentive regulation may work against the type of long-term productive efficiency improvements that can be achieved by RD&D investments (Bauknecht 2011).

In summary, additional mechanisms must be implemented within the regulatory framework to adequately incentivize network utilities to invest in RD&D projects that are inherently risky and uncertain but that are necessary for achieving dynamic efficiency in electric power systems.

5.1.5 Challenges in cybersecurity, privacy regulation, and standards

Finally, the proliferation of DERs and ICTs in electricity networks will continue to increase the vulnerability of these systems to cyber attacks. The grid is a cyber and physical system, and future threats may involve both dimensions. Strong cybersecurity regulations, standards, and best practices are needed for all levels of the system—including the bulk power system (central generation and transmission), distribution systems, DERs, smart meters, and electrical devices with Internet connectivity in industrial, commercial, and residential buildings.

The grid is a cyber and physical system, and future threats may involve both dimensions. Strong cybersecurity regulations, standards, and best practices are needed for all levels of the system — including the bulk power system (central generation and transmission), distribution systems, DERs, smart meters, and electrical devices with Internet connectivity in industrial, commercial, and residential buildings.

Cybersecurity investments suffer from a type of market failure in which the benefit of secure operations is shared among electricity providers and network users, but each actor's private incentive to invest in adequate cybersecurity preparedness is too low to justify the cost. Network utilities, many of which face slow or negative electricity demand growth and relatively uncertain financial futures, tend to make investments that increase financial returns rather than spend money on cybersecurity, which has limited direct financial benefit. Network utilities will therefore require additional regulatory incentives to invest in adequate levels of cyber protection.

Current cybersecurity regulations vary across transmission and distribution systems, across regions, and across utility-owned and independently-owned DERs. In the United States, the North American Electric Reliability Corporation (NERC) has developed cybersecurity regulations at the bulk power and transmission levels (NERC 2016). In Europe, the first European legislation on cybersecurity, the Network and Information Security (NIS) Directive, entered into force in July 2016 (European Union 2016a). Regulations at the distribution-system level (including DERs and customer connection points) are lacking, therefore a baseline set of cybersecurity standards for all distribution networks is needed. Continuous improvements in best practices will also be required to meet evolving cyber threats. Achieving good cybersecurity in the future demands good governance as well as regulation because rule-making is not sufficiently nimble to keep up with evolving threats. Instilling good governance is difficult. Industry organizations such as the Institute of Nuclear Power Operations have been reasonably successful for the nuclear industry, but there is currently no equivalent for the cyber and physical security of electricity networks.

In addition to strengthening cybersecurity standards and regulations, privacy regulations also may need to be updated to address issues raised by the influx of DERs and Internet-connected devices. As network utilities gather more detailed information about network users and as the number of Internet-connected devices in distribution networks rises, privacy concerns may be exacerbated. Regulators may therefore need to adopt new approaches to privacy regulation. For example, EU privacy regulations approved in April 2016 give citizens control of their personal data and create a uniformly high level of data protection (European Union 2016b).

5.1.6 New regulatory tools for a more distributed power system

To address the combination of challenges described above, regulators can employ a novel combination of established best practices or state-of-the-art regulatory tools. The remainder of this chapter describes key regulatory priorities and possible mechanisms for accomplishing these objectives. At the broadest level, these regulatory priorities can be divided into two categories. In **Section 5.2**, we present a set of priorities for improving traditional approaches to establishing the core remuneration of distribution utilities, accounting for new cost drivers, new demands and uses of the distribution network, and increased uncertainty while establishing incentives for cost-saving efficiencies. **Section 5.3** then discusses additional measures necessary to incentivize utilities to improve specific outputs that are not captured by core remuneration methods, including quality of service, losses, asset health, cybersecurity preparedness, and long-term innovation.

5.2 New Methods for Establishing the Core Remuneration of Distribution Utilities

The emergence of DERs, as well as of ICT-enabled technologies and systems, has led to increased uncertainty about the evolution of network needs, cost drivers, and opportunities. As a result, the fundamental challenges of economic regulation—namely information asymmetries between the regulator and the utility—are significantly exacerbated. Moreover, the emergence of new technologies and DERs creates new opportunities for cost-saving efficiencies in utility planning and operations and heightened trade-offs between operational and capital expenditures with regard to achieving economically efficient outcomes at lowest cost. Network services that historically have been supplied primarily with capital expenditures in network assets (conductors, transformers, etc.) may now be most efficiently supplied with some combination of operational and capital expenditures, including payments to DERs for network services. To remove barriers to the cost-effective integration of DERs into power systems, the core business incentives of electricity distribution utilities must be aligned with these new conditions. That requires proactive regulatory reforms to establish efficient incentives for distribution utilities, enable business model innovation by utility companies, and effectively manage uncertainty and information asymmetry in an evolving power sector. Given the ongoing pace of change in many jurisdictions, delaying such reforms may carry significant costs to society. We recognize the significant diversity of regulatory practices in different jurisdictions and note that some of the regulatory proposals we discuss may be novel to some readers. Nevertheless, we propose that best practices, wherever they are found, should inform both the goals and practices of regulation in any jurisdiction. What follows is thus a survey of key regulatory challenges and potential solutions, drawing from best practices throughout the world.

5.2.1 Establishing efficiency incentives for distribution utilities

The electricity landscape is changing. To keep pace with the new demands and capabilities of DERs and the expanded opportunities enabled by information and communications technologies, distribution utilities must innovate, develop new network planning and management practices, and find creative ways to cost-effectively serve an increasingly diverse set of network users. To facilitate these changes to utility business models and practices, regulators need to establish appropriate incentives that reward utilities for efforts to deliver desired reliability and quality of service at lower cost to network users. In order to align the business incentives of distribution utilities with these efficient outcomes, regulators can establish earnings sharing mechanisms or efficiency incentives that reward utilities for cost-saving efforts, as well as reforms that equalize utilities' financial incentives with regard to operational and capital expenditures.

5.2.1.1 Efficiency incentives

In the absence of competition, efficient regulation of network monopolies should strive, to the extent possible, to mimic the outcomes of a competitive marketplace. Without competitors to discipline prices, for example, regulators can focus on ensuring regulated monopolies do not exercise their privileged market position to extract outsized profits from their customers. A focus on ensuring that prices charged for network use reflect the real costs of service and not the exercise of monopoly power (i.e., improving allocative efficiency) is thus central to the regulatory task. Economic regulation could also concern itself with replicating for distribution utilities the incentives that stimulate incremental innovation[10] and cost saving efficiencies in competitive markets, encouraging firms to pursue new ways to lower their bottom lines while maintaining the quality and value customers expect. These incentives are also critical to ensuring lower costs for electricity network users over time (i.e., improving productive efficiency) and are of increased importance in the context of rapidly evolving capabilities and demands on electricity networks.

Traditional cost of service regulation provides a poor substitute for the consistent pressure competitive markets supply to find cost-saving efficiencies. Regulators may disallow cost recovery for particularly imprudent or unwise investments, but in general, utilities face only modest financial incentives to reduce costs between rate

10 Throughout this chapter we make the distinction between incremental or process innovation and medium- to long-term innovation. Incremental or process innovation, which produces cost savings that can be realized in the near term (i.e., within a single regulatory period), can be incentivized under a core regulatory framework with appropriate efficiency incentives such as the framework described here. Medium- to long-term innovation, on the other hand, necessitates upfront costs for which the economic benefits may not be realized until several years after the investment (i.e., beyond the current regulatory period). Additional incentives may be needed to encourage this kind of long-term innovation, as discussed in **Section 5.3.2**.

Forward-looking, multi-year revenue trajectories with profit-sharing mechanisms can reward distribution utilities for cost-saving efficiency improvements, aligning utilities' business incentives with the continual pursuit of novel solutions.

cases. The more frequently rate cases occur, and the more closely allowed revenues match a utility's actual costs, the less incentive utility managers have to reduce cost of service over time (Lowry and Woolf 2016). At the same time, the steady returns on capital afforded by cost of service regulation can incentivize utilities to find ways to justify increased investment, rather than pursue cost-saving efficiencies (Averch and Johnson 1962).[11]

By adopting state-of-the-art regulatory mechanisms, regulators can better align utility incentives for cost-saving efficiency efforts and ensure that the benefits of improved utility performance are shared between utility shareholders and ratepayers. The principal mechanism used to reward efficiency efforts is the multi-year revenue trajectory with profit sharing. This regulatory mechanism is a key component of regulatory regimes that go by many names, including "revenue caps," "RPI-X," "multi-year rate plans," and others.[12] In each case, the basic objective is the same: Utilities should be assured that, over some definite period of time, their revenues will be to some degree decoupled from their costs, such that they will be able to retain some portion of any cost savings they accrue during that time. An essential characteristic is that such multi-year trajectories are set in advance and are forward-looking.

In the most extreme form, a pure "revenue cap" can be established, fixing a utility's revenues over a period of time, such as three to five years. This can be established by granting a utility a predefined "holiday" between rate cases (i.e., fixing allowed revenues at the level established in the previous cost of service review) or in a forward-looking manner by establishing an expected revenue requirement for an efficient utility (e.g., using a forward test-year to establish cost of service or using benchmarking and forecasting techniques to define an efficient revenue trajectory). In either case, once a utility's allowed revenues are fixed for a defined period, its profitability will depend entirely on how well it controls costs during that time: Utilities will retain any cost reductions below the cap as profit and incur any costs above the cap as loss. A pure revenue cap thus establishes a powerful incentive for cost-saving productive efficiency but also entails much greater volatility in utility earnings. In addition, if utilities succeed in reducing costs well below the revenue cap, they may retain significant profits, reducing allocative efficiency.

Profit-sharing mechanisms are therefore typically used in conjunction with multi-year revenue trajectories to share potential profits and risks between utilities and ratepayers and balance incentives for productive efficiency with concerns for allocative efficiency. Under a profit-sharing mechanism, utilities retain only a portion of any reductions in cost below the revenue trajectory,

11 Indeed, in an extreme case where allowed utility revenues precisely track costs, the only way for a utility to increase profitability is to increase investments whenever allowed returns on capital invested exceed the utility's cost of capital.

12 These regulatory regimes are sometimes classified as "incentive regulation." We avoid using that term here because all forms of regulation incentivize some behaviors, and performance-based output incentives can be incorporated into a cost of service regulatory regime to incentivize key outputs as well. Under the umbrella of "incentive regulation," we distinguish between (1) "performance-based" mechanisms, which focus on incentivizing key outputs or improvements in quality-of-service metrics (e.g., reducing the number of minutes per year of non-served energy), and (2) multi-year revenue trajectory or revenue cap mechanisms, which specifically promote production cost reductions over a multi-year period (e.g., three to five years or more). This section is focused on the latter, discussing mechanisms for balancing incentives for productive and allocative efficiency and managing the challenges involved in this task, while **Section 5.3.1** focuses on performance-based or output-based incentive mechanisms. Establishing appropriate "incentives" for desired outcomes is a central focus throughout both sections.

with the remaining share accruing to ratepayers in the form of lower rates. Likewise, if actual expenditures exceed the revenue trajectory, utilities bear only a portion of the excess cost, with rates increasing to share the remainder of the burden with ratepayers. In this manner, the profit-sharing mechanism retains utility incentives for cost reduction and improved performance (productive efficiency) but does not fully decouple allowed revenues from realized utility costs, thus improving rent extraction and allocative efficiency (if costs fall below the revenue cap) and mitigating utility exposure to uncertainty (if costs rise above the revenue cap) (Gómez 2013; Joskow 2014; Schmalensee 1989). The regulator can choose a precise profit-sharing factor (or efficiency incentive rate) ranging between 1.0 (corresponding to a pure revenue cap) and 0.0 (corresponding to a pure cost of service approach). The choice of profit-sharing factor can thus be selected to manage regulatory trade-offs between incentives for productive efficiency and cost savings on the one hand (which argues for a higher factor) and rent extraction or allocative efficiency on the other (which argues for a lower factor).

Furthermore, the regulator can improve on a single profit-sharing factor by offering a regulated utility a menu of regulatory contracts with a continuum of different sharing factors (Laffont and Tirole 1993; Crouch 2006; Cossent and Gómez 2013; Jenkins and Pérez-Arriaga 2017). A menu of contracts allows the firm to play a role in selecting the strength of cost-saving incentives. If constructed correctly, this menu will establish "incentive compatibility"—that is, the design of the menu ensures that, *ex ante*, a profit-maximizing utility will always be better off (i.e., earn the greatest profit and return on equity) when the firm accurately reveals its expectations of future costs.[13] Incentive compatibility thus eliminates incentives for firms to inflate their cost estimates while rewarding firms for revealing their true expected costs to the regulator, which helps minimize strategic behavior and information asymmetries. Furthermore, a profit-motivated firm with less opportunity to reduce costs will choose a low-powered incentive, while a firm with large efficiency opportunities will choose a high-powered incentive, thereby helping to establish an appropriate balance between productive and allocative efficiency. The theoretical foundations for an incentive-compatible menu of contracts were outlined by Nobel Prize–winning economist Jean Tirole and colleague Jean Jacques Laffont (Laffont and Tirole 1993), and the method has been successfully implemented by the UK Office of Gas and Electricity Markets (Ofgem), which regulates electricity network utilities in Great Britain (Ofgem 2004, 2009, 2010a, 2013b; Crouch 2006).[14] Cossent and Gomez (2013) and Jenkins and Pérez-Arriaga (2017) describe in detail how to establish an incentive-compatible menu of contracts for distribution utility regulation; this approach is briefly summarized in **Box 5.1**.

13 Note that once the regulatory contract is set and the revenue trajectory established, the utility will maximize profit by reacting efficiently to the *ex post* realization of demand for network services and the availability of new technologies or practices that lower costs below its initial expectations. Incentive compatibility thus ensures utilities make their best projections available *ex ante*, while a multi-year revenue trajectory with earnings sharing incentives ensures utilities are rewarded for reducing costs as much as possible *ex post*.

14 Under Ofgem's regulations, the incentive-compatible menu of contracts is known as the "information quality incentive" (IQI); it debuted for capital expenditures in the 2004 distribution price control review process (DPCR4; Ofgem 2004; Crouch 2006). The mechanism is now an integral part of Ofgem's RIIO regulatory framework (RIIO stands for "revenue set to deliver strong incentives, innovation, and outputs" [Ofgem 2010c, 2013b]) wherein the IQI is used to establish a utility's profit-sharing and total allowed expenditures.

Box 5.1: Establishing an Incentive-Compatible Menu of Contracts

A menu of contracts is an important tool for incentivizing utilities to accurately reveal their expectations of efficient future expenditures. Employing a menu of contracts in the remuneration process can thus help reduce information asymmetry between regulators and utilities and mitigate incentives for utilities to engage in strategic behavior (such as inflating expectations of future costs to increase allowed revenues or returns).

A menu of contracts works by establishing the strength of a profit-sharing factor that rewards cost-saving efficiency efforts such that the sharing factor reflects the ratio between an estimate of efficient network costs over the regulatory period submitted by the utility as compared to the regulator's own estimate of efficient expenditures. Using the method introduced in Cossent and Gómez (2013) and described further in Jenkins and Pérez-Arriaga (2017), the regulator only needs to select four discretionary regulatory parameters to create a continuous, incentive-compatible menu of contracts:

1. The weight placed on the regulator's estimate of efficient network expenditures relative to the utility's estimate, ω. This weight should depend on how reliable the regulator believes his or her estimate of future expenditures is likely to be relative to the accuracy of the firm's estimate. A higher value places more weight on the regulator's estimate, while a lower value places more weight on the firm's estimate.

2. The reference value for the profit-sharing factor (the portion of cost savings/increases to which the utility is exposed, also known as the efficiency incentive rate), SF_{ref}, which corresponds to the case in which the utility's estimate of future expenditures aligns with the regulator's estimate. A profit-sharing factor of 1.0 corresponds to a pure revenue cap contract while a value of 0.0 corresponds to a cost of service contract. The regulator can thus tune the sharing factor to establish the strength of efficiency incentives faced by utilities in order to manage trade-offs between incentives for efficiency and rent extraction. This parameter also plays an important role in managing the effects of forecast errors. Regulators should thus take into account the degree of uncertainty about future network costs when establishing this factor. In general, a higher profit-sharing factor (i.e., the firm is exposed to most of the risks and rewards of cost savings) performs better under lower levels of uncertainty, while a lower profit-sharing factor (which shares most risks and rewards with ratepayers) performs better under greater uncertainty (Schmalensee 1989; see also Ofgem 2010a, pp. 84-87, for further discussion of regulatory considerations in establishing the sharing factor or incentive rate).

3. When the method is applied to several utilities simultaneously, the value of this parameter, SF_{roc}, is set to control the spread in efficiency incentives faced by different utilities during the regulatory period. The parameter determines the change in the profit-sharing factor as the ratio between the changes in the utility's estimate and the regulator's estimate. A larger value results in a wider range of profit-sharing factors offered, while a smaller factor results in a tighter range.

4. The reference value for an additional income payment, AI_{ref}, is used to ensure the utility will always earn the greatest profit when the firm accurately reveals its expectations of future costs—that is, it ensures "incentive compatibility" of the menu of contracts. This reference value specifies the additional income (on top of the allowed cost of capital) that applies when the utility's estimate of future costs aligns exactly with the regulator's estimate. The selected value can be used to tune expected profit margins for the utility.

Using these four parameters and a set of formulas described in Cossent and Gómez (2013) and Jenkins and Pérez-Arriaga (2017), the regulator can then calculate the remaining initialization parameters necessary to construct a full menu of contracts. These formulas can also compute the appropriate allowed expenditure trajectory (the *ex ante* estimate of efficient total expenditures necessary to serve network users) and the strength of the profit-sharing incentive (the portion of realized over- or under-spending shared with the utility's shareholders).

An example menu of contracts is shown in **Table 5.1**. The first row of the table describes the ratio between the regulator's estimate of efficient expenditures and the estimate submitted by the utility. Given this ratio, the regulator sets the utility's allowed *ex ante* expenditure baseline (Row 2) as well as the profit-sharing factor (Row 3) and the additional income (Row 4) awarded to ensure incentive compatibility. The first four rows of Table 5.1 thus fully define the menu of contracts, while the remaining rows illustrate the efficiency incentive earnings (or penalties) associated with the level of actual network expenditures the utility manages to achieve *ex post*. Shaded cells in these rows correspond to cases in which the utility's *ex ante* expenditure forecast matches actual expenditures, demonstrating the incentive-compatible nature of this matrix. For any realized value of network costs (the horizontal row in the bottom half of Table 5.1), the utility will earn the greatest revenues when realized costs match its *ex ante* forecast. Efficiency incentives are also preserved, as lowering realized costs below the utility's forecast (i.e., moving up in a vertical column) will increase the utility's final revenues (and vice versa). Note that while this table shows discrete values in each column, the formulas actually generate a continuous menu of contracts for any ratio between regulator and utility expenditure estimates.

Table 5.1: Sample Incentive-compatible Menu of Profit-sharing Contracts[15]

Ratio of Firm's Efficient Expenditure Estimate to Regulator's Estimate [%]	$O_{ex\ ante}$	90	95	**100**	105	110	115	120
Ex Ante Allowed Expenditure Baseline [% of Regulator's Cost Estimate]	$X_{ex\ ante}$	96.6	98.3	**100.0**	101.7	103.4	105.1	106.8
Sharing Factor [%]	SF	80.0	75.0	**70.0**	65.0	60.0	55.0	50.0
Additional Income [% of Regulator's Expenditure Estimate]	AI	3.2	2.2	**1.0**	-0.2	-1.5	-2.9	-4.4
Ratio of Realized *Ex Post* Expenditures to Regulator's *Ex Ante* Estimate [% of *Ex Ante* Estimate]	$O_{ex\ ante}$	Profit-sharing Efficiency Incentive [% of Regulator's *Ex Ante* Expenditure Estimate]						I
85		12.5	12.1	**11.5**	10.6	9.5	8.1	6.5
90		8.5	8.4	**8.0**	7.4	6.5	5.4	4.0
95		4.5	4.6	**4.5**	4.1	3.5	2.6	1.5
100		0.5	0.9	**1.0**	0.9	0.5	-0.1	-1.0
105		-3.5	-2.9	**-2.5**	2.4	-2.5	-2.9	-3.5
110		-7.5	-6.6	**-6.0**	-5.6	5.5	-5.6	-6.0
115		-11.5	-10.4	**-9.5**	-8.9	-8.5	8.4	-8.5
120		-15.5	-14.1	**-13.0**	-12.1	-11.5	-11.1	11.0
125		-19.5	-17.9	**-16.5**	-15.4	-14.5	-13.9	-13.5

Source: Jenkins and Pérez-Arriaga (2014)

15 The menu of contracts in this table uses the following discretionary parameters: $\omega = 0.66$; $SF_{ref} = 0.7$; $SF_{roc} = -0.01$; $AI_{ref} = 1.0$.

5.2.1.2 Equalizing incentives for operational and capital expenditures

In addition to incentivizing utilities to pursue cost-saving efficiencies, regulatory mechanisms should take care to avoid distorting a utility's incentives to invest in capital assets rather than operational expenditures. As discussed in **Section 5.1.3**, utilities face increased trade-offs between traditional capital investments in network assets and novel operational and network management strategies that harness DERs. To encourage utilities to find the most efficient combination of capital and operational expenditures, financial incentives related to CAPEX and OPEX need to be equalized.

Incentives are typically skewed by conventional regulatory approaches, which add approved capital expenditures directly to a utility's regulated asset base or rate base, while operational expenditures are expensed annually. Even if utilities are properly incentivized to pursue cost savings via a profit-sharing incentive, under this financial framework, a dollar in reduced CAPEX will also involve a reduction in the utility's regulated asset base and thus a reduction in the allowed return on equity and a corresponding decline in net profit for the utility's shareholders. This decline in net profit will offset some portion of any efficiency-related income awarded by the regulator, distorting trade-offs between OPEX and CAPEX and potentially encouraging overinvestment in conventional network assets (Ofgem 2009, 2013c; Jenkins and Pérez-Arriaga 2014; NYDPS 2016b).

Ofgem has developed a mechanism for equalizing these incentives that is known as the total expenditure or "TOTEX-based" approach (Ofgem 2009, 2013c). Under a TOTEX-based approach, the regulator establishes a fixed portion of total utility expenditures (TOTEX), referred to as "slow money," that will be capitalized into the utility's regulated asset base (from which depreciation and cost of capital revenue allowances are calculated). The remainder of TOTEX is designated as "fast money," which is treated as an annual expense. Critically, the regulator fixes these shares at the start of the regulatory period based on an estimate of the efficient split between CAPEX and OPEX in total expenditures. As such, the share of CAPEX and OPEX in actual utility expenditures is free to depart from this expected share without impacting the utility's return on equity. Under this approach, both OPEX and CAPEX savings face the same efficiency incentives—that is, a dollar of OPEX savings and a dollar in CAPEX savings produce the same improvement in utility earnings—freeing the utility to exploit the expanding frontier of cost-saving trade-offs between both types of expenditure.[16]

Alternative measures have been enacted by the New York Department of Public Service (NYDPS 2015a, 2016b) consistent with its cost of service–based regulatory framework and US accounting practices. New York establishes allowed revenues or "base rates" based on a "forward test year" or a forward-looking estimate of expected utility costs. This estimate is based in part on a capital investment plan submitted by the utility and reviewed by the regulator in each rate case. Once base rates are set for a given year, a utility could conceivably increase earnings by withholding funds from capital projects included in the base rates. While this provides an efficiency incentive, New York regulators were historically concerned that this incentive could lead to underinvestment and degradation of network quality, or to rewards for utilities that inflate their estimates of future expenditures during rate cases. Therefore, a "clawback" provision was established wherein regulators automatically reduced a utility's allowed revenues if capital expenditures fell below approved levels, returning

16 For additional discussion of the TOTEX-based approach and resulting incentives for utility cost savings, see Ofgem (2009), pp. 117-120, and Ofgem (2013c), pp. 30-32.

all such earnings to ratepayers. Recognizing that the clawback created an inherent disincentive to pursue cost-saving measures, however, the NYDPS reformed this mechanism in 2016. Its new Reforming the Energy Vision framework aims to incentivize utilities to "pursue cost-saving DER-based alternatives to capital expenditures" (NYDPS 2016b). Via this framework, New York explicitly pledges to allow utilities to retain earnings on capital included in the base rates for the regulatory period, freeing the utility to pursue cost-effective operational expenditures and DER alternatives to planned capital projects. During the next rate case, such DER expenses will be incorporated into base rates and the earnings associated with avoided capital projects will be removed, allowing ratepayers to benefit from any net savings in total expenditures achieved.[17]

Whatever mechanism is pursued, the goal is important: Utilities should be free to find the most cost-effective combination of conventional investments and novel operational expenditures (including payments to DERs) to meet demand for network services at desired quality levels.

5.2.2 Managing uncertainty and information asymmetry

In addition to incentivizing cost-saving efficiency, regulators need new tools to confront the uncertain future facing distribution utilities, improve regulatory certainty for utilities to preserve desired incentives, and manage heightened information asymmetry between regulators and utilities. This will require new ways to establish more accurate forward-looking benchmarks for efficient utility expenditures while reducing information asymmetry. While such measures should improve the accuracy of benchmarks in an evolving power sector, there will inevitably be errors in these benchmarks as new technologies and new techniques change utility practices. Fortunately, the strength of profit-sharing incentives and the length of the multi-year revenue trajectory can also be tuned to appropriately distribute exposure to such benchmark errors (balancing reduced risk exposure against reduced efficiency incentives). Finally, mechanisms to automatically adjust allowed revenues in light of inevitable errors in the forecasted evolution of network uses or other cost drivers can improve regulatory certainty for utilities and improve allocative efficiency for ratepayers by ensuring that utility earnings are tied to each firm's own efforts, not to trends and forces outside a utility's control.

Several state-of-the-art regulatory tools, including an incentive-compatible menu of contracts, an engineering-based reference network model, and automatic adjustment factors to account for forecast errors, can better equip regulators for an evolving and uncertain electricity landscape.

17 The New York Public Service Commission recognizes that the clawback reform discussed here is "not a complete solution to issues around capital and operating expenses" and that a TOTEX-based approach that eliminates the distinction between capital and operational expenditures is "a more comprehensive way to address the potential capital bias" (NYDPS 2016b, pp. 100-101). In addition, the NYDPS restricts waivers of the clawback mechanism specifically to cases "directly linked to a demonstration of the DER alternative that replaced the capital project." This implies increased regulatory risk for utilities, as they must demonstrate, on a case-by-case basis, the particular measures undertaken to harness DERs in lieu of network investments. A multi-year revenue trajectory with profit-sharing incentives and a TOTEX approach to capital bias would thus provide greater regulatory certainty, a reduced regulatory burden, and enhanced efficiency incentives.

5.2.2.1 Forward-looking benchmarks for efficient utility expenditures

As DERs proliferate and network users become more diverse and active, the demands placed on electricity networks will evolve. Likewise, the emergence of DERs, active system management, and ICT capabilities in distribution systems will change the set of tools available to utilities, shifting the "efficient frontier" for utility best practices. Methods for establishing the allowed revenues for distribution utilities should therefore become forward-looking, based not on the best practices and customer preferences of the past, but rather on estimates of what an efficient utility can accomplish and what diverse network users are likely to demand in tomorrow's electricity landscape.

The future will not look like the past, so the past no longer provides a useful benchmark for efficient future expenditures. Forward-looking benchmarking tools are needed.

Diverse regulatory bodies have pioneered a number of forward-looking approaches to establish allowed revenues. Fourteen US state regulatory commissions have used "forward test years" to establish utility cost of service, including commissions in California, New York, Wisconsin, and Illinois, and the experience has been overwhelmingly positive (Costello 2013, 2014). Regulators in Europe and elsewhere commonly employ multi-year revenue or price cap regulation that requires the establishment of forward-looking *ex ante* estimates of efficient utility costs as well, and they have developed a variety of tools to produce higher quality forecasts and benchmarks.

Several state-of-the-art methods can better equip regulators to establish forward-looking benchmarks of efficient utility expenditures, contending with both heightened information asymmetry—the likelihood that utilities know more about their expected costs and opportunities than regulators—and the likelihood that the future will be different from the past, meaning that prior best practices fail to offer a good guide for future costs.

One approach to reduce information asymmetry between regulators and utilities has already been described: the incentive-compatible menu of contracts. With this approach, utilities profit most when their actual expenditures match their reported forecasts; utilities are thus rewarded for submitting their most accurate and truthful forward-looking estimates of efficient network costs to regulators during each regulatory review period or rate case. This method has been successfully implemented by the Ofgem since 2004 (where it is known as the Information Quality Incentive) and offers a powerful tool to reduce information asymmetry (Ofgem 2004, 2009, 2013b).

Regulators can also use other ways to incentivize utilities to submit high-quality business plans. For example, Ofgem expedites the regulatory review process for the utilities that submit the highest-quality business plans during each distribution price control review process (Ofgem 2010a, 2013d).[18] Utilities fast-tracked during the first RIIO electricity distribution price control review process, which was completed in 2015, finished the bulk of the regulatory process a year ahead of other utilities (Lowry and Woolf 2016). This significant reduction in regulatory costs and improved regulatory certainty offers a more intangible but still important incentive for utilities to submit high-quality business plans and reduce information asymmetry.

18 In the first RIIO electricity distribution price control review (ED1), Ofgem assessed business plan quality according to five broad criteria: process, outputs, efficient expenditure, efficient finance, and uncertainty and risk. Process criteria included how clear the business plan document was, how extensively stakeholder input was engaged in developing the plan, and how reasonable the plan was. The efficient expenditure criterion was based on a regulator benchmark process that involved expert and consultant review of the business plan and how well the utility justified proposed expenditures. The finance metric involved review of the utility's regulatory compliance strategy, consistency with best practices, and justification of its finance plan (including debt and equity sources). Finally, the uncertainty and risk metric captured how well the business plan identified and articulated the key risks and uncertainties the utility faced and the strength of the mitigation measures proposed in its plan (Ofgem 2010a, 2013d).

Finally, regulators can equip themselves with a forward-looking engineering-based reference network model (RNM) to assist in developing benchmarks for efficient network expenditures. An RNM emulates the network planning practices of an efficient utility and equips the regulator with a forward-looking benchmark that accommodates expected evolutions in network use, technology performance and costs, and network management practices (Domingo et al. 2011). The forward-looking capabilities of an RNM-based benchmarking approach contrast with more commonly employed statistical benchmarking techniques, which analyze previous costs incurred by similar utilities to develop statistical estimates of efficient utility expenditures (controlling for various factors outside the utility's purview, such as density of network topology, etc.) (Cossent 2013). These statistical approaches are necessarily backward-looking since they depend on raw data from past utility expenditures and thus cannot capture the dynamic changes now unfolding in the electricity distribution sector. In contrast, an RNM gives the regulator a tool with which to "peer into the future" and estimate efficient costs given expected changes in network uses (e.g., changes in electricity demand and DER adoption) and in the cost and availability of different network assets and management practices over the regulatory period. In addition, an RNM directly accounts for each utility's existing network topology and customer demand profile, a particularly important feature as DER penetration is likely to increase the heterogeneity of distribution networks.[19]

The information requirements necessary to employ an RNM can be significant, but a number of countries, including Spain, Chile, and Sweden, have used RNMs to assist in forward-looking benchmarking (see Jamasb and Pollitt [2008] and Cossent [2013] for details). Plus, as electric utilities adopt electronic equipment inventories and geographic information systems, it is likely to become easier for them to meet the reporting requirements necessary for the regulator to employ an RNM. RNMs have already demonstrated their ability to assess the impact on distribution planning and costs due to large-scale deployment of distributed generation (DG) and electric vehicles as well as the use of active network management (Cossent et al. 2011; Fernández et al. 2011; Olmos et al. 2009; Vergara et al. 2014).[20] Note that Jenkins and Pérez-Arriaga (2014, 2017) describe in detail the step-by-step application of an RNM to establish a multi-year revenue trajectory with profit-sharing incentives.

19 Note that the RNM is only a tool, and it must be applied with appropriate regulatory discretion as part of an overall benchmarking and remuneration process. The RNM does not estimate several important company cost drivers, including depreciation of existing assets, business support costs, overhead costs, staff salaries, investments in refurbishments, financial costs, property taxes, etc.

20 We have seen in **Chapter 4** that network reinforcements have an important impact on the profitability of DER investments. As network users increasingly consider investments in technologies, the benchmark tool should take into account the optimal coordination between network and DER investments. One way of implementing this is through a simulated iterative investments discovery mechanism including an RNM and an optimal private DER investment model.

Box 5.2: How Does a Reference Network Model Work?

An RNM typically takes as input the location and power injection/withdrawal profile of all network users as well as a catalog containing technical and cost information about available equipment, probability of component failure, and the cost and time burden of maintenance. Given these inputs, the RNM constructs a network to serve all users while minimizing total costs (including capital expenditures, operational expenditures, and a specified penalty for thermal network losses) and meeting three specified quality of service constraints: (1) maximum system average interruption duration index (SAIDI); (2) maximum system average interruption frequency index (SAIFI); and (3) maximum and minimum acceptable voltage range at every node.

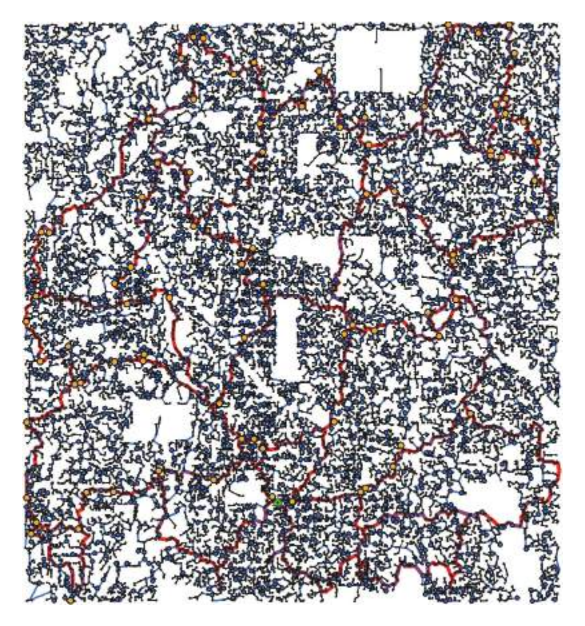
For regulatory benchmarking purposes, it is important to take into account the established layout of the utility's network and sunk investments in network components. The RNM should thus be run in a "brownfield" or network expansion mode, taking as inputs the existing network layout and location of the utility's network users and specifying the layout of network reinforcements and extensions necessary to serve projected changes in network use over the regulatory period. Regulators would thus have to require utilities to report information on their existing networks in a standard format including the location, voltage level, contracted capacity, and injection/withdrawal profile of all existing network connections (loads and DERs); the layout, impedance, and capacity of electrical lines and protection devices; and the capacity and location of transmission interconnection substations, high-/medium-voltage substations, and transformers. Explicitly taking into account the heterogeneous nature of distribution networks is thus another key advantage of the RNM over statistical benchmarking techniques.

The regulator must also maintain a detailed library of the standard network components used by the RNM, including cost and performance characteristics of cables, overhead lines, distribution transformers, substation components, and protection devices. This catalog should adequately characterize the real investment alternatives the utility may face. The library should be updated regularly to reflect the current cost of standard components and expanded to include any new components entering common use, such as ICT equipment and advanced power electronics. To minimize opportunities for the strategic inflation of reported component costs, the regulator or suitable government agency[21] should determine costs for library components by benchmarking efficient unit costs across multiple utilities.

To ensure the RNM is suitable for use in regulatory proceedings, it must be accurate enough to simulate established industry best practices. Over time, as novel, emerging techniques such as active system management, coordinated dispatch of DERs, or other measures become standard practice, the model should be updated to include these practices. However, it is appropriate for the RNM used in benchmarking to reflect current best practices at the beginning of the regulatory process. New methods adopted during the regulatory period that successfully reduce total system costs (or improve performance) will then be rewarded appropriately, and the model can be updated periodically to reflect any practices that have become widespread. In this manner, this proposed method works well despite the evolving nature of distribution network management, provided a suitable RNM is available that captures the best practices of an efficient network utility at the beginning of the regulatory period.

21 In many US states, there may be a very limited number of distribution utilities from which any individual regulator could draw data. A federal agency such as the US Department of Energy or a designated independent entity may assist regulators in developing RNM frameworks and standard inputs for use by regulators.

Figure 5.1: Sample Synthetic Network Output of a Reference Network Model



Source: Jenkins and Pérez-Arriaga (2014)

5.2.2.2 Tuning the multi-year revenue trajectory and profit-sharing incentives to distribute the risk of benchmark errors

The regulatory tools described above can help overcome information asymmetry and improve the accuracy of regulatory benchmarks and forward-looking revenue trajectories. Yet in today's uncertain era, benchmark errors can still occur when regulators fail to anticipate the emergence of new cost-saving technologies or practices within the regulatory period that shift the efficient frontier. To some degree, the regulator's forward-looking benchmark of efficient expenditures can lag behind the best practices likely to emerge over the next regulatory period. For example, if the regulatory benchmark is based on best practices at the beginning of the regulatory process, then the adoption of any new methods that successfully reduce total system costs will be rewarded, and regulators can update their estimates of best practices accordingly. However, if the regulator's benchmark becomes too outdated, utilities may earn substantial rents, distorting allocative efficiency.

Regulators employing multi-year revenue trajectories have two tools to manage the impact of such benchmark errors (beyond improving the quality of the benchmark itself using the tools described above).

First, the length of the regulatory review process directly affects its exposure to uncertainty and benchmark error (Ofgem 2010a, 2013a). The longer the regulatory period, the greater the opportunity for the efficient frontier to evolve and move out of step with the established benchmark. Conversely, the shorter the period, the more frequently regulators can update their estimates of efficient utility costs to reflect evolving best practices (although at the cost of reduced incentives for productive efficiency). Second, the strength of any profit-sharing incentive also helps manage exposure to uncertainty about future costs and demand. In particular, under lower levels of uncertainty, a higher profit-sharing factor (meaning the utility is exposed to most of the risks and rewards of cost savings) performs better, while a lower profit-sharing factor (which shares most risks and rewards with ratepayers) performs better under higher levels of uncertainty (Schmalensee 1989). In both cases, however, mitigating exposure to benchmark error and thus improving allocative efficiency comes at the cost of reduced incentives for pursuing cost-saving measures and improving a utility's productive efficiency. The twin goals of allocative efficiency and productive efficiency should therefore be carefully balanced. Regulators can tune both the length of the regulatory period and the strength of the profit-sharing factor to achieve the desired balance.

5.2.2.3 Automatic adjustment mechanisms to resolve forecast errors

In addition to benchmark errors, there will always be uncertainty regarding the evolution of network uses, cost of capital, and network component costs over the regulatory period. This uncertainty can lead to "forecast errors," where costs rise or fall unexpectedly due to new network uses (e.g., growth in solar PV penetration or electric vehicle adoption) or other cost drivers (e.g., changes in cost of capital or commodity prices) that fall outside the utility's direct influence. If forecast errors are unmitigated, the economic sustainability of regulated utilities could be jeopardized if allowed revenues are set too low; alternatively, utilities could profit from significant economic rents if the revenue trajectory is too generous. For example, the rapid and unanticipated growth of electric vehicle adoption could drive new demand for network upgrades that a multi-year revenue trajectory fails to anticipate. Alternatively, anticipated load growth could fail to materialize due to economic recession. The general principle should be that utility earnings are tied to the exercise of managerial effort and quality of service, not to factors outside the utility's direct influence.

Fortunately, regulators have developed a range of mechanisms designed to manage impacts when the actual evolution of network uses and cost inputs diverges from forecasts. For a discussion of uncertainty mechanisms available for *ex ante* or forward-looking regulatory approaches, see Ofgem (2010a, 2013a). Options for managing uncertainty include:

- Indexing provisions to adjust allowed expenditures based on changes in economy-wide price indexes (e.g., inflation of input costs, fuel price changes, or changes in average cost of debt).
- Revenue trigger provisions that specify a rule to increase or decrease revenues by a specified amount if certain pre-specified trigger events occur (e.g., substantial departures from forecasted network uses).
- Reopener thresholds that specify clear conditions under which expenditure and revenue determinations will be revisited and revised (e.g., to account for major new legislative requirements, changes in the tax code, substantial departures from forecasted network uses, or other major unanticipated cost drivers).
- Reviews of output requirements halfway through the regulatory period. For example, Ofgem's RIIO multi-year revenue trajectory period lasts eight years but includes a mid-period review at four years to capture changes in output requirements.
- Automatic adjustment factors or forward-looking volume drivers—i.e., predetermined formulas to adjust revenues if key cost drivers (e.g., number of customer connections, load growth, DER penetration) depart from *ex ante* forecasts.

An RNM can be employed not only to assist in initial benchmarking of efficient utility expenditures, but also to calculate *ex ante* automatic adjustment factors to correct the estimate of efficient network expenditures to account for any deviations from the forecast for load growth and DER adoption. Jenkins and Pérez-Arriaga (2017) describe in detail how an RNM can be used to determine the appropriate formulas—in short, regulators can employ an RNM to estimate network costs across a range of uncertainty scenarios designed to capture likely load patterns, DG penetration, electric vehicle adoption, or other important and uncertain cost drivers. The regulator then runs the RNM in a brownfield mode to calculate efficient network expansion costs under each of these scenarios. Finally, the regulator determines the relationship between deviations in cost driver values and efficient network costs by performing a multivariate linear regression on the resulting estimated total network expenditures for each scenario. The coefficients of this regression function, which we call "delta factors," prescribe simple formulas to adjust the estimated TOTEX baseline *ex post* based on the realized evolution of network uses or other key cost drivers. Delta factors reduce the risk that the revenue determination will need to be reopened during the regulatory period, significantly increasing regulatory certainty. These delta factors also align incentives for the utility to connect and serve new DERs by ensuring cost recovery even if DER penetration grows more rapidly than expected. Finally, by automatically reducing allowed revenues if projected network use fails to materialize, delta factors improve allocative efficiency and prevent utilities from profiting from events beyond their influence.

5.2.2.4 Progressive profit-sharing mechanisms

Finally, establishing progressive profit- or earnings-sharing mechanisms can help attenuate the impact of both benchmark and forecast errors. Under a progressive profit-sharing mechanism, the strength of the profit-sharing efficiency incentive is progressively reduced as the deviation between expected and actual expenditures increases (Malkin and Centollela 2013). For example, the percentage of earnings shared with customers could increase as the utility's actual return on equity exceeds authorized levels. Similarly, a symmetrical approach would protect the utility from a confiscatory result by sharing an increasing percentage of any reasonable, unanticipated costs with ratepayers if earnings fell below anticipated levels. The attenuation of the profit-sharing incentive rate could be smooth (e.g., the share of retained profits could reduce linearly or by some other continuous function as total savings increase) or through a stair-step or tiered pattern (e.g., the profit-sharing incentive reduces in stages as total savings pass some threshold). In general, a continuous function for the change in the profit-sharing rate will prevent discontinuous jumps in the marginal incentives facing utilities and is preferable. Progressive profit-sharing mechanisms can help determine the maximum possible impact of benchmark or forecast errors on utility earnings and can be another effective tool for mitigating the effects of regulatory uncertainty.

5.3 Regulatory Tools to Promote Other Important Outcomes: Quality of Service, Innovation, and Cybersecurity

Making the transition to an actively managed distribution grid with greater DER penetration will require tools that fall outside the scope of the core remuneration framework. First, additional incentives will be needed to move network utilities toward critical objectives or outcomes that are not sufficiently incentivized under the core remuneration framework. These include objectives such as maintaining and improving the quality of electrical service, reducing energy losses, proactively stewarding network assets, and ensuring high levels of workplace and public safety. The design and implementation of a sound set of incentives for outcomes such as these will play an important role, both in the transition to actively managed networks and in ensuring the continued safe, reliable, and efficient operation of electricity networks.

Second, additional incentives are needed to promote long-term innovation—that is, investment in demonstration projects that provide technological learning and in the dissemination of that knowledge among utilities. The regulatory strategies discussed in **Section 5.2** are critical for incentivizing cost-saving measures in a rapidly evolving power sector, but they typically create incentives that work only over the relatively short term. The core remuneration framework described above is thus effective at driving incremental or process innovation. To achieve medium- and long-term economic efficiency, networks need to be incentivized to invest in demonstration projects that are inherently risky and promise longer-term returns.

Finally, new precautions and standards are needed to protect against cyber attacks. Since the proliferation of DERs and ICTs in electricity networks will continue to increase the vulnerability of these systems, strong cybersecurity regulations and standards are needed for all levels of the power system.

5.3.1 Outcome-based incentives

Short-term economic efficiency is not the only goal of regulation. Other important outcomes include improving the quality of electrical service, reducing energy losses, proactively stewarding network assets, ensuring workplace and public safety, and others. These outcomes are not incentivized by core remuneration frameworks since achieving them imposes costs on distribution companies, and such frameworks focus on reducing costs. Therefore, many regulatory authorities have created additional incentives to promote these critical outcomes.

Outcome-based performance incentives can reward utilities for improvements in quality of service, such as enhanced resiliency, reduced distribution losses, and improved interconnection times.

Ideally, such mechanisms should aim to replicate the outcomes of efficient markets, where consumers are willing to pay more for products or services that deliver greater value. Value for money is the overarching concern, not simply lowest cost. Outcome-based incentives are thus a critical complement to the cost-saving efficiency incentives discussed above.

In this section, we discuss and make recommendations regarding regulatory strategies for the design of outcome-based performance incentives. We focus on four performance areas—commercial quality, continuity of electrical supply, voltage quality (which together comprise quality of service), and energy loss reduction—drawing from best practices in Europe and the United States before delving into specific performance areas. However, we first present some general considerations for the design and implementation of outcome-based incentives, which will be useful for the discussion.

Figure 5.2: Stylized Illustration of the Implementation of an Outcome-based Performance Incentive

Source: Aggarwal and Harvey (2013)

5.3.1.1 General considerations for the design and implementation of outcome-based incentives

The basic regulatory design elements and stages of implementation of outcome-based performance incentives are illustrated in **Figure 5.2**.[22] The left side of the figure enumerates the basic regulatory elements that may be used in the design of outcome-based incentives:

1. A suitable performance metric and established reporting requirements for the network utility
2. A minimum performance target and specified time period for achieving that target
3. A reward-penalty ("bonus-malus") scheme that rewards the utility for exceeding the performance target and penalizes the utility for failing to meet the target

The right side of Figure 5.2 illustrates the implementation of an outcome-based performance incentive that includes all three of the aforementioned elements. The distribution company is given a period of time during which it must improve its performance in meeting the pre-established target or face a penalty (or possibly a sanction if there is severe underperformance). If each of the elements of the outcome-based incentive is well-designed, then the network utility will successfully meet or exceed the target at a cost that is no greater than the benefit received by the performance improvement. Considerations for establishing an efficient performance target are discussed in **Box 5.3**, while considerations for establishing an incentive formula for a reward-penalty scheme are discussed in **Box 5.4**.

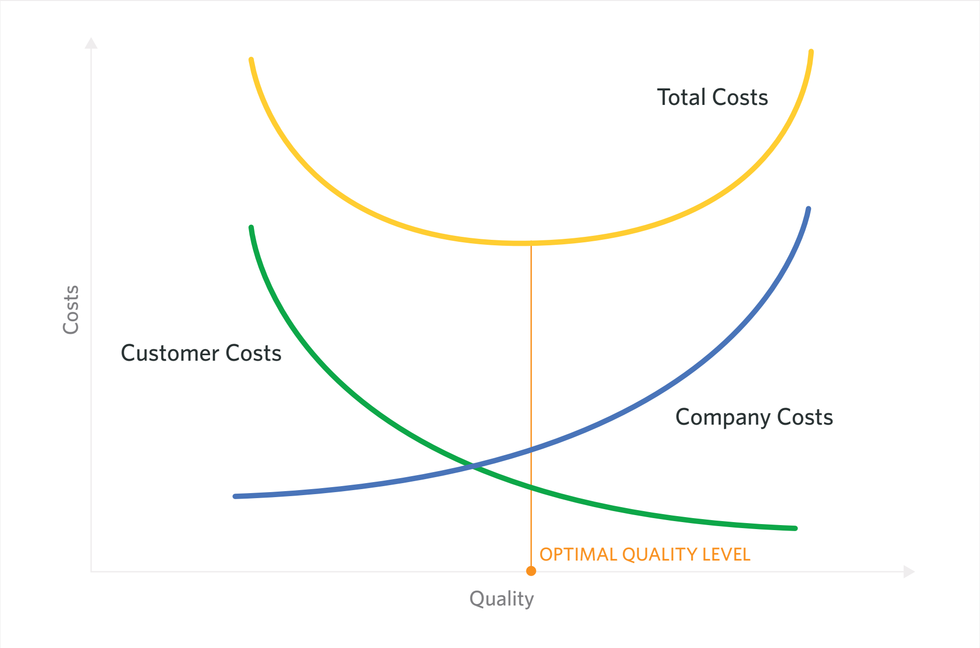
Note that while Figure 5.2 depicts the implementation of an incentive that includes all possible elements, in practice many outcome-based performance incentives are designed with only a subset of the elements described above. For instance, in some cases a utility may not be subject to meeting a specific target but may simply be required to report its performance to the public on a regular basis. There is good evidence that the mere dissemination of information regarding a company's performance provides an incentive for investing in improvements. In other cases, a utility may be subject to a minimum performance target but not to a reward–penalty scheme, or may be subject to a reward-only or penalty-only scheme.

22 This figure was adapted from Aggarwal and Harvey (2013).

Box 5.3: Establishing an Economically Optimal Performance Target

Careful analysis is essential to setting any performance target for an outcome-based incentive. From a social perspective (i.e., a perspective that includes the interests of both customers and companies), the optimal level of performance is the level at which the marginal benefit of additional performance equals the marginal cost of supplying it. The benefit—or willingness to pay (WTP)—of an increment of performance improvement is usually difficult to quantify, so in practice WTP is often approximated by its inverse: the cost of—or willingness to avoid (WTA)—incurring an increment of performance deterioration. From this perspective, the optimal level of performance is that which minimizes the total cost function—that is, the sum of the company's cost function and the customers' cost function (Fumagalli et al. 2007; Gómez 2013; Rivier and Gómez 2000). This concept is illustrated in **Figure 5.3**.

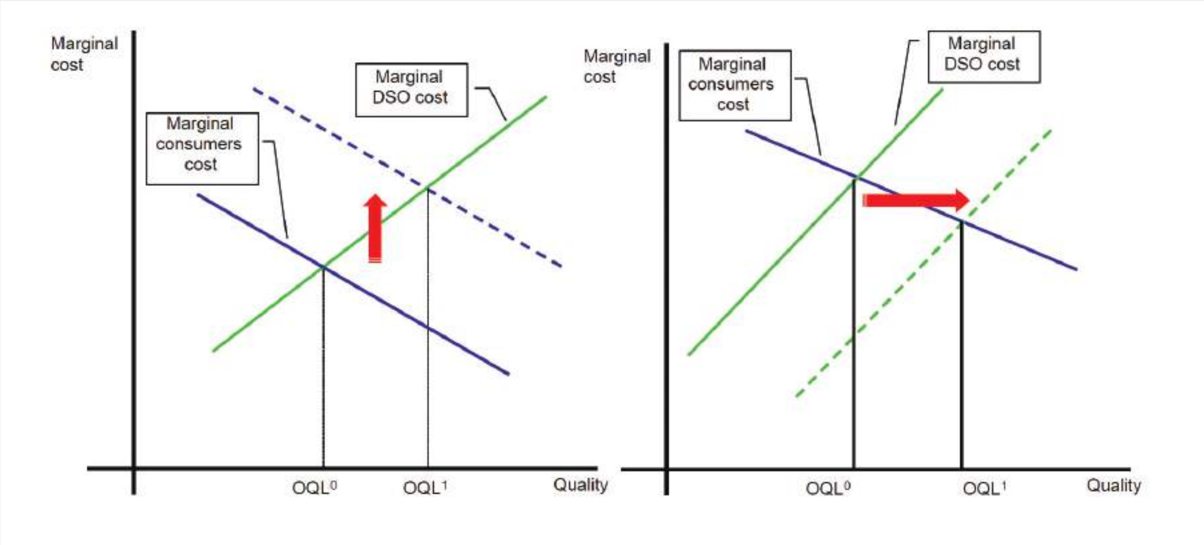
Figure 5.3: Initial Optimal Level of Performance Quality



In practice, of course, determining these values is not straightforward. Establishing customers' WTA is an inexact science that usually depends upon customer surveys and can vary widely within a given distribution network. Likewise, while estimates of a company's cost for delivering various levels of performance can be more accurate, these estimates too are usually rough (at best) due to the complexity of distribution networks.

Regulators should also bear in mind that optimal levels of quality will change and may be very different for different types of customers; therefore, performance target levels should evolve over time and may need to become more granular. Over time, the perceived importance of a given performance goal for network users may vary. Likewise, the cost to the distribution utility of implementing a given performance goal may vary over time as technology costs change and as economies of scale are realized. Consider the example of reducing connection times for network users with solar PV. If prices are more cost-reflective, as more solar PV is installed in a network (or in a certain area of a network), the economic incentive for a network user to install solar PV may decline (since the value of solar PV in a given network location decreases as a function of the total installed capacity). Therefore the user's willingness to pay for a speedy connection time may also decline. Similarly, as distribution companies become more efficient at interconnecting solar PV and realize economies of scale, the marginal cost of ensuring timely connections may decrease. A final determination about the optimal level of quality will depend on the direction and magnitude of these changes. As depicted in **Figure 5.4**, the optimal level of quality will increase if network user WTP increases, or if the marginal cost of performance delivery decreases. Conversely, the optimal level of quality will decrease if network user WTP decreases or if the marginal cost of performance delivery increases.

Figure 5.4: Evolution of Optimal Level of Performance



Source: Cossent (2013)

Finally, regulators must be wary of the fact that oftentimes achieving a specific performance outcome may increase a company's OPEX and/or CAPEX efficiency. Therefore, in jurisdictions that include cost efficiency incentives in the core regulatory framework (such as the efficiency incentives described in **Section 5.2.1**), regulators must ensure that efficiency incentives and outcome-based incentives are tightly coupled—that is, that the expected costs associated with the performance outcome are built into the cost-efficiency trajectory to which the utility is subject. This becomes increasingly important as the share of earnings that a network utility can achieve via outcome-based incentives becomes larger.

Box 5.4: Establishing Output-based Incentive Formulas in a Reward–Penalty Scheme

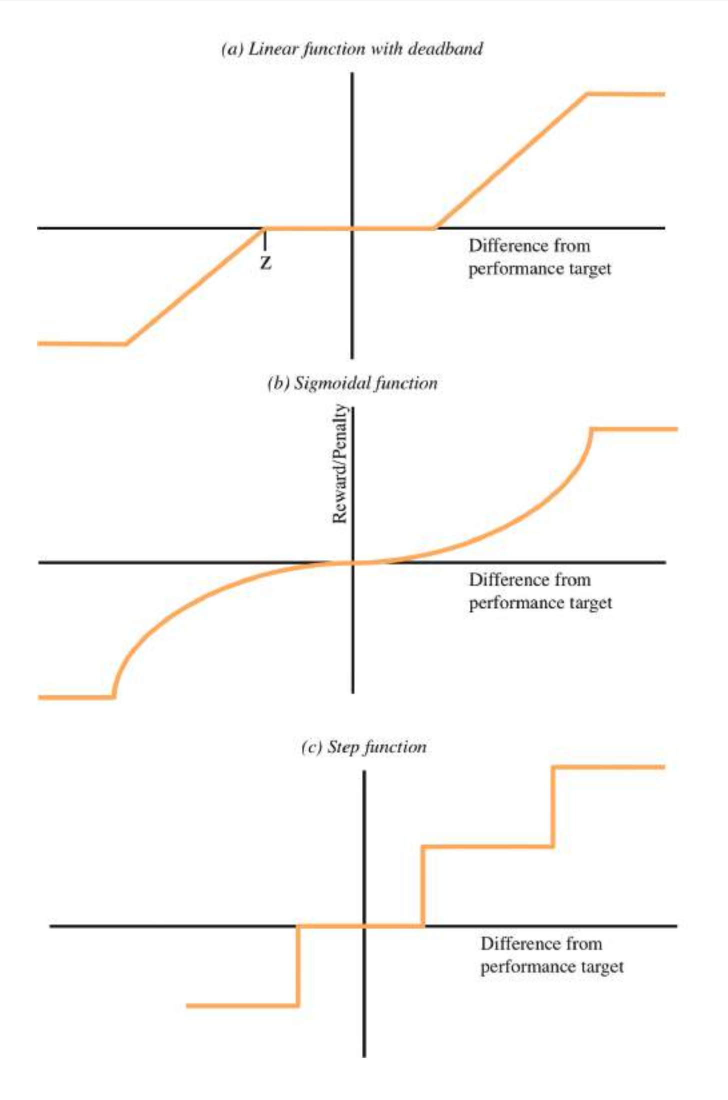
The inclusion of a financial reward–penalty scheme in a performance incentive allows the regulated company some degree of discretion: Given the performance target and the associated financial incentives, the company is expected to employ its superior knowledge of costs to deliver an efficient level of performance (Sappington 2005). Financial reward–penalty schemes have delivered positive results where they have been applied (Fumagalli et al. 2007). However, they are complex to design and therefore should be approached with appropriate deliberation and analytical rigor.

The incentive formula of a financial reward–penalty scheme can take numerous forms, including linear, quadratic, and step function. As much as possible, the shape and the slope of the formula should reflect customers' willingness to pay for increases in the performance level of a given outcome. Practical considerations, such as ease of implementation, may also come into play in the choice of an incentive formula. **Figure 5.5** presents examples of possible functional forms that reward–penalty schemes may take, as well as some of their potential benefits and drawbacks.

- A *linear function* is a popular functional form since it is simple to understand and to administer. **Figure 5.5a** displays a linear function with a deadband and caps. The deadband helps reduce the risk of the distributor achieving a reward/incurring a penalty due to a change in performance that occurs as a result of factors outside the distributor's control. A potential drawback of a linear function with a deadband is that it might incentivize the distributor to perform at the most negative performance level within the deadband (Point Z). A linear function need not have a deadband or caps. For example, incentive formulas for TOTEX efficiency are typically linear with no deadband or caps.
- A *sigmoidal function* (**Figure 5.5b**) may be used as an alternative to a linear function with a deadband, since the slope of the sigmoidal function (i.e., the rate of change of the reward/penalty) is relatively low at small deviations from the performance target. The parameters of the function can also be tweaked so that the maximum reward and penalty occur at the same performance level as a similar linear function with a deadband. Moreover, the fact that the slope of the sigmoidal functions increases/decreases as the deviation from the target increases/decreases provides an additional incentive to strive for high performance levels and avoid low performance levels. However, a sigmoidal function may be more difficult to administer than a linear function.

- A *step function* (**Figure 5.5c**) may be used if the objective is to provide a "sudden" reward/penalty if the distributor deviates from the performance target by a certain value. A step function may have as few as two steps or as many as the regulator deems appropriate. While step functions are easy to administer, the fact that they are associated with "sudden" rewards or penalties can lead to contentious debate about whether the distributor did or did not deviate from the performance target by the set amount. Moreover, this feature may incentivize the distributor to engage in unsound practices either to receive a large reward or to avoid a large penalty.

Figure 5.5: Examples of Functional Forms for Reward–Penalty Schemes



5.3.1.2 Quality-of-service incentives

Service quality in the electricity distribution and retail sectors spans many technical and non-technical aspects.[23] The aspects that are normally regulated can be grouped into three areas, one of which is non-technical and two of which are technical. First, commercial quality includes ensuring adequate performance on a number of service quality measures, such as the provision of new electrical connections, meter reading, billing, and the handling of customer requests and complaints. Second, continuity of supply is related to events during which the voltage at a customer connection drops to zero.[24] Finally, voltage quality is concerned with minimizing deviations from desired levels of voltage characteristics. The design of regulation for each of these three services will be significantly affected by the emergence of DERs. Moreover, the transition to actively managed networks will mean that DERs may increasingly participate in the provision of some of these services. Each of these three performance dimensions, and best practices for their regulation, are briefly discussed below.

Commercial quality regulation addresses the non-technical aspects of quality-of-service regulation and is concerned with maintaining minimum levels of quality for services that relate to the interaction between service providers (the distribution utility or retailer) and customers. Commercial quality regulation employs all three of the regulatory elements described in **Section 5.3.1.1**: the establishment of performance metrics (and reporting requirements), the establishment of a minimum performance target, and (rarely) reward and penalty schemes. However, it is recommended that when introducing a new commercial quality incentive, regulators do not adopt all three elements at once but rather introduce them incrementally, beginning with reporting requirements and moving (if necessary) to the establishment of a performance target and (again, if necessary) to the establishment of a reward–penalty scheme. Introducing regulatory elements in succession provides time to monitor and fine-tune each of the elements, which are often quite complex.

Table 5.2 provides a list of some of the most frequently regulated commercial quality services in the European Union. These include services that are provided before the supply of electricity begins and those that are provided during an active contract. Services that are provided during an active contract can be further divided into those that are provided regularly and those that are provided occasionally. Services that are provided before supply begins include connection, meter installation, and responses to requests for information. The metric typically used to regulate these services is the waiting time for the provision of a service following a service request.

Services that are regularly provided during an active contract include billing, meter reading, and customer service. The performance of these services is usually assessed with metrics that measure regularity and accuracy such as the number of incorrect bills, the frequency of meter readings, customer satisfaction with respect to the precision of information provided by call center staff, etc. Services that are occasionally provided include responding to a customer's request for technical assistance (e.g., to address problems with the meter or the supplied voltage), responding to information requests, etc. Metrics used to regulate these services typically include the amount of time it takes for the service provider to respond to requests and results of customer satisfaction surveys.

23 For a richer discussion of service quality regulation, see Fumagalli et al. (2007).

24 According to the European Norm EN 50160, a supply interruption is a condition in which the voltage at the supply terminals is below 1 percent of the declared voltage. The declared voltage is normally the nominal voltage of the system (i.e., the voltage by which the system is designated or identified), unless a different voltage is applied by agreement between the supplier and the customer (CENELEC 1999).

Table 5.2: Commonly Regulated Commercial Quality Performance Areas in the European Union

BEFORE SUPPLY	DURING VALIDITY	
	REGULAR	OCCASIONAL
Providing (supply and meter)	Accuracy of estimated bills	Responding to failures of a distributer's fuse
Estimating charges for connection*	Accuracy of meter readings	Voltage complaints
Execution of connection-related works*	Service at customer centers	Meter problems
	Service at call centers	Appointment scheduling
		Responding to customer requests for information Responding to customer complaints Reconnection following lack of payment Notice of supply interruption

*Requested also during contract validity

Source: Fumagalli et al. (2007)

Increasing use of DERs will usher in a variety of new considerations for commercial quality regulation. As electricity customers increasingly become electricity service providers, their ability to quickly and efficiently gain access to electricity services markets will become more important. Additional issues such as the availability of and ease of access to information about electricity services markets will also grow in importance. Network users will increasingly look to distribution utilities or retailers to provide platform services that enable them to enter electricity services markets. Commercial quality regulation will need to keep pace with these evolving needs and priorities, and if it is successful in doing so will contribute to innovation and to the transition to actively managed networks.

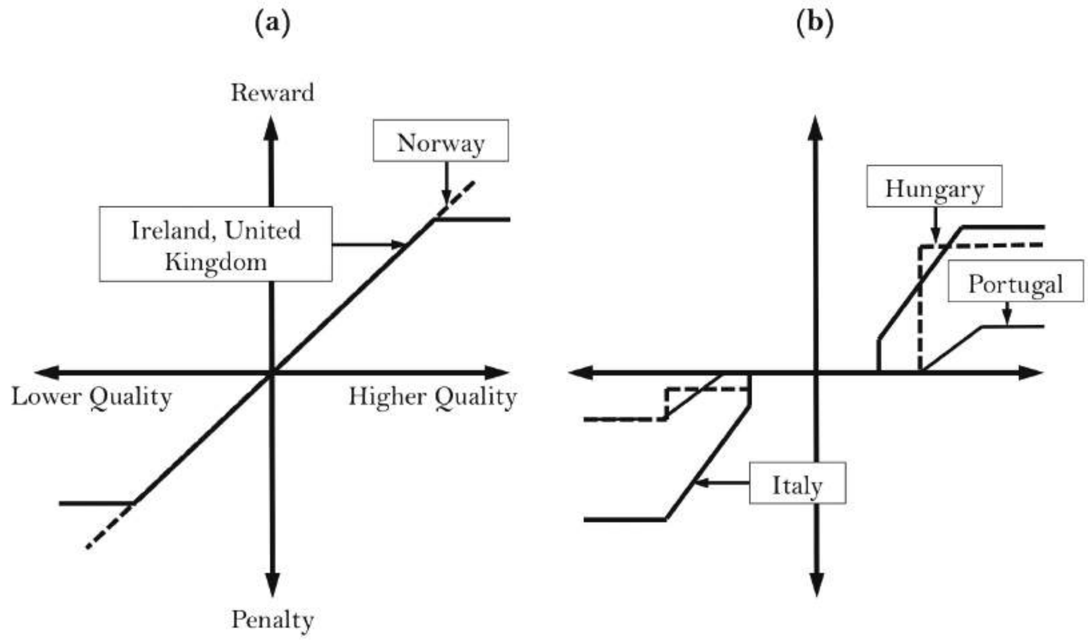
Continuity of supply regulation seeks to minimize interruptions of electrical supply—that is, events during which the voltage at a customer connection drops to zero.[25] In practice, continuity of supply regulation uses all three of the elements described in **Section 5.3.1.1**: the establishment of performance metrics (and reporting requirements), the establishment of a minimum performance target, and reward and penalty schemes. However, as with commercial quality regulation, we do not recommend the adoption of all three regulatory elements at the outset. Rather, we recommend the gradual implementation of each regulatory element in succession to allow ample time to fine-tune each element. The introduction of sound continuity of supply regulation can take one to two years' work (three to four years if reliable data are unavailable) on the part of both the regulator and the distribution companies.

Continuity of supply regulation is primarily concerned with three performance categories: the number of outages experienced by customers in a given time period, the duration of outages (which includes both the average duration of outages as well as the total duration of outages in a given period), and the intensity of outages (i.e., total amount of load that was curtailed). Moreover, continuity of supply regulation is concerned both with "long" unplanned interruption events and with "short" events. Unplanned interruptions are classified as "long" when they last more than three minutes, and as "short" when they last for up to three minutes (CENELEC 1999).

With respect to long unplanned interruptions, three of the most frequently used system-level statistical indicators are: average number of interruptions per customer per year (also known as the system average interruption frequency index, SAIFI); average interruption duration per customer

25 For a richer discussion of continuity of supply regulation, see Cossent (2013).

Figure 5.6: Shape of the Incentive Formula for Continuity of Supply Regulation in Different Countries



Source: Fumagalli et al. (2007).

per year (also known as system average interruption duration index, SAIDI); and energy-not-supplied. Regulation to minimize short interruptions (less than or equal to three minutes[26]) has been less common than regulation to minimize long interruptions. Nonetheless, in recent years this metric has received more attention from customers (particularly industrial customers, since they are the most sensitive to short interruptions) and regulators. As with long interruptions, when data are available for the number of customers affected by short interruptions, system-level indicators of performance can be calculated as the average number of short interruptions per customer per year (also known as the momentary average interruption frequency index, MAIFI).

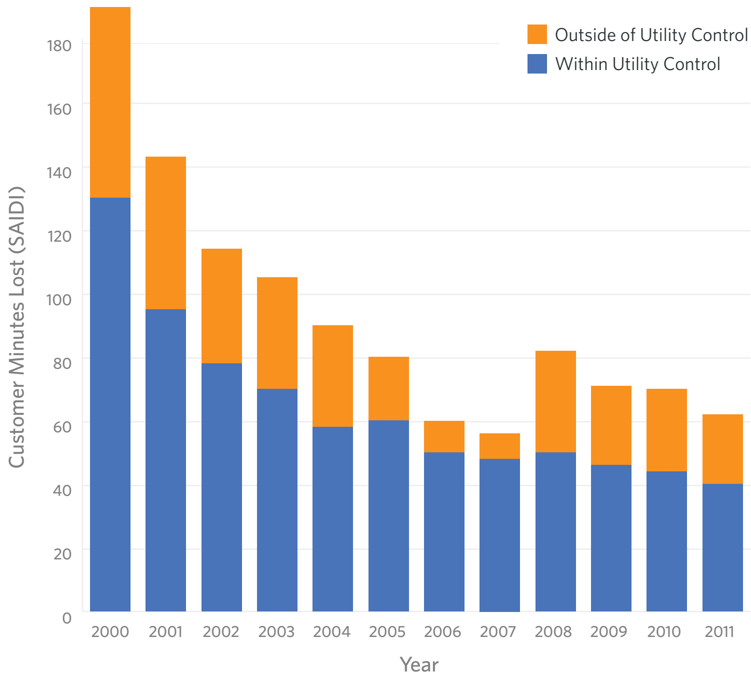
After a target has been set, the regulator must determine the functional form that relates level of quality to financial reward or penalty. To the degree possible, the functional form should reflect customers' WTP for increases in quality as reflected, for example, in customer surveys.

In practice, there is significant diversity in the incentive functions chosen by the different countries that have implemented reward–penalty schemes for continuity of supply. **Figure 5.6** shows the shape of incentive functions for a variety of European countries. **Figure 5.6a** shows countries that use linear functions (one with no cap and two with caps), while **Figure 5.6b** shows countries that utilize caps as well as deadbands.

Continuity of supply regulation has been widely adopted both in the United States and in the European Union and has proven extremely successful in many jurisdictions in significantly reducing outages (both in number and frequency). For example, as shown in **Figure 5.7**, in Italy over a span of three regulatory periods (2000–2011), the number of customer minutes lost was reduced from more than 180 in 1999 to less than 60 in 2011 at an average cost to customers of 2.5 euros per customer per year. Figure 5.7 distinguishes between customer minutes lost that are "within utility control" and customer minutes lost as a result of events that are outside the control of the distribution operator ("outside of utility control"). Italian regulators consider outage events that are under the control of the distribution utility to be those that originate

26 "Micro-cuts," losses of supply that last only fractions of a second, are generally regulated under voltage quality standards and regulations and not under continuity of supply regulation. Micro-cuts can be caused by a fault in a piece of equipment, changes in network configuration, and other reasons; they can be a serious problem for some industrial applications.

Figure 5.7: Average Interruption Duration per Customer per Year (SAIDI) in Italy (2000-2011)



Source: Lo Schiavo (2016)

in the distribution network (low or medium voltage) and those that are not caused by a *force majeure* event (i.e., an event that stresses the distribution network beyond its designed resilience limits). Outage events that are beyond the control of the distribution utility include those that originate in the transmission system, those that trigger a system-wide response, and those that are caused by *force majeure* events.

Increasing penetrations of DERs and ICT-enabled smart grid technologies will provide new opportunities and challenges for continuity of supply regulation that, if acted upon, will contribute to innovation and to the transition to actively managed networks. For example, the presence of DERs may contribute to system reliability through islanded operation in which a portion or portions of the electricity network may function even if there are electrical outages in other areas of the network (Cossent 2013; McDermott and Dugan 2003). In this case, well-designed output-based performance incentives for distribution utilities—combined with regulations that equalize incentives for OPEX and CAPEX—could lead to markets for distribution-level reliability services in which DERs play an active role.

Moreover, regulators will have to carefully consider how to update continuity of supply regulation in light of the increasing penetrations of technologies into distribution networks. For an increasing number of network users, the cost of a service interruption is not only the cost of lost energy consumption but also the cost of foregone revenues from distributed generation production (Cossent 2013).

Also, the increasing use of electric vehicles will link mobility to electricity supply, which may significantly increase customers' WTP for continuity of supply. Finally, continued deployment of ICT-enabled smart grid technologies and systems will very likely reduce the costs associated with providing continuity of supply. All of these changes mean that regulators that have established outcome-based performance incentives for continuity of supply will have to regularly reevaluate the target levels, incentive structures, and reward–penalty amounts associated with these incentives.

Finally, **voltage quality regulation** is concerned with a subset of factors that describe deviations of voltage from its ideal waveform. Such deviations can lead to damage or to the malfunctioning of customers' electrical equipment.

To date, the economic regulation of voltage quality has been relatively rare; voltage has typically been regulated using technical standards. Nonetheless, increasing penetrations of advanced metering infrastructure (AMI), and thus the increasing ability to gather distribution feeder-level data, is creating new opportunities for voltage quality regulation. At the same time, interest in voltage quality regulation has increased in recent years due to a proliferation of electronic devices that are extremely sensitive to voltage quality. Voltage quality also lends itself to all three of the regulatory elements discussed at the beginning of this section: performance metrics (and reporting requirements), minimum performance targets, and reward and penalty schemes.

The types of data required for voltage quality regulation are more technical and more difficult to obtain than data for the other two types of service quality regulation. Voltage supplied to customers has a number of characters—including frequency, magnitude, waveform, and symmetry of three-phase voltage—each of which exhibits variations during the normal operation of an electrical system (CENELEC 1999). Deviations of these characteristics from their nominal values are called voltage disturbances.

Due to these complexities and the large costs associated with voltage measurement at every connection point in a network, the regulation of voltage quality has typically been achieved via technical standards and bilateral arrangements between affected customers and distribution utilities. The use of outcome-based performance incentives for the regulation of voltage quality has been relatively rare. The few cases that we are aware of are in Latin America. In Argentina, for example, economic regulation for power quality indicators was introduced in the 1990s following the privatization of distribution companies in Greater Buenos Aires. Other Latin American countries introduced similar schemes following the privatization of publicly-owned national utilities. Moreover, several jurisdictions have implemented outcome-based performance standards that rely on "non-technical" metrics of voltage quality to enforce voltage quality regulation. These include, for example, customer waiting time before receiving a response following a voltage-related request.

Despite the complexities associated with voltage quality regulation, increased AMI deployment will create new opportunities for this area of regulation. The widespread deployment of distributed power electronics on distribution feeders will enable new volt and var (volt-ampere reactive) optimization (VVO) capabilities that were previously not available. Specifically, these new capabilities will make it possible to regulate voltage in real time and within a much narrower band than is currently required under most existing technical standards. New VVO capabilities can contribute to significant reductions in technical losses, improvements in energy efficiency, and load reduction, for example via conservation voltage reduction (CVR), whereby voltage levels are kept at or very close to the minimum safe voltage level. It is plausible that output-based incentives for CVR performance could begin to emerge in jurisdictions with the requisite technical capabilities.

5.3.1.3 Minimization of energy losses

Electrical losses are defined as "the difference between the amount of electricity entering the transmission system and the aggregated consumption registered at end-user meter points" (ERGEG 2008) and are comprised of technical or physical losses (i.e., those that occur as a result of heat and noise as electricity passes through network components) and non-technical or commercial losses (i.e., electricity consumption that is not metered such as consumption at distribution substations, energy theft, non-metered consumption, and metering errors). This section only considers technical or physical losses, which are the most common type of losses in most developed countries.

Technical electricity losses in distribution networks increase costs for customers and may increase CO_2 emissions (and emissions of other pollutants from power plants).[27] In a vertically integrated setting (in which the utility is the owner/operator of power plants, transmission, and distribution networks), the utility has an incentive to reduce technical losses to a level at which the cost of further loss reduction is equal to the marginal benefit of that reduction (where benefit is defined as the upstream

27 Of course, the degree to which technical losses increase CO_2 emissions depends on the carbon intensity of the electricity generation fleet.

savings associated with increased power plant efficiency). However, in jurisdictions with competitive retail markets, electricity distribution companies do not have an adequate incentive to reduce technical losses since they do not bear the full cost of these losses.[28] Network losses across their networks are in effect externalized, and additional regulation is required in these jurisdictions to incentivize utilities to account for the impact of network planning and operational decisions on the cost of losses born by network users.

Technical losses can be improved via specific operational and planning strategies (Arritt et al. 2009), therefore regulations that target such strategies may be desirable. However, many of the approaches to reducing losses through targeted operational and planning strategies present trade-offs with other regulatory objectives and incentives. For this reason it may be more appropriate to regulate losses via outcome-based performance incentives (Cossent 2013). As is the case with other forms of outcome-based performance regulation, the targeted level of losses is that at which the marginal benefit to society of reducing energy losses is equal (in magnitude) to the marginal cost to the utility of achieving this reduction. Nonetheless, the use of performance incentives for the regulation of technical losses is distinct from other forms of quality regulation (such as quality-of-service regulation) in that network users are not directly affected by the occurrence of technical energy losses (except for the increase in overall system costs). In practice, regulating the reduction of technical energy losses has proven to be challenging, and some of these challenges are likely to be exacerbated by increasing penetrations of DERs (Cossent 2013). Yet, DERs also may present important opportunities for reducing energy losses.

Increasing penetrations of DERs pose significant challenges for the regulation of technical loss reduction. DERs can significantly change power flow patterns and have a particularly large effect on variable technical losses.[29] Losses that arise as a result of DERs may thus perversely penalize the distribution utility for losses that the utility did not cause (Cossent 2013). There have been proposals to compensate network utilities for costs that are incurred by the utility but caused by distributed generation, including losses (de Joode et al. 2009). However, in addition to this type of compensation, it will also be important to incorporate the effect of DERs on technical losses when determining the methodologies used to assess the reference value of losses. Richer descriptions of how this might be accomplished and how it has been done in practice can be found in Cossent et al. (2009) and Cossent (2013).

Increasing penetrations of DERs also pose challenges for determining a metric that could be used in an outcome-based performance incentive for technical loss reduction. There are two obvious metrics that could be used for measuring, reporting, and regulating losses: an absolute value expressed in units of energy or a percentage expressing losses as a proportion of total energy injected into the distribution network or total energy consumed. To make the metric comparable across services areas, it is preferable to use the latter metric. However, the presence of DERs makes the use of this metric potentially problematic. For example, consider a scenario in which the reference values of losses are determined as a percentage of energy consumed (i.e., using the latter metric), and in which there is a high level of distributed generation (DG) in the system that has the effect of reducing net consumption and total variable losses in the system. In this case, due to the presence of fixed (transformer) losses that are unaffected by DG, the distribution utility would be penalized if DG loss reductions were higher (in percentage terms) than the reference loss value (Shaw et al. 2010). Moreover, the higher the proportion of fixed losses in the network, the more likely it is that this would occur. One solution to this potential problem would be to meter electricity consumption and generation separately. Another possible solution would be to use the first of the two metrics introduced above (i.e., an absolute value expressed in units of energy).

28 If the network utility is also the default or primary retailer in its service territory, it has at least some incentive to reduce losses in order to minimize the costs of retail supply for its customers. However, if other competitive retailers use the utility's distribution networks, at least some of the cost of losses is externalized for the utility, which can distort incentives for efficient management of network losses.

29 Variable technical losses are those that occur in electrical wires and are distinct from fixed losses that occur in transformer cores. Variable technical losses can range between 66 percent and 75 percent of total distribution losses (KEMA 2009; Ofgem 2003).

On the other hand, DG and DERs may also present opportunities for energy loss reduction. For instance, DG could contribute to energy loss reduction if it were sited close to consumption points. Demand-response technologies could also contribute to energy loss reductions, although the significance of this contribution may depend on the price elasticity of the electricity demand of customers that are participating in the demand-response program (e.g., Shaw et al. 2009 and Venkatesan et al. 2012). The use of ICT-enabled smart grid technologies may also contribute to technical loss reductions, since these technologies will assist network utilities in real-time network monitoring and automation and in the dispatch of DERs that contribute to loss reduction. However, it is likely that a significant share of the loss reductions that are enabled by smart grid technologies could be achieved as a byproduct of technology deployment that occurs via quality of supply regulation (Cossent 2013). With respect to electric vehicles (EVs), there is ample evidence to suggest that high penetrations of EVs are very likely to contribute to increased losses (Clement-Nyns et al. 2010; Fernández et al. 2011; and Peças Lopes et al. 2011, 2009a, 2009b).

Due to the challenges associated with regulating energy loss reductions, and the fact that these challenges will only be exacerbated by increasing penetrations of DERs, regulators must tread extremely carefully with respect to the design of performance-based incentives. Incorporating the effect of DERs on technical losses into the methodologies used to determine the reference value of losses will prove critical in the design of any incentive mechanism aimed at loss reduction.

5.3.1.4 New frontiers for outcome-based performance regulation

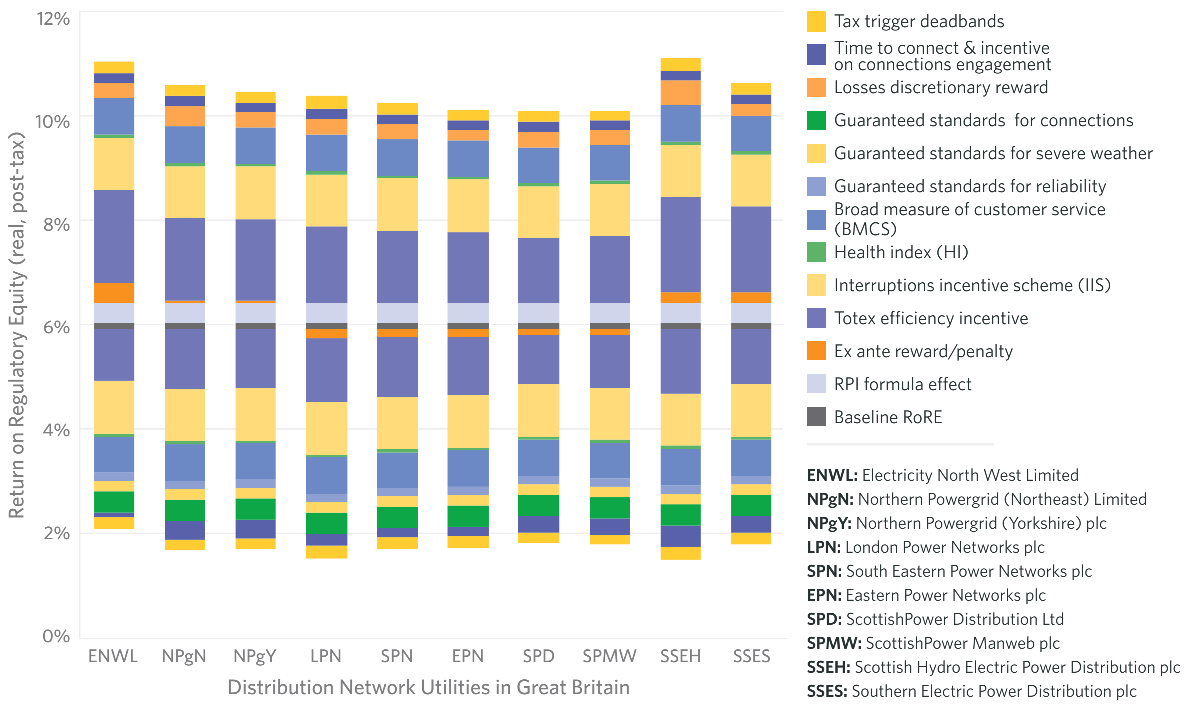
The preceding paragraphs have discussed some of the best practices with respect to outcome-based regulation in four critical performance areas: commercial quality, continuity of supply, voltage quality (which together comprise quality of service), and energy loss reduction. In each of these areas, the implementation of outcome-based performance incentives has helped steer network utilities in directions that are beneficial for network customers and that have led to safer, more reliable, and more efficient networks. It is also very likely that the use of outcome-based incentives for the improvement of these services has contributed to the creation of new knowledge, technologies, and processes that will play an important role in the transition to actively managed electricity networks.

These facts have led some authorities to move toward regulatory frameworks in which outcome-based performance incentives play a major role. Perhaps the best example is in the United Kingdom, where in 2011 the UK regulator (Ofgem) launched a new regulatory framework—titled "Revenue = Incentives + Innovation + Outputs" (RIIO) —in which outcome-based performance incentives play a central role in determining network utility profits.

Figure 5.8 shows the relative strength of incentives associated with the core remuneration framework—and with additional outcome-based performance incentives—for distribution utilities in the United Kingdom under RIIO. The vertical axis is the rate of return on regulated equity (ROR), and the horizontal axis shows different UK distribution companies. As Figure 5.8 shows, the allowed ROR of network utilities is linked to core efficiency incentives (the "RPI formula effect," "*ex ante* reward/penalty," and the "TOTEX efficiency incentive") as well as to a variety of important outcomes: improving network reliability ("interruptions incentive scheme," "guaranteed standards for reliability," "guaranteed standards for severe weather"), improving asset health ("health index")[30], improving customer service ("broad measure of customer service"), improving connection times and network user

30 Performance incentives may be put in place to ensure that the health of distribution network assets remains sufficiently high and the risk of network asset failure remains sufficiently low. The United Kingdom is a good example of a jurisdiction that has implemented this type of incentive. In the first regulatory period under the United Kingdom's new RIIO framework (RIIO-ED1), the UK regulator introduced an asset "risk index" (RI) that is a composite of an asset "health index" (HI) and an asset "criticality index" (CI). The RI is assessed on an increasing scale of 1 to 5 (Ofgem 2013e). UK distribution utilities are required to demonstrate how their allowed investments will contribute to network asset risk mitigation, and a specific "risk score" is assigned to each target a utility must meet. Under this scheme, if a distribution utility achieves its risk mitigation performance target, it will not receive a penalty or reward at the end of RIIO-ED1. On the other hand, if a distribution utility fails to meet its risk mitigation target and is not reasonably justified in underperforming, it will be subject to a penalty of 2.5 percent of the avoided costs that are associated with under-delivery. Likewise, if a utility exceeds its target, it will be rewarded 2.5 percent of the incremental costs associated with over-delivery. Notably, in the United Kingdom, ensuring adequately low levels of asset failure risk is considered secondary to the primary objectives of safety (public and workplace safety) and network reliability (Ofgem 2013e).

Figure 5.8: Range of Possible Returns on Investment for Distribution Utilities in Great Britain



Source: Ofgem (2014)

engagement ("guaranteed standards for connections," "time to connect and incentive on connections engagement"), and reducing electricity losses ("losses discretionary reward").

Figure 5.8 reveals that success or failure in meeting minimum targets for certain outcomes can lead to an increase or decrease (respectively) of 2 to 3 percentage points in the network utility's allowed rate of return on regulated equity, comparable to the possible gains and losses associated with efficiency incentives that are part of the core remuneration framework (Ofgem 2014).

As suggested in **Box 5.3**, regulators must recognize that achieving a specific performance outcome often increases a company's OPEX and/or CAPEX efficiency; they therefore must link outcome performance incentives (which fall outside the core remuneration framework) to cost efficiency incentives in the core remuneration framework. The United Kingdom, which has moved toward a more outcome-based approach to regulation, has done a good job of coupling outcome-based performance incentives and cost efficiency incentives.

During the *ex ante* establishment of future costs, Ofgem carefully considers the impact that achieving outcomes might have on these costs, and network utilities are unable to claim certain cost efficiency gains if they do not also successfully meet performance targets.

The United Kingdom is not the only jurisdiction that envisions outcome-based performance incentives playing an important role in the transition to actively managed electricity networks. In 2014, New York launched Reforming the Energy Vision (REV), an ambitious policy and regulatory initiative aimed at decarbonizing the state's energy sector. A major goal of the REV initiative is the transformation of the electric power system from a passively managed one to an actively managed system in which DER generators participate in the operation and planning of networks. Efforts are under way to completely rethink the regulation of distribution utilities to align revenue streams with the goals of active system management.

One of the mechanisms by which New York's regulatory authority seeks to align the incentives of network utilities with the goals of the REV initiative is by developing a set of outcome-based performance incentives (which it has termed "earnings adjustment mechanisms" or EAMs) (NYDPS 2016). In particular, regulatory staff have proposed outcome-based incentives that would apply to five performance areas seen as critical to achieving the objectives of the REV initiative: peak load reduction, energy efficiency, customer engagement (i.e., the education of, engagement with, and provision of data to customers), affordability (i.e., the promotion of low-income customer participation in DERs and a reduction in the number of terminations and arrearages), and interconnection (i.e., improvements in the speed and affordability of distributed generation units). The regulatory mechanisms that would incentivize performance improvements in these areas are explicitly temporary and would be updated and reassessed regularly (NYDPS 2016).

Notably, some of the performance areas that the regulatory commission has decided to focus on (namely peak load reduction and energy efficiency) are largely beyond the control of the distribution utility, since events and actions external to the utility will affect how these performance areas change. It is partly for this reason that regulators decided to link these performance incentives to financial rewards but not to financial penalties. Distribution utilities will be remunerated if performance improves but will not be penalized if performance degrades. This is a significant departure from the design of outcome-based performance incentives in other jurisdictions, in which incentives are linked to outcomes that can be fully or mostly attributed to distribution utility performance.

5.3.2 Explicit incentives for long-term innovation

Finally, increased uncertainty about the evolution of network needs, cost drivers, and opportunities will intensify the need for long-term innovation, including expanded investment in demonstration projects that produce technological learning and dissemination of knowledge among network utilities.[31] Since it is unclear how networks will evolve, it is also uncertain what technological solutions will lead to the greatest levels of productive efficiency in the medium to long term (i.e., lengths of time that are typically longer than the regulatory period). The technologies and systems that will be most efficient for facilitating active network management in distribution networks with high penetrations of DERs are simply not known with precision today. Therefore, there is a need for greater investment in demonstration projects and applied research and development (NAS 2016).

One of the most straightforward regulatory approaches for achieving increased levels of innovation is to offer explicit financial incentives, outside of the core remuneration framework, for network utilities to spearhead demonstration projects. The most common way to do this is via so-called "input-based" financial incentives, whereby demonstration projects are capitalized and included in the regulated asset base (in a cost of service regulatory context). It is also possible to incentivize demonstration projects via output-based financial incentives such as those discussed in **Section 5.3.1**. However, this is not recommended, since the outcomes of such projects are inherently uncertain. Nonetheless, as was discussed in Section 5.3.1, outcome-based performance incentives will inevitably incentivize some additional near-term innovation. This portion of the report presents case studies of three jurisdictions that have had success promoting innovation in electricity networks via input-based regulatory mechanisms. The case studies illuminate best practices from Europe and the United States and may provide guidance to regulatory authorities considering the adoption of incentives for innovation.

Incentives for longer-term innovation are needed to accelerate investment in applied R&D and demonstration projects that can provide lessons about the capabilities of novel technologies that may be risky or may have long-term payback periods.

31 Here we refer to activities focused on supporting potential solutions that may not be commercially ready within one regulatory period, or that are extremely uncertain with respect to performance or returns. The objective is for utilities to become consistent adopters and integrators of novel solutions.

CASE STUDY 1: UNITED KINGDOM

The United Kingdom was one of the earliest jurisdictions to explicitly embed significant incentives for innovation into its regulatory framework. By 2004, two new mechanisms that created dedicated funds for RD&D projects had been established (Bauknecht 2011; Lockwood 2016; Müller 2012), mainly on the level of the distributed system operator (DSO). One was the Innovation Funding Incentive (IFI), which covered "all aspects of distribution system asset management" (Ofgem 2004). The IFI was capped at 0.5 percent of allowed revenue and was available on a use-it-or-lose-it basis. Ofgem (the UK regulator) allowed 90 percent of the costs of IFI projects to be recovered in the first year of the price control, but this tapered off through the period to 70 percent in the fifth year, in order to incentivize early uptake. The second mechanism was Registered Power Zones (RPZs)—a scheme aimed at demonstrating innovative solutions for the connection of new DG on sections of the network. Distribution utilities were allowed additional revenue for each kilowatt of DG connected, capped at a total of £500,000 per company per year. These mechanisms were small relative to the size of total network utility expenditures, but they created a new niche of activity that subsequently led to a much larger scheme.

Once launched, the IFI quickly produced a response. Spending by network utilities under the IFI increased from around £2 million in 2003–2004 to around £12 million in 2008 (Jamasb and Pollitt 2008). By contrast, RPZs were less successful, with only a small number of projects materializing during the price control period (Bolton and Foxon 2011; Woodman and Baker 2008).

In 2010, a new mechanism for RD&D projects was introduced in the fifth distribution price control review: the Low Carbon Network Fund (LCNF). This was a competitive mechanism that allowed distribution companies to bid for up to £500 million over five years (Ofgem 2010), equivalent to 2.3 percent of allowed revenue, an order of magnitude larger than the IFI and a very substantial increase relative to levels of innovation spending a decade earlier. There were two tiers of funding, one allowing distribution companies to recover most of the costs of smaller projects in allowed revenue, and another for larger projects in the form of a competitive fund of £64 million a year. Tier 2 funding requires companies to cooperate with ICT firms, suppliers, generators, and consumers in projects to encourage cross-sector collaboration and innovative partnerships. Essentially the same structure for RD&D funding is continuing into the first RIIO price control period (2015–2023), with the name of the program changed to the Electricity Network Innovation Competition (NIC).

The LCNF and NIC have led to a step change in levels of RD&D activity by UK utilities as well as to much larger-scale demonstration projects. Moreover, the LCNF and NIC require the sharing of knowledge gained from trials between utilities and other market participants and have led to a website and an annual conference, which is now a major event that attracts several hundred participants. In this manner, these competitions are accelerating important learning, knowledge-sharing, and networking processes that can speed up the spread of best practices and successful innovations. There is some evidence that the competitions have also had a significant effect on distribution company thinking and culture, albeit to varying degrees among companies. The LCNF and NIC have required distribution companies to work together with suppliers, ICT firms, renewable generators, and consumers on concrete demonstration projects. These funding opportunities have engaged board-level interest from network companies in the smart grid agenda and have made distribution companies aware of potential new commercial relationships and opportunities (eg., in demand response).

In addition to the NIC, RIIO has also led to the creation of a Network Innovation Allowance (NIA) that every distribution company receives (each utility receives a fixed and predetermined amount over the course of the regulatory period). The central purpose of the NIA is to fund smaller technical, commercial, or operational projects that are directly related to each company's network and that have the potential to deliver financial benefits to the company and its customers. Distribution companies are required to publish information about the projects (as well as frequent updates) on a website known as the Smarter Network Portal (Energy Networks Association 2016).

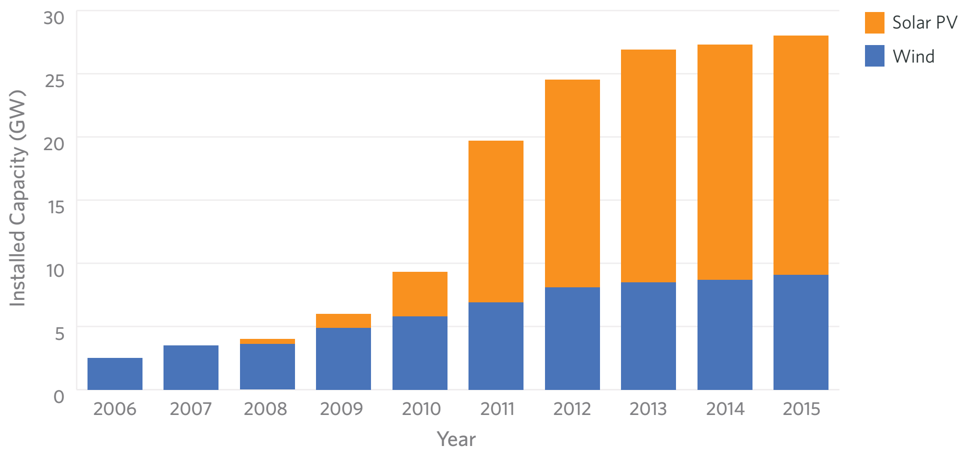
CASE STUDY 2: ITALY

Italy has also been particularly active in promoting long-term innovation in distribution networks. The rate and scale at which DG has increased in Italy is astounding (**Figure 5.9**). In 2006 the installed capacity of DG was 2.5 gigawatts (GW), comprised entirely of wind turbines. By 2015, installed DG capacity had ballooned to 28 GW, two-thirds of which was distributed solar PV. The Italian electricity system, with a maximum demand of approximately 53 GW and a minimum demand of approximately 20 GW, has so far accommodated this growth via traditional network investments, i.e., a "fit and forget" approach. Nonetheless, the challenges facing distribution network utilities in Italy have increased, and this is reflected in the percent of time that primary (high- and medium-voltage) substations are experiencing reverse power flows. As **Figure 5.10** shows, the number of primary substations facing reverse power flows more than 1 percent of the time and more than 5 percent of the time in the Enel distribution network[32] dramatically increased over a period of five years (2010 to 2014).

As a result of these challenges, regulators in Italy have implemented a number of new regulations, including output-based regulatory incentives (as described in **Section 5.3.1**), new standards and codes (e.g., network codes that specify frequency variation tolerance requirements for DG units), and input-based incentives for innovation and pilot projects (Lo Schiavo et al. 2013).

32 Enel is the largest distribution utility in Italy, serving 86 percent of the country's total electricity demand in 2011.

Figure 5.9: Installed Capacity of Intermittent Renewable Energy Sources in Italy, 2006–2015



Source: Lo Schiavo (2016)

For this discussion we focus solely on innovation-related incentives. The first of these established in Italy, in 2009, was to commission a research study to investigate the hosting capacity of DG in Italian medium-voltage (MV) distribution networks, where the majority (75 percent) of solar PV was deployed (AEEG 2009). The research found that MV networks in Italy have a very large hosting capacity at the nodal level—85 percent of buses in the sample were individually able to host at least 3 megawatts of distributed generation—but also found that there were certain problems that would limit system-wide DG hosting capacity (Delfanti et al. 2010). Another outcome of the research was the identification of an indicator of network "activeness" called "reverse power-flow time" (RPT), the percentage of time in a year during which power flows from medium to high voltage.

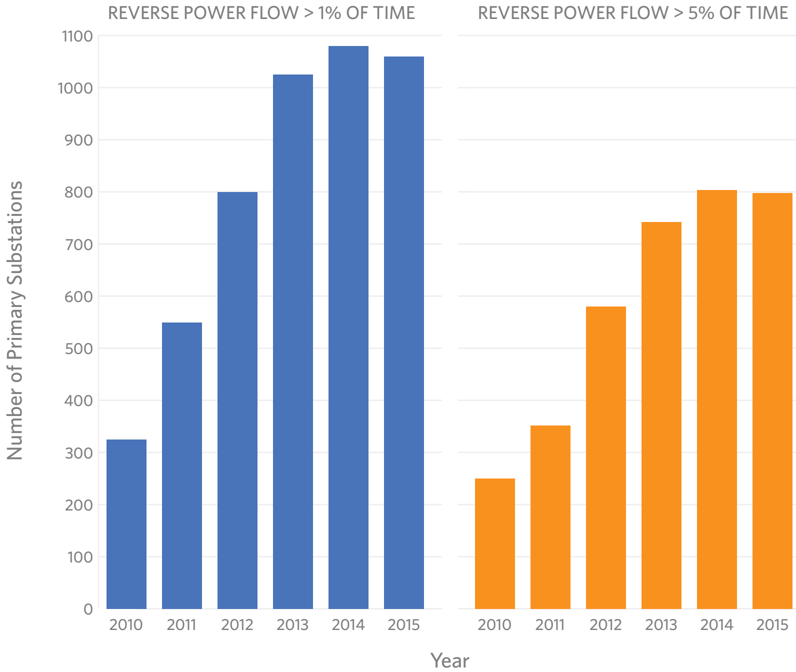
The results of this research motivated the Italian regulatory authority to solicit competitive offers for innovative demonstration projects with the primary objective of reducing RPT and thereby improving the DG hosting capacity of distribution networks (AEEG 2010).[33] A committee of experts conducted a selection process on behalf of the regulatory authority, and selected demonstration projects were capitalized and benefited from an additional 2 percent weighted average cost of capital (WACC) on top of the standard rate of return for a period of 12 years.[34] The committee assessed the different proposals using a variety of parameters, including qualitative indicators and technical scores, the cost of the project, and an indicator known as Psmart. Psmart is defined as the increase in DG production that can be connected to the grid without undermining safe operating conditions (voltage, current, frequency), as a result of the pilot project.

Eight pilot projects were deployed under this regulatory framework, and as a result, several functionalities pertaining to active network management have been tested. In particular, the pilot projects have shed light on the interactions between distribution system operators (DSOs) and transmission system operators (TSOs), smart voltage control, active power modulation, anti-islanding functionalities, fast MV fault isolation, and electricity storage. The Italian regulatory authority is currently focused on using lessons from the pilot projects to develop outcome-based incentives related to the transition to active system management. A detailed analysis of various active management functionalities was undertaken in 2015, and it was proposed that two functionalities—TSO-DSO data exchange and smart voltage control in MV networks—should be considered for outcome-based regulation (AEEG 2015a, 2015b).

33 Reverse power flows can contribute to overvoltage in electricity networks and therefore can pose significant operational, safety, and security risks. The DG hosting capacity of a distribution network is constrained by these risks. Therefore innovations that can reduce reverse power flows will increase the DG hosting capacity of distribution networks.

34 To participate in the selection process, demonstration projects had to meet three main requirements: (1) distribution networks in which the projects were deployed had to show an RPT of at least 1 percent annually; (2) projects had to focus on development and deployment (i.e., not basic research); and (3) projects had to meet standard protocols for any communication applications involving network users.

Figure 5.10: Number of Primary Substations in Italy Experiencing Reverse Power Flows for More than 1 Percent and More than 5 Percent of the Year



Source: Lo Schiavo (2016)

CASE STUDY 3: NEW YORK

In the United States, New York State has been at the vanguard of developing regulations to promote long-term innovation. Under the REV initiative, the goal is for electricity distribution companies to become "distribution system platform" (DSP) providers, meaning the majority of their revenues will be generated from providing market or platform services rather than from expenditures on capital equipment and from sales volumes (NYDPS 2016). The intent is to change the regulated utility business such that distribution utilities facilitate DER participation in electricity markets and such that network utilities are agnostic between traditional network upgrades and DER-based "non-wires" solutions.

As part of this transition, distribution network utilities in New York have recently submitted distribution system implementation plans (DSIPs), documents intended to demonstrate (to the regulatory commission) how the network utilities intend to move toward actively managed networks while ensuring that traditional utility objectives (e.g., reliability, safety, etc.) continue to be met. Each utility's DSIP documents the utility's plans over a five-year period, and there will be a formal DSIP filing every two years. As one of the components of a DSIP, each utility is required to explain how it "expects to maximize option value of the distribution system for customers through better planning, system operations and management and vastly scaled integration of DER—without making unnecessary investments" (NYDPS 2015b). Distribution utilities are also charged with including information about relevant current and near-term RD&D pilot projects. Utilities are encouraged to include pilot projects in their DSIPs because the regulatory authority recognizes that data collected from REV demonstration projects will "assist the process of integrating DER resources into system planning, development, and operations on a system- and statewide scale" (NYDPS 2015b).

As a result, all the major distribution utilities in New York have submitted proposals for pilot projects, and there are currently (at the time of writing) more than 10 such projects under way.[35] Examples include CenHub Marketplace, a project in which a distribution utility and a technology company have teamed up to build an online portal for energy products and services that provides customers with personalized recommendations and offers enhanced data analytics; Clean Virtual Power Plant, a project in which a utility is partnering with DER providers that bundle solar with storage to aggregate electrical supply and thus serve as a virtual power plant to provide grid services on clear days; Flexible Interconnect Capacity Solution, a project in which a utility is partnering with a technology company to offer a new, less costly, and faster way for customers and third parties to connect large DG

35 A full list of demonstration projects can be found at the DPS website: www.dps.ny.gov.

projects to the grid by providing an "infrastructure as a service" alternative to traditional interconnection; and others.

SUMMARY OF CASE STUDIES

The case studies described above shed light on some of the novel ways in which regulatory authorities in jurisdictions in Europe and the United States have created input-based incentives and competitive rewards to promote long-term innovation in electricity distribution networks. By enabling distribution utilities to invest in RD&D projects, these regulations are helping to generate new technological knowledge pertaining to the integration of DERs into network operation and management. Much of the knowledge that emerges from these pilot projects will contribute to the deployment of new technologies, systems, and processes that will be critical to the transition to active system management, thereby leading to much greater efficiencies in the medium to long term. We therefore strongly encourage regulatory authorities to establish appropriate incentives for greater utility investment in RD&D projects, as well as mechanisms for knowledge sharing so that the knowledge gleaned from the projects is widely disseminated across utilities, regulators, and technology providers.

5.3.3 Cybersecurity preparedness

Widespread connection of DERs will increase digital complexity and attack surfaces, and therefore require more intensive cybersecurity protection. A multi-pronged approach to cybersecurity preparedness is required, including enhanced information sharing on cyber threats and possible responses. Although cybersecurity regulations and standards have been adopted, they will need to be updated and enhanced. Moreover, because cyber attack strategies will evolve, complying with regulations will only be a starting point; it will be necessary to improve cyber defense best practices continually. The power grid is an inherently open system. From a cybersecurity standpoint, the population of relevant cyber and physical devices is extremely large. Every electric vehicle, building energy management system, smart thermostat, and electrical device with a connection to the Internet and every interconnected infrastructure—gas, communications, water—could potentially be used in an attack on the stability of the power grid. A challenge is to have the capacity to operate, maintain, and recover a system of subsystems and devices that will never be fully protected from cyber attacks. In the longer term, relevant issues that need to be addressed include cloud security, machine-to-machine information sharing, advanced cybersecurity technologies, the possibility of adding chief information security officers to distribution companies, outcome-based regulation to avoid prolonged outages, and international approaches to cybersecurity.

Although some DER systems may have been deployed without robust cybersecurity protections, future cybersecurity protection standards and privacy concerns can become a barrier to the deployment of DERs until clarification is provided regarding what cybersecurity standards will be required and how much meeting those standards will cost DER systems and DER aggregators. In addition, evolving privacy concerns and regulations need to be taken into account alongside deployment of aggregated demand-response and other DER systems that gather large amounts of private and corporate energy-use information. The privacy and ownership of users' data must be protected by DER owners and operators using appropriate systems and safeguards.

5.3.3.1 Current approaches for information sharing in Europe

Information sharing can make DERs more reliable because cybersecurity and interface problems—and potential solutions—can be communicated more quickly to mitigate widespread issues. Information sharing will play a key role in the development and deployment of cybersecurity standards and solutions in the European Union. Fortunately, certain initiatives are already under way. For example, a network of national Computer Security Incident Response Teams (CSIRTs) has been established under the NIS Directive. The CSIRTs network will be composed of representatives of the member states' CSIRTs and CERT-EU (Computer Emergency Response Team for European Union institutions, agencies, and bodies). The European Union Agency for Network and Information Security (ENISA) will actively support cooperation among CSIRTs. Among other tasks, the CSIRT network will support the exchange of information on CSIRTs' services, operations, and cooperation capabilities. A detailed list of the CSIRT network's tasks is provided in article 12(3) of the NIS Directive (European Union 2016).

Several other initiatives have been launched in Europe at the national and international levels to share information on risks, vulnerabilities, and threats. For example, the European Commission has established the Energy Expert Cyber Security Platform and the Thematic Network on Critical Energy Infrastructure Protection, and it has financed the Distributed Energy Security Knowledge (DENSEK) project. One of the deliverables of the DENSEK project was the establishment of the European Energy - Information Sharing & Analysis Centre, which plays an important role in cybersecurity information sharing in the energy sector in Europe. Over time, cybersecurity information sharing could be augmented and coordinated with the support of other organizations, such as the European Network of Transmission System Operators for Electricity.

In the United Kingdom, cybersecurity information sharing is achieved via a government–industry consortium known as the Energy Emergencies Executive Committee Cyber Security Task Group, as well as via the Cyber Security Information Sharing Partnership, an initiative that allows users to share real-time cyber threat information.

5.3.3.2 Expanding cybersecurity regulation in the United States and Europe

To improve cybersecurity in the United States, state regulatory commissions could work together to establish standards applicable to distribution utilities and modeled on the existing NERC Critical Infrastructure Protection standards. This would remove the uncertainty surrounding what cybersecurity standards will be applied to DERs. Using an approach similar to NERC, authorities could announce regulations that apply to distribution utilities (and large distributed generators) along with a date that the regulations would come into effect. State public utility commissions would be responsible for new critical infrastructure protection guidelines applicable to distribution utilities. These cybersecurity standards would apply to DERs owned by both utilities and non-utilities. Harmonizing regulatory standards at the transmission, distribution, and end-user levels would contribute to economies of scale and reduce the cost of cybersecurity preparedness.

In Europe, the European Commission encourages member states to make the most of NIS coordination mechanisms. Building on those, the Commission will propose how to enhance cross-border cooperation in the case of a major cyber-incident. Given the speed with which the cybersecurity landscape is evolving, the Commission will also evaluate ENISA, which will possibly lead to the adoption of a new mandate. Common coordinated standards are likely to be more effective at reducing risks for involved member states than piecemeal guidelines that vary by state.

5.3.3.3 Longer term approaches for building cybersecurity resilience

To address the long-term cybersecurity challenges faced by network utilities, we recommend a multifaceted approach.

First, cybersecurity regulations or standards should consider network and market operations that are performed in the cloud in order to make data interfaces consistent, clear, and secure. Performing network and market operations in the cloud with consolidated information would require enhanced security, robust monitoring, and artificial-intelligence and machine-learning technologies to monitor out-of-norm activities.

Individual DER operators and aggregators would provide secure, authenticated data to a private or hybrid public/private cloud data repository and network utilities with appropriate cybersecurity protections, and market operators would assess the veracity of the information. In this way, operations would be consolidated; instead of relying on separate systems with different protocols and control software, this plan would ensure all data and controls pass through major portals. If executed well, this type of configuration could lead to more secure and efficient operations and would enable DER providers to interact with and utilize utility data. However, the integration of DER communications with network operations could create risks of wider outages in the case of attack; therefore, it is recommended that DERs be required to have adequate built-in security measures before connecting to the grid and to cloud systems that control network and market operations.

Second, trusted capabilities for sharing threat information from machine to machine will need to be deployed in the United States and Europe to enable timely monitoring of and response to cyber incidents that can develop within minutes. However, high-speed monitoring and information transfer significantly increases the need to ensure that threat data and response plans are coming from verified official sources.

Third, enhanced cybersecurity technologies are needed to address the evolution of more sophisticated cyber attacks and cyber technology and reduce the likelihood of successful attacks on DERs and the grid. Such technologies include advanced encryption, programs to ensure information integrity, artificial intelligence and behavioral analysis, moving target techniques to randomize cyber system components, and cyber-secure architecture (including processor memory safety, unauthorized application denial, and compiler validation). With the layering of industrial Internet of Things technologies on top of existing infrastructure (including decades-old programmable logic controllers and remote terminal units), security technologies will be needed for disaggregated single-purpose devices. Model-based engineering and virtualized risk simulators can be used to evaluate the security and performance of utility systems.

Fourth, each network utility should have a chief security officer—an executive responsible for addressing key cybersecurity, physical security, and privacy challenges; for ensuring that regulations are understood and followed within the organization; and for ensuring that DERs and control systems are safely connected to the Internet. This executive can ensure that cybersecurity and physical controls are not insurmountable barriers to connecting DER systems to the grid. Corporate boards of directors also need to understand and evaluate cybersecurity-related risks, governance, and operations; review results of security audits; evaluate cyber-event prevention and recovery plans and operations; act when a cybersecurity breach is known; and evaluate cyber-insurance options. Companies need to establish formal cybersecurity risk management processes in which senior managers and board members are engaged. All parts of the corporation, from the board to the general workforce, need training to understand and address these risks.

Fifth, regulations to avoid and mitigate cyber attacks could be considered to augment baseline cybersecurity standards. To achieve increased levels of innovation with the goal of avoiding prolonged outages over a broad geographic area due to a cyber attack, explicit financial incentives—outside of the core remuneration framework—could be provided for network utilities to spearhead pilot projects. An "input-based" financial incentive, whereby pilot projects are capitalized and included in the regulated asset base (in a cost of service regulatory context) could work to mitigate holistic security risk and minimize the chance and impact of high-cost, low-probability events such as a widespread cyber attack. These regulations will bring enhanced security concerns into utility reliability and disaster recovery planning. In addition, because regulators face an extremely difficult task when asked on a state-by-state basis to evaluate the reasonableness and effectiveness of utility security expenditures, there could be requirements for utilities to maintain private cyber insurance and to provide external validation of utility cyber governance. If a successful cyber attack caused a network utility to face a penalty, there would be a penalty cap, such as that used for utility supply, that limits the penalty to a percent of the company's remuneration so as not to send a company into bankruptcy.

Finally, because of the critical importance of electricity, including electricity from DERs, to economic stability and to the health and welfare of citizens worldwide, an international approach to cybersecurity is recommended as part of a comprehensive strategy. A global international agreement would be difficult to achieve; however, an intergovernmental initiative on cybersecurity would not be without precedent. Examples include the Budapest Convention on Cybercrime, the Chemical Weapons Convention, the Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and Their Disposal, and the Montreal Protocol on Substances that Deplete the Ozone Layer. To achieve international support, an organization such as the United Nations, working with nation states, worldwide electric utilities, Internet service providers, and groups such as the Organization for Security and Co-operation in Europe and international law enforcement organizations, could initiate this effort and determine how to monitor any agreement that might be reached.

References

ACCC. 2012a. "Benchmarking Opex and Capex in Energy Networks." Australian Competition and Consumer Commission and Australian Energy Regulator working paper No. 6. May 2012.

ACCC. 2012b. "Regulatory Practices in Other Countries: Benchmarking Opex and Capex in Energy Networks." Australian Competition and Consumer Commission and Australian Energy Regulator. May 2012.

AEEG. 2009. Monitoraggio dello sviluppo degli impianti di generazione distribuita in Italia per l'anno 2006 ed analisi dei possibili effetti della generazione distribuita sul sistema elettrico nazionale. Autorita per l'Energia Elettrica e il Gas. Regulatory decision ARG/elt 25/09 (only in Italian).

AEEG. 2010. Procedura e criteri di selezione degli investimenti ammessi al trattamento incentivante di cui al comma 11.4 lettera d) dell'Allegato A alla deliberazione dell'Autorit a per l'energia elettrica e il gas 29 dicembre 2007, n. 348/07. Autorita per l'Energia Elettrica e il Gas. Regulatory decision ARG (only in Italian).

AEEG. 2015a. Criteri per la definizione delle tariffe per l'erogazione dei servizi di trasmissione, distribuzione e misura dell'energia elettrica nel quinto periodo regolatorio. Autorita per l'Energia Elettrica e il Gas. Regulatory consultation 544/2015/R/eel (only in Italian).

AEEG. 2015b. Smart distribution system: promozione selettiva degli investimenti nei sistemi innovativi di distribuzione di energia elettrica. Autorita per l'Energia Elettrica e il Gas. Regulatory consultation 255/2015/R/eel (only in Italian).

Aggarwal, S. and H. Harvey. 2013. "Rethinking Policy to Deliver a Clean Energy Future." *The Electricity Journal* 26: 7–22. doi:10.1016/j.tej.2013.08.009.

Armstrong, C.M., S. Cowan, and J. Vickers. 1994. *Regulatory Reform: Economic Analysis and British Experience.* Cambridge, MA: MIT Press.

Averch, H. and L. L. Johnson. 1962. "Behavior of the Firm Under Regulatory Constraint." *American Economic Review* 52: 1052–1069.

Bailey, E.E. 1974. "Innovation and Regulation." *Journal of Public Economics* 3: 285–295.

Bauknecht, D. 2011. "Incentive Regulation and Network Innovations." EUI working paper RSCAS 2011/02. Florence, Italy: European University Institute.

Beesley, M.E. and S.C. Littlechild. 1989. "The Regulation of Privatized Monopolies in the United Kingdom." *The RAND Journal of Economics* 20(3): 454. doi:10.2307/2555582.

Bolton, R. and T. Foxon. 2011. "Governing Infrastructure Networks for a Low Carbon Economy: Co-evolution of Technologies and Institutions in UK Electricity Distribution Networks." *Competition and Regulation in Network Industries* 12: 2–27.

Bourreau, M. and P. Dogan. 2001. "Regulation and Innovation in the Telecommunications Industry." *Telecommunications Policy* 25(3): 167–184.

CENELEC. 1999. "*Voltage Characteristics of Electricity Supplied by Public Distribution Systems.*" European Standard EN 50160.

Clemenz, G. 1991. "Optimal Price-Cap Regulation." *The Journal of Industrial Economics* 39(4): 391–408.

Conchado, A. and P. Linares. 2012. "The Economic Impact of Demand-Response Programs on Power Systems: A Survey of the State of the Art." In A. Sorokin et al., eds., *Handbook of Networks in Power Systems.* Berlin, Germany: Springer. dx.doi.org/10.1007/978-3-642-23193-3_11.

Cossent, R. 2013. "Economic Regulation of Distribution System Operators and Its Adaptation to the Penetration of Distributed Energy Resources and Smart Grid Technologies." PhD thesis. Madrid, Spain: Universidad Pontificia Comillas.

Cossent, R. and T. Gómez. 2013. "Implementing Incentive Compatible Menus of Contracts to Regulate Electricity Distribution Investments." *Utilities Policy* 27: 28–38. dx.doi.org/10.1016/j.jup.2013.09.002.

Cossent, R., T. Gómez, and P. Frías. 2009. "Towards a Future with Large Penetration of Distributed Generation: Is the Current Regulation of Electricity Distribution Ready? Regulatory Recommendations under a European Perspective." *Energy Policy* 37(3): 1145–1155. dx.doi.org/10.1016/j.enpol.2008.11.011.

Cossent, R., L. Olmos, T. Gómez, C. Mateo, and P. Frías. 2011. "Mitigating the Impact of Distributed Generation Advanced Response Options." *International Transactions on Electrical Energy Systems* 21(6): 1869–1888.

CPUC. 2015. Order Instituting Rulemaking Regarding Policies, Procedures and Rules for Development of Distribution Resources Plans Pursuant to Public Utilities Code Section 769. California Public Utilities Commission.

Crouch, M. 2006. "Investment under RPI-X: Practical Experience with an Incentive Compatible Approach in the GB Electricity Distribution Sector." *Utilities Policy* 14(4): 240–244. dx.doi.org/10.1016/j.jup.2006.05.005.

Delfanti, M., D. Falabretti, M. Merlo, G. Monfredini, and V. Olivieri. 2010. "Dispersed Generation in MV Networks: Performance of Anti-islanding Protections." ICHQP 2010–14th International Conference on Harmonics and Quality of Power. Bergamo, Italy. doi:10.1109/ICHQP.2010.5625446.

Denholm, P., J. Jorgenson, T. Jenkin, M. Hummon, D. Palchak, B. Kirby, and M. O'Malley. 2013. "The Value of Energy Storage for Grid Applications." NREL Technical Report (No. TP-6A20-58465). Golden, CO: National Renewable Energy Laboratory.

Energy Networks Association. 2016. Smarter Networks Portal (website). www.smarternetworks.org. Accessed October 15, 2016.

ERGEG. 2008. "Treatment of Losses by Network Operators." Position paper for public consultation No. E08-ENM-04-03. The European Regulators' Group for Electricity and Gas.

Eurelectric. 2013. "Active Distribution System Management: A Key Tool for the Smooth Integration of Distributed Generation." Brussels, Belgium.

European Union. 2016a. Directive (EU) 2016/1148: Concerning Measures for a High Common Level of Security of Networks and Information Systems across the Union. Brussels, Belgium. July 6, 2016.

European Union. 2016b. Regulation (EO) 2016/679: On the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation). Brussels Belgium. April 27, 2016.

Fernández, L. P., T. Gómez, R. Cossent, C.M. Domingo, and P. Frías. 2011. "Assessment of the Impact of Plug-in Electric Vehicles on Distribution Networks." *IEEE Transactions on Power Systems* 26(1): 206–213. dx.doi.org/10.1109/TPWRS.2010.2049133.

Fumagalli, E., L. Lo Schiavo, F. Delestre. 2007. *Service Quality Regulation in Electricity Distribution and Retail*. Berlin, Germany; Heidelberg, Germany; New York, NY: Springer. doi:10.1007/978-3-540-73444-4.

Gómez, T. 2013. "Electricity Distribution." In I.J. Pérez-Arriaga, ed., *Regulation of the Power Sector*. London, United Kingdom: Springer-Verlag. dx.doi.org/10.1007/978-1-4471-5034-3_5.

Gómez, T., I. Momber, M.R. Abbad, and A. Sánchez Miralles. 2011. "Regulatory Framework and Business Models for Charging Plug-in Electric Vehicles: Infrastructure, Agents, and Commercial Relationships." *Energy Policy* 39(10): 6360–6375. dx.doi.org/10.1016/j.enpol.2011.07.037.

Hurley, D., P. Peterson, and M. Whited. 2013. "Demand Response as a Power System Resource: Program Designs, Performance, and Lessons Learned in the United States." Cambridge, MA: Synapse Energy Economics Inc. www.synapse-energy.com/Downloads/SynapseReport.2013-03.RAP.US-Demand-Response.12-080.pdf.

IEC. 2016. "Resilience and Security Recommendations for Power Systems with Distributed Energy Resources (DER) Cyber-Physical Systems." IEC TR 62351-12:2016, International Electrotechnical Commission. April 2016.

Jamasb, T., P. Nillesen, and M. Pollitt. 2003. "Gaming the Regulator: A Survey." *The Electricity Journal* 16(10): 68–80. dx.doi.org/10.1016/j.tej.2003.09.008.

Jamasb, T., P. Nillesen, and M. Pollitt. 2004. "Strategic Behaviour under Regulatory Benchmarking." *Energy Economics* 26(5): 825–843. dx.doi.org/10.1016/j.eneco.2004.04.029.

Jamasb, T. and M. Pollitt. 2003. "International Benchmarking and Regulation: An Application to European Electricity Distribution Utilities." *Energy Policy* 31(15):1609–1622. doi:10.1016/ S0301-4215(02)00226-4.

Jamasb, T. and M. Pollitt. 2008. "Liberalisation and R&D in Network Industries: The Case of the Electricity Industry." *Research Policy* 37: 995–1008. doi:10.1016/j.respol.2008.04.010.

Jenkins, J.D. 2014. "Economic Regulation of Electricity Distribution Utilities Under High Penetration of Distributed Energy Resources: Applying an Incentive Compatible Menu of Contracts, Reference Network Model and Uncertainty Mechanisms." Master's thesis. Cambridge, MA: Massachusetts Institute of Technology. June 2014.

Jenkins, J.D. and I.J. Pérez-Arriaga. 2014. "The Remuneration Challenge: New Solutions for the Regulation of Electricity Distribution Utilities Under High Penetrations of Distributed Energy Resources and Smart Grid Technologies." MIT Center for Energy and Environmental Policy Research, working paper WP 2014-005. Cambridge, MA: Massachusetts Institute of Technology. September 2014.

Jenkins, J.D., and I.J. Pérez-Arriaga. 2017. "Improved Regulatory Approaches for the Remuneration of Electricity Distribution Utilities with High Penetrations of Distributed Energy Resources." *The Energy Journal* 38(3): 89-117.

Joskow, P.L. 1989. "Regulatory Failure, Regulatory Reform, and Structural Change in the Electrical Power Industry." *Brookings Papers on Economic Activity*.

Joskow, P.L. 2014. "Incentive Regulation in Theory and Practice: Electricity Distribution and Transmission Networks." In N.L. Rose, ed., *Economic Regulation and Its Reform: What Have We Learned*. Chicago, IL: University of Chicago Press.

Kahn, A.E., T.J. Tardiff, and D.L. Weisman. 1999. "The Telecommunications Act at Three Years: An Economic Evaluation of Its Implementation by the Federal Communications Commission." *Information Economics and Policy*.

Kind, P. 2013. "Disruptive Challenges: Financial Implications and Strategic Responses to a Changing Retail Electric Business." Washington, DC: Edison Electric Institute. Retrieved from www.eei.org/ourissues/finance/Documents/ disruptivechallenges.pdf.

Laffont, J-J. and J. Tirole. 1993. *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA: MIT Press.

Littlechild, S.C. 2006. "Beyond Regulation." CWPE 0616. EPRG 0516. www.eprg.group.cam.ac.uk/wp-content/uploads/2014/01/eprg0516.pdf.

Lo Schiavo, L. 2016. "Incentive Regulation of Smart Distribution Systems in Italy: from Input-based Demonstration to Output-based Deployment." Presentation given to the Utility of the Future Study team at MIT on May 6, 2016.

Lo Schiavo, L., M. Delfanti, E. Fumagalli, and V. Olivieri. 2013. "Changing the Regulation for Regulating the Change: Innovation-driven Regulatory Developments for Smart Grids, Smart Metering, and e-Mobility in Italy." *Energy Policy* 57: 506–517. doi:10.1016/j.enpol.2013.02.022.

Lockwood, M. 2016. "Creating Protective Space for Innovation in Electricity Distribution Networks in Great Britain: The Politics of Institutional Change." *Environmental Innovation and Societal Transitions* 18: 111–127. doi:10.1016/j.eist.2015.05.007.

Lowry, M.N. and T. Woolf. 2016. "Performance-based Regulation in a High Distributed Energy Resource Future." *Lawrence Berkeley National Laboratory, Future Electric Regulation Series Report No 3*. LBNL-1004130. Berkeley, CA. January 2016.

MA DPU. 2014. Order 12-76-B: Investigation by the Department of Public Utilities on Its Own Motion into Modernization of the Electric Grid. Boston, MA: Massachusetts Department of Public Utilities. June 12, 2014.

Malkin, D. and P. A. Centolella. 2013. "Results-Based Regulation: A Modern Approach to Modernize the Grid." Atlanta, GA: GE Digital Energy.

Meeus, L., Saguan, M., Glachant, J.-M., and R. Belmans. 2010. "Smart Regulation for Smart Grids." EUI working paper RSCAS 2010/45. Florence, Italy: European University Institute.

Meeus, L. and Saguan, M. 2011. "Innovating Grid Regulation to Regulate Grid Innovation: From the Orkney Isles to Kriegers Flak via Italy." *Renewable Energy* 36(6): 1761–1765. dx.doi.org/10.1016/j.renene.2010.12.005.

Momber, I., T. Gómez, and L. Söder. 2013. "PEV Fleet Scheduling with Electricity Market and Grid Signals." IEEE 10th International Conference on the European Energy Market: 1–8.

Müller, C. 2012. "Advancing Regulation with Respect to Dynamic Efficient Network Investments: Insights from the United Kingdom and Italy." *Competition and Regulation in Network Industries* 13: 256–272.

NAS. 2016. *The Power of Change: Innovation for Development and Deployment of Increasingly Clean Electric Power Technologies.* Washington, DC: National Academy of Sciences.

NERC. 2016. "Critical Infrastructure Protection Standards." Washington, DC: North American Electric Reliability Corporation.

NSF. 2011. NSF Industrial Research and Development Information System (Historical Data 1953-2007). National Science Foundation. www.nsf.gov/statistics/iris/. Accessed August 25, 2016.

NYDPS. 2015a. Order Adopting Regulatory Policy Framework and Implementation Plan. Case 14-M-0101: Reforming the Energy Vision. Albany, NY: State of New York Public Service Commission, Department of Public Service. February 2015.

NYDPS. 2015b. Staff proposal: Distributed System Implementation Plan Guidance. Albany, NY: State of New York Public Service Commission, Department of Public Service.

NYDPS. 2016a. Proceeding on Motion of the Commission in Regard to Reforming the Energy Vision. Albany, NY: State of New York Public Service Commission, Department of Public Service.

NYDPS. 2016b. Order Adopting a Ratemaking and Utility Revenue Model Framework. Case 14-M-0101: Reforming the Energy Vision. Albany, NY: State of New York Public Service Commission, Department of Public Service. May 2016.

Ofgem. 2004. "Electricity Distribution Price Control Review—Final Proposals." London, United Kingdom: Office of Gas and Electricity Markets.

Ofgem. 2009. "Electricity Distribution Price Control Review Methodology and Initial Results Paper." London, United Kingdom: Office of Gas and Electricity Markets.

Ofgem. 2010a. "Handbook for Implementing the RIIO Model." London, United Kingdom: Office of Gas and Electricity Markets.

Ofgem. 2010b. "Low Carbon Networks Fund Governance Document v.1." London, United Kingdom: Office of Gas and Electricity Markets.

Ofgem. 2010c. "RIIO: A New Way to Regulate Energy Networks." London, United Kingdom: Office of Gas and Electricity Markets.

Ofgem. 2013a. "Strategy Decision for the RIIO-ED1 Electricity Distribution Price Control: Uncertainty Mechanisms." London, United Kingdom: Office of Gas and Electricity Markets.

Ofgem. 2013b. "Strategy Decision for the RIIO-ED1 Electricity distribution Price Control: Outputs, Incentives and Innovation." London, United Kingdom: Office of Gas and Electricity Markets.

Ofgem. 2013c. "Strategy Decision for the RIIO-ED1 Electricity Distribution Price Control: Financial Issues." London, United Kingdom: Office of Gas and Electricity Markets.

Ofgem. 2013d. "Strategy Decision for the RIIO-ED1 Electricity Distribution Price Control: Business Plans and Proportionate Treatment." London, United Kingdom: Office of Gas and Electricity Markets.

Ofgem. 2013e. "Strategy Decision for the RIIO-ED1 Electricity Distribution Price Control: Reliability and Safety." London, United Kingdom: Office of Gas and Electricity Markets.

Ofgem. 2014. "Final Determinations for the Slow-Track Electricity Distribution Companies." London, United Kingdom: Office of Gas and Electricity Markets.

Olmos, L., R. Cossent, T. Gómez, C. Mateo, J. de Joode, M. Scheepers, F. Nieuwenhout, J. Poot, M. Bongaerts, D. Trebolle, B. Doersam, S. Bofinger, U. Cali, and N. Gerhardt. 2009. "Case Studies of System Costs of Distribution Areas." IMPROGRES No. WP 4, Deliverable 5. Petten, Netherlands: Improvement of the Social Optimal Outcome of Market Integration of DG/RES in European Electricity Markets Project.

Pérez-Arriaga, I. J. and A. Bharatkumar. 2014. "A Framework for Redesigning Distribution Network Use of System Charges Under High Penetration of Distributed Energy Resources: New Principles for New Problems." CEEPR Working Paper No. 2014-006. Cambridge MA: MIT Center for Energy and Environmental Policy Research.

Poudineh, R. and T. Jamasb. 2014. "Distributed Generation, Storage, Demand Response and Energy Efficiency as Alternatives to Grid Capacity Enhancement." *Energy Policy* 67: 222–231. dx.doi.org/10.1016/j.enpol.2013.11.073.

Pudjianto, D., M. Aunedi, P. Djapic, and G. Strbac. 2014. "Whole-Systems Assessment of the Value of Energy Storage in Low-Carbon Electricity Systems." *IEEE Transactions on Smart Grid* 5(2): 1098–1109. dx.doi.org/10.1109/TSG.2013.2282039.

Rivier, J., T. Gómez. 2000. "A Conceptual Framework for Power Quality Regulation." *Proceedings of the IEEE Ninth International Conference on Harmonics and Quality of Power*: 469–474.

Sappington, D. 2005. "Regulating Service Quality: A Survey." *Journal of Regulatory Economics* 27: 123–154.

Schisler, K., T. Sick, and K. Brief. 2008. "The Role of Demand Response in Ancillary Service Markets." IEEE Transmission and Distribution Conference and Exposition, 2008: 1–3. dx.doi.org/10.1109/tdc.2008.4517087.

Schmalensee, R. 1989. "Good Regulatory Regimes." *The RAND Journal of Economics* 20(3): 417–436. dx.doi.org/10.2307/2555580.

Strbac, G., M. Aunedi, D. Pudjianto, P. Djapic, F. Teng, A. Sturt, D. Jackravut, R. Sansom, V. Yufit, and N. Brandon. 2012. *Strategic Assessment of the Role and Value of Energy Storage Systems in the UK Low Carbon Energy Future.* London, United Kingdom: Carbon Trust. Retrieved from www.carbontrust.com/media/129310/energy-storage-systems-role-value-strategic-assessment.pdf.

Trebolle, D., T. Gómez, R. Cossent, and P. Frías (2010). "Distribution Planning with Reliability Options for Distributed Generation." *Electric Power Systems Research* 80(2): 222–229. dx.doi.org/10.1016/j.epsr.2009.09.004.

Vergara, C., I. Perez-Arriaga, C. Mateo, and P. Frías. 2014. "Estimating the Aggregate Impact of Distributed Photovoltaic Generation over Distribution Networks." IIT-Comillas Working Paper No. IIT-14-034A. Madrid, Spain: Institute for Research in Technology (IIT). Comillas University.

Weisman, D. 2005. "Price Regulation and Quality." *Information Economics and Policy* 17: 165–174.

Wirth, H. 2016. "Recent Facts about Photovoltaics in Germany." Frieberg, Germany: Fraunhofer Institute for Solar Energy Systems. April, 22 2016.

Woodman, B. and P. Baker. 2008. "Regulatory Frameworks for Decentralised Energy." *Energy Policy* 36: 4527–4531.

PART 2: A FRAMEWORK FOR AN EFFICIENT AND EVOLVING POWER SYSTEM

06

Restructuring Revisited: Electricity Industry Structure in a More Distributed Future

6.1 Restructuring Revisited

After nearly a century of exclusive monopoly provision of electricity services, a wave of restructuring began in the 1980s and spread across much of the world in the 1990s and 2000s. With the apparent exhaustion of economies of unit scale in power generation due to the emergence of smaller-scale electricity generation technologies, such as gas-fired plants, wind turbines, and combined heat and power generation, policy makers began to reconsider the natural monopoly characteristics of bulk power generation, and in some cases, even retail supply (Joskow and Schmalensee 1983). This era saw considerable debate over the proper structure of the electricity sector and the appropriate roles and responsibilities, in particular, of electricity network utilities, system operators, and market operators (FERC 1996, 1999; European Commission 2003; Millán 2006).

Until rather recently, the debate about restructuring has been mainly framed within the boundaries of the bulk power system (i.e., wholesale generators and high-voltage transmission networks). Today, the growth and future potential of a variety of distributed energy resources—including solar photovoltaics (PV), other forms of distributed generation, and electrical and thermal energy storage devices—and more price-responsive and flexible electricity demand has sparked a new wave of debate (CPUC 2014; CEER 2015a; Corneli and Kihm 2015; de Martini and Kristov 2015; European Commission 2009, 2010; Pérez-Arriaga, et al. 2013; NYDPS 2014, 2015a, 2015b).

Market platforms, network providers, and system operators perform three critical functions that sit at the center of all transactions in electricity markets. Properly assigning responsibilities for these core functions is thus critical to an efficient, well-functioning electricity sector and for establishing a level playing field for competitive provision of electricity services by both traditional generators and network providers and new businesses harnessing DERs.

While there are many close parallels to restructuring debates of the past, contemporary challenges require extending discussions about electricity industry structure "all the way to the bottom"—to encompass not only distribution network owners and operators, but also end consumers, aggregators (including retailers), and new competitive business models that harness distributed energy resources (DERs). The first round of electricity restructuring focused on establishing competitive markets for large-scale power generators and hinged on the role of transmission utilities and bulk power system operators and their relation to competitive market segments. This new round of debate revolves around the proper role of electricity distribution utilities, the potential need for a distribution system operator (DSO), the establishment of market structures or platforms that enable the provision of multiple electricity services by DERs of all kinds, and the emerging role of DER aggregators and other new DER business models.

This chapter considers the question of electricity industry structure from four different perspectives. First and most fundamentally, the chapter offers different views on how to best allocate the core responsibilities of market platform, system operation, network ownership, and data management at the distribution and retail level of the power system in order to enable an efficient and competitive sector. The new and more complex relationship between "bulk" system operators at the transmission level and distribution system operators is examined next. The third section of this chapter analyzes how the aggregation of DERs may create value, presently and in a foreseeable future. The chapter concludes with a mostly descriptive presentation of existing and proposed non-traditional business models for harnessing DERs, some of which may play a significant role in providing electricity services in the future.

6.2 Electricity Industry Organization in a More Distributed Future

6.2.1 Three crucial electricity industry functions

In electricity markets, the trading and provision of services occurs at three levels. First, producers and consumers of electricity (or their representatives) buy and sell energy from one another, often taking advantage of market platforms, such as power exchanges or centralized markets run by system operators. Second, a physical transmission and distribution network must be built and maintained to deliver energy from generators to consumers. Third and finally, system operators need to plan network development, procure certain technical services from market agents, and coordinate the dispatch of these market agents and network assets to reliably and efficiently operate electricity systems.

Market platforms, network providers, and system operators thus perform three critical functions that sit at the center of all transactions in electricity markets. This makes properly assigning responsibilities for these core functions critical to an efficient, well-functioning electricity sector and critical to establishing a level playing field for competitive provision of electricity services by traditional generators and network providers, and by new businesses that harness DERs. Considering the appropriate industry structure and assignment of responsibilities in a more distributed future thus requires familiarity with these three core functions, together with an understanding of which agents and actors are best equipped to carry out each function.

6.2.1.1 Market platforms

Market platforms provide the financial infrastructure to match buyers and sellers of electricity services. While bilateral transactions are possible (and occur in electricity markets), centralized market platforms often improve the efficiency, competitiveness, and transparency of commodities markets, including markets for electricity (Mansur and White 2012). Examples of market platforms in other sectors include the Chicago and New York Mercantile Exchanges, New York Stock Exchange, and the Royal Exchange in London (now the London Stock Exchange), each of which emerged over time to facilitate multilateral trading and settlement of commodities or equities. In markets where standard, fungible products or services can be established, multilateral market platforms minimize transaction costs, increase liquidity, minimize information asymmetries, improve transparency, and allow parties to exchange risks, while creating a centralized infrastructure for settling transactions.[1] As in some other sectors, electricity market platforms are often implemented as a series of temporally linked markets, including futures contracts, various derivatives and hedges, and spot purchases.[2]

6.2.1.2 Network providers

Electricity services cannot be provided without the physical network infrastructure that connects buyers and sellers of these services. The network provider function thus involves building and maintaining the transmission and distribution networks over which electricity services are delivered. In the long term, this function involves procuring and constructing network assets. In the medium term, the network provider must properly maintain and steward the health of network assets. Finally, the network must also be repaired and services restored in a timely fashion after unscheduled outages due to weather or other events or from failures of network components.

6.2.1.3 System operators

While electricity market platforms often closely mirror the structure and function of other commodities markets, several peculiar characteristics differentiate the electricity sector and give rise to the sector's third core function, that of system operator. System operators plan network development and operate the system to ensure that the schedule of electricity production and consumption that results from market trading and associated physical power flows across electricity networks is feasible, taking into account thermal and voltage constraints in all parts of the network. Where flows are expected to be infeasible, interventions are required, such as re-dispatching generators or loads in the short term or expanding the network in the long term. In addition, power inputs and withdrawals must be balanced at all times to maintain the synchronous frequency of alternating current power systems within tolerable ranges (e.g., 60 Hertz [Hz] in much of the Americas and parts of Asia; 50 Hz elsewhere, with generally less than 0.036 Hz and no greater than 0.5 Hz deviation at any time). A combination of uncertainty in demand (and, with the growth of variable renewable energy sources, supply uncertainty as well), unplanned "contingencies" or failures in generators or transmission and distribution lines, and technical constraints that affect how quickly generators can start up and change output levels makes balancing system frequency at all times a challenging and critical task.

The core tasks of the system operator typically give rise to markets for several "security-related" services, including several classes of operating reserves and voltage control. Likewise, the need to ensure the physical security of real-time electricity production and consumption schedules leads to one of two options. In the first option, system operators take responsibility for day-ahead and intraday market platform responsibilities, co-optimizing market dispatch and system operation via security-constrained unit commitment and economic dispatch algorithms

1 To a significant extent, multilateral markets can also reduce transaction costs and improve market efficiency for products or services that are not entirely standardized and fungible (e.g., AirBnB, Ebay, Amazon). However, in such markets an additional curation or search function is required to match differentiated products or services to buyer preferences. In the electricity context, two types of platform markets may emerge: a multilateral commodity platform market for electricity products as well as a services platform that links customers with service providers, such as energy management companies or clean energy suppliers, that match additional customer preferences. This chapter focuses primarily on markets for commodity electricity services.

2 Financial traders and brokers also participate in these markets, connecting market agents, offering financial hedges (e.g., contracts-for-differences), and transacting in futures and spot markets (e.g., as "virtual" bidders in day-ahead markets), where they generally increase liquidity and reduce information asymmetry. While virtual bidders can improve market outcomes, in some cases, virtual bidders may arbitrage against inconsistencies between forward and spot markets in multi-market settlement processes, thereby profiting without adding to improved system performance. See Parsons et al. (2015).

(common in the United States).[3] In the second option, system operators create new markets to re-dispatch and balance supply and demand when market platform transactions would lead to infeasible power flows or the real-time physical position of generators or loads differ from market schedules (common in European electricity markets).[4] In practice, the line between "security-related" markets and "regular" markets is somewhat fluid and may shift over time.[5] In addition, uncertainty about load growth over time; long lead times to build new generators; lack of trust in the stability of policy or regulatory conditions; risk aversion on the part of generators, regulators, and consumers; and lack of full market participation by demand may give rise to a market for "firm capacity" contracts (e.g., capacity remuneration mechanisms), another security-related market for which the system operator may be responsible.

Finally, electricity network assets have traditionally been operated "statically," with network components sized to accommodate peak demands (a "fit and forget" strategy) and lines only occasionally switched in or out of service. More "active" network management practices are emerging, however, and could yield considerable improvements in cost and performance, including dynamically switching lines (Hedman et al. 2010, 2011). System operators may therefore also be responsible for network topology control.

6.2.2 Lessons from restructuring bulk power systems

Although debates over industry structure at the bulk power system level have not fully ended (and perhaps never will), decades of experience in restructuring have highlighted a number of necessary structural and regulatory practices that underpin healthy electricity markets and operations. Today's debates, over the proper role of distribution system operators, owners, and users and over proposed market platforms to serve them, can thus benefit from more than 30 years of experience with restructuring in bulk generation and transmission networks.

6.2.2.1 Market platforms alone are insufficient

First, as the restructuring of wholesale electricity markets unfolded, it quickly became apparent that establishing a competitive market platform (or "power pool") alone was insufficient to ensure competitive electricity generation and supply. In practice, both the system operator and network provider functions can significantly affect the ability of market agents to buy or sell electricity services. The entities responsible for system operation and network provision are in a position to exercise "vertical foreclosure"[6] by planning, building, or operating the system in a manner that negatively impacts the ability of certain upstream suppliers to access downstream customers or vice versa. As the US Federal Energy Regulatory Commission (FERC) observed:

> "*When utilities control monopoly transmission facilities and also have power marketing interests, they have poor incentives to provide equal quality transmission service to their power marketing competitors. …The inherent characteristics of monopolists make it inevitable that they will act in their own self-interest to the detriment of others by refusing transmission and/or providing inferior transmission to competitors in the bulk power markets to favor their own generation, and it is our duty to eradicate unduly discriminatory practices.*" (FERC 1999)

The European Commission (EC) has concurred, noting:

> "*Without effective separation of networks from activities of generation and supply (effective unbundling), there is an inherent risk of discrimination not only in the operation of the network but also in the incentives for vertically integrated undertakings to invest adequately in their networks.*"[7] (European Commission 2009)

3 For discussion and definition of the roles of regional transmission organizations (RTOs) or independent transmission system operators (ISOs) in the United States, see FERC (1999).

4 For discussion and definition of the roles of transmission system operators (TSOs) in Europe, see Slot et al. (2015).

5 The definition of security-related markets usually reflects some technical constraints on the speed at which generators or demand can respond to security-related market prices or control signals and practical concerns about transaction costs, all of which can change over time.

6 Vertical foreclosure, a term from the industrial organization literature, encompasses a variety of practices whereby a vertically-integrated firm exercises its position in the market to provide discriminatory preference to affiliated upstream or downstream firms, either by excluding or negatively impacting the ability of downstream buyers to access upstream suppliers that compete with the firm or by excluding or negatively impacting the ability of upstream suppliers to access downstream customers that do business with the firm (Tirole 1988).

7 In this context, vertical integration refers to transmission or distribution network utilities that are vertically integrated with generation and/or retail supply functions.

Establishing a competitive market platform is insufficient to ensure competitive electricity generation and supply. In practice, both the system operator and network provider functions can have significant impacts on the ability of market agents to buy or sell electricity services.

Network planning and expansion must be performed in a transparent and impartial manner under appropriate regulatory oversight to ensure that network expansion and interconnection processes do not privilege certain agents over others. Less immediately obvious but perhaps equally critical, actions by system operators to resolve security-related concerns (e.g., re-dispatch of loads and generators) or procure security-related system services (e.g., reserves) can also advantage some market participants and penalize others.

6.2.2.2 Independence of system operation and network planning are critical

Given these concerns, regulators in the United States, Europe, and elsewhere have established requirements for "functional and legal unbundling" of transmission network and system operation functions from competitive market activities (FERC 1996; European Commission 2003). Functional and legal unbundling typically entailed open access to transmission systems by all generators, establishment of transparent and non-discriminatory transmission tariffs (including purchases of ancillary services), creation of a functionally separate legal entity responsible for transmission business activities, and other mandates to prevent transmission network utilities from preferentially advantaging certain generators or load-serving entities (such as subsidiaries of the network utility's owner) or foreclosing market access by certain agents.

However, after initial restructuring efforts generated several years of experience, both FERC and the EC ultimately concluded that open access tariffs and functional and legal unbundling were generally insufficient to establish effective incentives and a well-functioning, competitive wholesale marketplace. As FERC has argued:

> *"Functional unbundling does not change the incentives of vertically integrated utilities to use their transmission assets to favor their own generation, but instead attempts to reduce the ability of utilities to act on those incentives."* (FERC 1999)

In light of this view, FERC (1999) concluded that a "better structured market where operational control and responsibility for the transmission system is structurally separated from the merchant generation function of owners of transmission" would best facilitate competition and eliminate incentives for vertical foreclosure in bulk power system activities. The EC reached a similar decision in 2009, finding that the "rules on legal and functional unbundling of [transmission system operators]" it established in 2003 "have not led to effective unbundling." This led the EC to call for "structural separation" that would "remove the incentive for vertically integrated undertakings to discriminate against competitors as regards access to the network, as regards access to commercially relevant information and as regards investments in the network" (European Commission 2009, 2010).

6.2.3 New challenges in distribution systems

While many lessons can be learned from prominent models for industry structure at the level of the bulk power system, these models cannot be simply "copied and pasted" into the distribution system context. A number of differences are important when considering how to translate these lessons to the realities of distribution system structure and operation. This section explores some of the salient differences between bulk and distribution systems and key challenges for restructuring in the distribution context.

6.2.3.1 Distribution networks are orders of magnitude more complex than transmission networks

While transmission network owners may manage a few hundred to tens of thousands of lines, substations, and generators, many distribution networks incorporate hundreds of thousands or millions of network components and users. The order-of-magnitude greater complexity of distribution networks challenges both planning and active system operation. For example, while transmission asset investments tend to be large, discrete investments that facilitate long-term planning, competitive procurement processes, and transparent consideration of "non-wires" alternatives, distribution investments are smaller, more numerous, and often must be completed in a shorter time period. This greater complexity also increases the regulatory burden associated with ensuring that vertically integrated utilities do not exercise vertical foreclosure and consistently provide non-discriminatory access to distribution networks and markets. Light-handed regulation of distribution utilities may therefore require structural changes that mitigate incentives for discriminatory behavior and/or changes to utility remuneration that establish incentives for utilities to efficiently serve DERs and consistently exhaust cost-saving and performance-improving opportunities across both capital and operational expenditures.[8] Finally, the complexity of distribution networks makes it more difficult to decouple network planning from operation. This increases economies of scope between network ownership and system operation, as compared to transmission systems.

As experience with restructuring in the bulk power system has demonstrated, ensuring the independence of market platforms, network providers, and system operators at the distribution level is important to a competitive industry landscape.

6.2.3.2 Markets for distribution system services are local and may be illiquid

Markets for distribution system services may also be far less liquid than markets for bulk power system operation and wholesale markets. Bulk power system operators maintain frequency across large geographic areas. While network constraints can segment these markets, reserves, balancing, capacity, and other ancillary services are generally fungible commodities that can be procured from a wide range of generators or loads across a system operator's territory. In contrast, distribution system operators are concerned primarily with local voltage constraints and local network reliability and resiliency issues. Only a small number of service providers with assets in very specific locations may be able to provide these services, particularly in the operational time scale. Markets for DSO services may therefore prove illiquid in practice and may require further regulation and oversight to ensure that participants do not exercise market power. Alternatively, DSO services could be procured via longer-term contractual arrangements, which, by allowing new entrants to bid, would increase competition and mitigate the exercise of market power during periods of system stress. Here again, lessons can be learned from bulk power systems, where the critical and highly location-specific nature of black-start capabilities and "must-run reliability resources" necessary for local voltage stability led to the development of regulated, long-term contracts for these illiquid, location-specific services. In addition, market monitors may be required to identify competitiveness issues and recommend corrective actions on an ongoing basis.

8 See **Chapter 5** and Jenkins and Pérez-Arriaga (2017) for more on economic regulation of distribution network utilities.

6.2.3.3 Distribution network users receive price signals via regulated tariffs

Generators, retailers, and large consumers that participate in wholesale energy markets are generally exposed to competitively determined locational marginal prices (at the zonal or nodal level[9]). These prices deliver information about the time- and location-specific value of energy supply and demand, helping to coordinate efficient investment and operations at the bulk system level. Closer to real time, wholesale generators also receive and respond directly to operational dispatch signals from a centralized system operator—the transmission system operator (TSO) or independent system operator (ISO). In contrast, at least to date, energy consumers and most DERs connected at distribution voltage levels do not directly participate in wholesale markets and instead receive price signals via regulated or competitive network and supply tariffs. These tariffs generally convey much less detailed (and sometimes non-existent) information about the time- and location-specific value of energy consumption and production at the distribution level—something that must change if DERs and more price-responsive consumption is to compete on a level playing field with conventional generators and network services.[10] Going forward, it is likely that operational decisions for many DERs will be more loosely coordinated via price signals or other incentives, rather than directly dispatched, as in bulk power systems. The exception may be DERs that commit to provide specific services to system operators in advance (e.g., forward options), which may be directly dispatched or receive dispatch instructions via intermediaries such as aggregators.

6.2.3.4 DERs can supply services to both distribution and bulk power system operators and end consumers

Finally, DERs connected at distribution voltages can supply services to distribution system operators and bulk system operators as well as end consumers. For example, an energy storage device could provide voltage management or network capacity deferral to the DSO, firm capacity or operating reserves to the TSO/ISO, enhanced reliability to end consumers, or some combination of the above (although a given increment of capacity cannot be committed to provide more than one of these products at a time). This raises the importance of a comprehensive system of prices and charges for energy, network and generation capacity, and system operator services (see **Chapter 4**), which can ensure that resources capable of providing multiple services commit their capabilities to those services with the greatest value. In addition, markets and system operation at the bulk level and the distribution level must be carefully coordinated in real time to ensure that conflicting dispatch signals are resolved in an orderly fashion and system reliability is maintained at all levels, as discussed in **Section 6.3**. Again, while the context is different, parallels with bulk power systems persist: Just as various markets for ancillary services and energy must be carefully coordinated or co-optimized to ensure efficient provision of multiple services by bulk generators, so too must markets for DSO services, energy and ancillary services at the bulk level, and end-user demands on distributed resources be coordinated.

9 Note that, as discussed in **Chapters 4** and **7**, the application of locational marginal prices differs in different jurisdictions. For example, zonal prices are established in European electricity markets, while most US markets employ nodal LMPs.

10 See **Chapter 4** for more on an efficient system of prices and charges for electricity system users.

Table 6.1: Salient Differences Between Bulk and Distribution Systems

	TRANSMISSION NETWORK	BULK SYSTEM OPERATOR (TSO/ISO)	WHOLESALE MARKETS	BULK GENERATORS
BULK SYSTEM LEVEL	Hundreds to thousands of lines and substations. A few large, discrete investments needed annually, making network expansion decisions easily supervised and contestable. Meshed network structure.	Focused primarily on regulating frequency. Security-related services (balancing, reserves) can be procured as fungible commodities across wide geographic areas (exceptions include black start and voltage regulation). Operations can be relatively easily decoupled from maintenance decisions.	Fungible commodities markets with many participants and prominent use of locational marginal prices (at nodal or zonal level).	Can supply services to wholesale markets and TSO/ISO security-related markets (ancillary services).

	DISTRIBUTION NETWORK	DISTRIBUTION SYSTEM OPERATOR (DSO)	RETAIL TARIFFS	DERs AND LOADS
DISTRIBUTION SYSTEM LEVEL	Thousands to millions of lines, substations, transformers, and other assets. Many smaller investments needed on a continual basis, making network expansion decisions more difficult to supervise or contest. Meshed and radial network structure.	Focused primarily on regulating voltage and restoring outages. Additional complexity due to need to maintain phase balance, topology changes, volatility in voltage, power quality concerns, and in some cases less complete electronic information and visibility of network assets. Location-specific and potentially illiquid markets for security-related services (voltage control, congestion management, reliability). May exhibit significant economies of scope with network maintenance.	End consumers (and most DERs) pay or are compensated based on regulated or competitive retail tariffs that currently lack cost-reflective locational and temporal granularity.	Can supply services to wholesale markets and TSO/ISO security-related markets, as well as to DSO security-related markets and directly to end consumers.

6.2.4 Three options for structuring the roles and responsibilities of distribution system operators

Just as in transmission and bulk generation, market platforms, system operation, and network provision are all critical functions that facilitate competition and efficient operation of distribution networks and DERs in lower voltages. Here we present three potential models for assigning these core functions to different actors, focusing in particular on the different roles and responsibilities of DSOs.

6.2.4.1 Distribution network owner and operator

The first option parallels the TSOs that have emerged to manage the bulk power system in Europe and other jurisdictions. It combines the functions of distribution network provider, distribution system operator, and market platform for distribution services within a single utility. This distribution network owner/system operator (DNO/SO) would be responsible for creating and maintaining the physical distribution network, operating the distribution system (including actively managing distribution network assets and coordinating the dispatch of DERs to provide distribution network services), and establishing and operating any markets or procurement processes for DSO services (e.g., voltage control, congestion relief, enhanced resiliency).[11]

As distribution system services are highly location-specific and may be illiquid, the DNO/SO could lead competitive and transparent procurement processes to establish longer-term contracts for system services.[12] This would reveal the price of alternatives to traditional network investments and allow the utility to effectively unlock the potential of DERs to reduce total system costs or improve system performance. Eventually, if metering, communication, and computational capabilities are sufficient (CAISO 2015), the DNO/SO would have the capability to compute and communicate locational marginal prices within the distribution system (in close coordination with the local ISO/TSO) that capture the marginal cost of local voltage or thermal constraints, distribution losses, and transformer wear and tear, and the capability to communicate the location- and time-specific value of DER-provided services (Caramanis et al. 2016). In addition, cost-reflective network tariffs can convey the long-term cost of network expansion in different locations and ensure adequate recovery of network costs. Prices and charges in distribution systems and the benefits and trade-offs associated with greater degrees of spatial granularity in energy and network charges are discussed further in **Chapter 4**.

To facilitate equal competition between traditional network investments and contracts or payments for DER services, the DNO/SO should be regulated in a manner that (1) rewards utilities for reducing total expenditures and improving network quality of service and (2) equalizes incentives between operational and capital expenditures, as discussed in detail in **Chapter 5**. Alternative or complementary measures include conducting transparent "distribution resource plans" (similar to the integrated resource plans undertaken by traditionally regulated, vertically integrated utilities), mandating full consideration of "non-wires" alternatives to all network investments through regulation (CPUC 2014; NYDPS 2015a, 2015b), and undertaking fully transparent procurement processes or organized auctions for DER services. While these measures improve incentives for DSOs to actively engage DERs and loads to reduce network expenditures, they do not align the fundamental business incentives of the utility (as performance-based regulation does, for instance). In addition, effective oversight of so many individual distribution network assets and investments can involve significant regulatory costs.

As with TSOs, this version of the DSO would have to be sufficiently independent of competitive activities, both upstream and downstream (e.g., centralized and decentralized generation upstream, and retailing/aggregating and DER ownership downstream), to ensure

11 The markets established by the DNO/SO would be specifically for services used *within* the distribution system, as distinct from existing markets for ancillary services used in bulk power systems, which are already well established by bulk system operators. Note that it is possible to separate responsibility for market platform operation from the system operation, planning, and maintenance functions of the DNO/SO (we discuss a variation of this model below). Indeed, dividing the market platform role from the system operator and network provider roles closely parallels the bulk power system structure common in Europe, where market platforms are managed by power exchanges and system operation and network provision is managed by TSOs.

12 See **Section 6.3.3** for more on how this procurement process could be structured.

impartial planning and operation of distribution systems and impartial procurement of services from DERs or their aggregators. Indeed, since the 3rd Energy Package (European Commission 2009), European regulation has recognized the importance of independence in distribution system operation and established a minimum level of separation between distribution utilities and retailing and generation activities. As the Council of European Energy Regulators explains:

> *"To create a level playing field in retail energy markets, all competitive actors need to compete on the same terms. DSOs should facilitate this by acting as neutral market facilitators. This requires a sufficient level of unbundling between suppliers and associated DSOs. As energy networks are regulated monopolies, DSOs have exclusive access to all customers within their geographic network area. Without sufficient unbundling, this has the potential of disturbing competition in the market." (CEER 2016)*

The New York Public Service Commission (PSC) has voiced similar concern that "unrestricted utility participation in DER markets presents a risk of undermining markets more than a potential for accelerating market growth" (NYDPS 2015a). The PSC's 2015 order establishing utilities as "distribution system platform" companies (similar to the DNO/SO model outlined here) thus restricts these companies from owning distributed generation or storage assets except in limited circumstances.[13] "Markets will thrive best where there is both the perception and the reality of a level playing field," the PSC has argued, "and that is best accomplished by restricting the ability of utilities to participate" in DER markets (NYDPS 2015a).

As experience in bulk power systems demonstrates, structural unbundling is most effective at minimizing incentives for discriminatory behavior by the network owner/operator and facilitating competition and more light-handed regulation. Structural unbundling in this context would entail financial independence from any companies active in wholesale energy or ancillary service markets and from retailing or DER ownership or aggregation *within* the service territory of the DNO/SO.[14] As an alternative to structural unbundling, independence can be approximated by the establishment of legal unbundling and effective restrictions on exchange of information and coordination between the DSO and any other subsidiary or sister companies engaged in competitive activities within the DSO's service territory. However, legal unbundling entails a greater regulatory burden to be effective[15] and, as experience in the restructuring of bulk power systems has demonstrated, legal and functional unbundling measures are second-best alternatives to structural independence. This is likely to be especially true in more complex distribution networks, as discussed above.

This DNO/SO model offers advantages in terms of maintaining potentially significant economies of scope between network ownership, planning, and operation, since both network provider and system operation functions are the domain of the same utility. In addition, if properly regulated and separated from competitive activities, as discussed above, the DNO/SO becomes an impartial facilitator of much greater competition between conventional and distributed energy resources and a new customer for DER services wherever they provide cost-effective alternatives to traditional network investments. The primary challenge to implementing this approach is the level of industry restructuring and regulatory reform that may be necessary in some jurisdictions.

---

13 The PSC restricts utility ownership of DERs only in cases where:

"1) procurement of DER has been solicited to meet a system need, and a utility has demonstrated that competitive alternatives proposed by non-utility parties are clearly inadequate or more costly than a traditional utility infrastructure alternative;

2) a project consists of energy storage integrated in to distribution system architecture;

3) a project will enable low or moderate income residential customers to benefit from DER where markets are not likely to satisfy the need; or

4) a project is being sponsored for demonstration purposes." (NYDPS 2015a at p. 70)

14 If the DSO or its parent company owns subsidiaries engaged in competitive activities outside the DSO's service territory, there is no possibility for conflicts of interest in the DSO's planning, operation, or maintenance of the network. Therefore, this kind of activity could be permitted under the industry structure described here.

15 For example, under Directive 2009/72/EC, the European Commission currently only requires legal unbundling of DSOs and retailers while establishing several requirements to approximate functional independence. Similarly, the NYDPS warns "participation by utility affiliates [in DER markets] within the service territory [of the distribution company] does present the risk of discriminatory treatment by the utility." The NYDPS nevertheless permits affiliates of distribution utilities to be active in competitive retail supply markets as energy services companies (ESCOs). In 2015, the DPS concluded that "affiliate ownership [of DERs] may be allowed under a less stringent set of conditions than direct ownership," while also requiring various "protections to ensure that affiliates' participation in [distribution system] markets do not represent market power abuses" (NYDPS 2015a at p. 71).

6.2.4.2 Independent distribution system operator

As a second model, the role of the DSO could mirror that of the ISOs that have been established in the United States and elsewhere at the bulk system level. A new independent distribution system operator (IDSO) could have responsibility for planning and operating distribution systems across a given geographic region, but would not own network assets (Friedrichsen 2015; Wellinghoff, Tong, and Hu 2015). Building and maintaining distribution network assets would remain the responsibility of one or more distribution network companies (the "wires" companies) that operate within the IDSO's territory. As an independent agent that does not own network assets and has no financial stake in competitive market activities within its service territory, the IDSO would plan and operate the distribution network and any markets for DSO services in an impartial manner to maximize overall efficiency. Individual wires companies within the IDSO's territory, meanwhile, would not be required to unbundle from competitive market segments such as retailing, generation, or DER ownership or aggregation.

The IDSO would facilitate competition between conventional network assets and "non-wires" DER solutions by planning and competitively procuring the most cost-effective combination of DSO services from both traditional wires companies within their territory and DER owners or aggregators. With its mandate limited to simply planning and operating a reliable and cost-effective system, the IDSO would have no inherent bias towards wires or non-wires providers of needed network expansion or congestion relief, voltage management, loss mitigation, and other DSO services. IDSOs could be established as non-profit entities charged with efficiently managing regional distribution systems, similar to ISOs in the United States. Alternatively, IDSOs could be created as new independent private utilities, subject to incentive regulation. This would ensure that IDSO managers have financial incentives to efficiently manage the system and facilitate effective competition between wires and non-wires services.

Like the distribution owner/operator in the DNO/SO model described above, the IDSO would likely rely, at least initially, on longer-term procurement of DSO services from DER providers. As metering, telemetry, and computational capabilities improve, the IDSO could also potentially become responsible for calculating and communicating distribution locational marginal prices (in close coordination with the local ISO/TSO) to DER aggregators, retailers, and end consumers within its system.

The advantage of the IDSO approach is that it would not require changes to the existing ownership of network assets or financial or legal unbundling of network wires companies and competitive market activities. Wires companies or their subsidiaries or affiliates would be free to engage in competitive activities, including DER ownership, while the IDSO would ensure impartial operation and planning of the distribution system.

This approach does, however, require the establishment of an entirely new entity, the IDSO, which is not a trivial challenge. Furthermore, it sacrifices significant economies of scope between network asset installation, ownership, and maintenance (which would remain the purview of the wires companies) and system planning and operation. While this is less of an issue for ISOs at the bulk level, the order-of-magnitude greater complexity of distribution networks renders it an open question whether separation of system operation from network provision could work in practice. To perform its planning and operational responsibilities, an IDSO would need to become intimately familiar with the complexities of multiple distribution networks owned and maintained by a variety of individual network companies. In addition, distribution network providers are almost continually deploying crews for preventative and restorative maintenance at all scales, which necessitates careful coordination with system operation and can require changes in system topology. The IDSO model also divides responsibility for delivering key quality-of-service outputs between the IDSO (responsible for planning and operational decisions impacting quality of service) and the wires companies (responsible for maintenance decisions). This may necessitate more complex legal or regulatory structures to ensure sufficient quality of service. At the very least, the IDSO model would impose significant coordination and transaction costs between system operator and network provider functions. Finally, due to their intimate knowledge of their networks, the wires companies operating within a given IDSO territory may have advantages over other competitors, such as DER providers or aggregators that could offer alternatives to network investments. This information asymmetry may act as a barrier to the efficient integration of non-wires alternatives to traditional network investments.

6.2.4.3 Closely regulated, vertically integrated utility

The final option we consider for restructuring at the distribution level is to incorporate all of the critical functions, including ownership of the distribution network, distribution system operation, and any markets for DSO services, into a vertically integrated, closely regulated utility.[16] This utility could also be responsible for retailing, generation, and/or transmission ownership and operation. The challenge in this model is to appropriately regulate the utility such that its incentives are aligned with the efficient integration of DERs and such that any potential for the utility to employ its central platform role to foreclose opportunities for DERs to access markets (i.e., vertical foreclosure) is minimized.

The most effective way to align the utility's incentives would be to manage the utility under a regulation that equalizes incentives for savings between operational and capital expenditures (similar to the distribution owner/operator model discussed above), as discussed in **Chapter 5**. If implemented effectively, the utility's distribution business unit could be made financially impartial between wires and non-wires options in network planning and operation and could be incentivized to pursue the most cost-effective strategies for building, maintaining, and operating the distribution system.

However, given the utility's vertical integration into generation and/or retailing activities, it will be impossible to render the utility fully impartial to DER adoption and integration, as DERs have the potential to compete directly with the utility's retailing, generation, or transmission businesses. This raises the potential for the utility to exercise vertical foreclosure by managing the distribution system in ways that disadvantage DERs where they would compete with the utility's business units. A vertically integrated DSO would therefore have to be subject to close regulation to minimize opportunities for vertical foreclosure, including requirements to procure network services through transparent and open auctions that allow DERs to compete on a level playing field. This option is similar to the way that vertically integrated, traditionally regulated utilities may be required by regulators to fairly and transparently consider signing contracts with independent power providers in lieu of building and owning their own generation. In addition, the vertically integrated DSO would be required to offer transparent, cost-reflective, open-access tariffs for connecting DERs to its networks and to adjudicate interconnection requests in a timely manner, paralleling the open-access tariffs and interconnection processes required of transmission system owners. Finally, the regulator would also need to closely oversee the utility to ensure that system planning accounts for DERs as another integral customer of the utility's services and as a provider of alternatives to utility-owned generation and network investments. This could be facilitated via long-term integrated system planning processes, similar to the integrated resource planning processes required of regulated utilities at the bulk power system level (CPUC 2014; Hawaii PUC 2014; Wilson and Biewald 2013).

The greater complexity of distribution networks significantly increases the regulatory burden of effectively overseeing each of these critical functions at the distribution level. The challenges associated with ensuring non-discriminatory behavior and facilitating open competition in bulk power systems ultimately led most regulators to pursue reforms that ensured the structural independence of transmission system operators. Similarly, the challenges and costs associated with sufficiently regulating a vertically integrated DSO should not be underestimated.

16 Note that this model is not compatible with current law in the European Union (European Commission 2009, 2010).

6.2.4.4 Variations of these models

Careful consideration of the key functions of market platform, system operation, and network provision suggests several possible variations of the three archetypal models for assigning distribution system responsibilities described above.

For example, while the practical feasibility of an independent distribution system operator (IDSO) remains untested, the challenges associated with this model could be reduced if system planning and operation responsibilities are limited to the higher voltage, meshed portions of the distribution network (e.g., “sub-transmission” assets, generally 34.5-115 kilovolt lines in the United States and associated substations and assets).[17] By limiting the scope of independent system operation and planning to the meshed, high-voltage sub-transmission assets, this approach sidesteps many of the practical modeling and market challenges associated with unbalanced phases, vastly increased network complexity, more frequent topology switching, etc. common at lower voltage levels. Assets at the sub-transmission scale are fewer in number and are operated similarly to transmission system assets that are already ably managed by ISOs and TSOs. Indeed, as discussed below, the division between “transmission” and “distribution” networks is arbitrary, especially at this scale, as assets managed by some ISOs or TSOs already range as low as 34.5 kilovolts (kV) while sub-transmission assets managed by distribution utilities include some lines rated at 115 kV or even higher. A *sub-transmission system operator* (StSO) model could thus prove more immediately feasible than an IDSO model in which the IDSO is responsible for the entirety of the distribution network. Indeed, given existing competency and experience with the operation of similar assets at the transmission level, responsibility for operating sub-transmission network assets could be assigned to existing bulk power system independent system operators in an *extended* ISO model.[18] This approach would capture economies of scope in operation and planning between transmission and sub-transmission assets.

One possible variation of the DNO/SO or vertically integrated utility models would be to assign responsibility for market platform design and operation to an independent entity, the *market platform operator.* The market platform operator could establish a common market platform for the purchase of distribution network services (e.g., network reserves for voltage control or congestion management) across a wider geographic area than that of a single distribution utility’s service territory.[19] While distribution network services purchased through this market platform would be local in nature and could only be offered by agents located at specific points within a given distribution utility’s network (in contrast to fungible markets for wholesale power), a common market platform could capture economies of scale in market operation, define consistent market rules that reduce transaction costs for agents selling services within multiple distribution networks, and improve the transparency of system services purchases by distribution utilities. Experience in Europe with decoupling market platforms (such as power exchanges) from transmission system operation and planning provides an analog for this potential variation. However, as experience in bulk power systems has demonstrated, an independent market platform alone is insufficient to ensure a competitive sector: Independence and/or appropriate transparency and oversight of distribution system planning and operation remain paramount.

17 The boundary between meshed and radial networks may also provide the basis for a more technically sensible division between the responsibilities of system operators, as discussed in **Section 6.3**.

18 This approach may raise a number of jurisdictional questions in the United States, as the Federal Power Act (FPA) generally assigns responsibility for regulating distribution networks to states and responsibility for regulating bulk transmission assets to the Federal Energy Regulatory Commission (FERC). However, the FPA is clear that transmission operators have jurisdiction over the interconnection of assets at any voltage level where those assets intend to sell output for resale. In addition, for planning purposes, 69 kV (and even some 34.5 kV) network investments for reliability may be assigned to the jurisdiction of the transmission operator.

19 This option is being considered in New York, where a common market platform may be created and managed by a consortium of individual distribution utilities. See Tabors et al. (2016).

Table 6.2: Benefits and Challenges of Different Industry Structures

	DISTRIBUTION NETWORK OWNER/SYSTEM OPERATOR (DNO/SO)	INDEPENDENT DISTRIBUTION SYSTEM OPERATOR (IDSO)	CLOSELY-REGULATED, VERTICALLY-INTEGRATED DISTRIBUTION UTILITY
BENEFITS	• Economies of scope from combining distribution market platform, system operation, and network provider functions. • Structural unbundling minimizes incentives for discriminatory behavior and ensures DNO/SO acts as neutral platform for competitive market activities. • As a second-best alternative, functional independence can be approximated via legal unbundling and sufficient "Chinese walls" between the network company and competitive affiliates.	• Independent system operator acts as neutral facilitator of markets and distribution system operation. • Does not require owners of distribution network assets to be unbundled from competitive affiliates.	• Captures economies of scope between all three distribution system functions as well as bulk system functions.
CHALLENGES	• Structural unbundling can be difficult to implement in practice. • Effective functional independence can entail significant regulatory burden.	• Hypothetical construct, untested in practice. • Loses economies of scope between IDSO and wires company activities, which could entail significant transaction and coordination costs – e.g., between IDSO planning and operation and wires company investment and maintenance.	• Significant regulatory burden. • Reduced opportunities for competitive provision of DER services.
VARIATIONS	• Market platform responsibility could be assigned to independent *market platform operator*, where the operator manages the market for multiple utility service territories. Captures economies of scale in market operation, defines common market rules that minimize transaction costs, and could improve transparency of distribution service purchases.	• *Sub-transmission system operator* (StSO) with planning and operation responsibility limited to higher voltage, meshed "sub-transmission" network assets. Avoids challenges associated with unbalanced phases, huge increases in network complexity, topology switching, etc. at lower voltage levels. • *Extended ISO*, similar to StSO but responsibility for sub-transmission asset operation and planning is assigned to existing independent transmission system operators (ISOs/RTOs). This model captures economies of scope and builds on the ISO's established experience.	• Market platform responsibility could be assigned to independent *market platform operator* that would manage market for multiple utility service territories (see DNO/SO column for detail.)

6.2.5 Data management: A fourth core function?

To facilitate level-playing-field competition between aggregators (including retailers), DER providers, and diverse competitive agents active within distribution systems, a fourth core function may become increasingly important: that of data platform or data hub. Smart metering and the infusion of information and communications technologies (ICT) into distribution systems are dramatically expanding the amount of data available on distribution system conditions and network user behavior.

As experience in retail markets in Europe and elsewhere has demonstrated, all market participants need equal and non-discriminatory access to a degree of customer information sufficient to facilitate a level playing field for competition (CEER 2015b, 2016). Likewise, timely and non-discriminatory access to data on network conditions and operation and planning decisions, as well as information on network customers, could be important to facilitate competition among DER service providers and aggregators (NYDPS 2015a).

A data hub function may therefore include responsibility for storing metered data on customer energy use, telemetry data on network operation and constraints, and other relevant information needed by competitive market agents, together with responsibility for effective stewardship of these data assets. The data hub would provide non-discriminatory access to these data; in addition, it could provide data analytics services to registered market participants while protecting the privacy and security of individual electricity consumers and other agents. Finally, the data hub should provide end consumers with timely and useful access to data on their own usage of electricity services, empowering consumers to seek new service providers.

Even as the data hub function aims to facilitate access to data to animate competitive market activities, managing customer data in a way that preserves privacy and security is paramount and must be assured (for further discussion, see US DOE [2015]). For example, CEER (2015b) recommends that "customer meter data should be protected by the application of appropriate security and privacy measures" and that "customers should control access to their customer meter data, with the exception of data required to fulfil regulated duties and within the national market model." With a similar objective of empowering consumers to control access to their own data, the New York PSC recommends the creation of "a data exchange to include monthly usage data and certain other customer information on an opt-out basis" while any "customer-specific data that is more granular than total monthly consumption would be provided on an opt-in basis only" (NYDPS 2015a).

A data hub or data exchange may constitute a fourth critical power system function, with responsibility for securely storing metered data on customer usage, telemetry data on network operation and constraints, and other relevant information; providing non-discriminatory access to this data to registered market participants; and providing end consumers with timely and useful access to data on their own use of electricity services. Responsibility for this function should be carefully assigned while considering multiple goals, including non-discrimination, efficiency, and simplicity.

This data hub function may ultimately become just as important as the three traditional core functions described above, motivating further consideration of appropriate industry structures. Economies of scope between metering, system operation, and data access argue for combining responsibility for the data hub with the distribution system operator function.[20] This once again places the focus on ensuring the effective independence of the DSO. If this independence cannot be guaranteed, an independent data hub agency is a

20 Note however that the DSO will not have immediate access to all information that may be ideally managed and exchanged by the data hub. Retailers and TSOs may also possess other information to be shared on the data hub.

second best option. If the DSO is independent of other competitive agents, the DSO can act as a neutral manager of the data hub. To the degree DSOs are integrated with competitive market segments, the importance of the data hub responsibility argues for further enhancement of functional independence or the establishment of an independent data hub manager (Smart DCC 2015). As the New York PSC contends, "asymmetry regarding system information if continued will result in a barrier to new market entry by third parties and ultimately impede innovation and customer choice" (NYDPS 2015a). Decisions about the governance of the data hub are thus strongly related to the other structural choices discussed above. Furthermore, the possibility cannot be ruled out that future ICT developments might render a centralized data hub unecessary, while making compatible the selective preservation of privacy, non-discriminatory access to data of commercial value, and DSO access to data needed to perform security functions.

6.2.6 Restructuring "all the way to the bottom"

To sum up, distribution networks and system operations are now at the heart of modern electricity markets. Just as restructuring transmission network utilities was essential to create a level playing field for competition at the bulk power system level, regulators and policy makers today must carefully reconsider the roles and responsibilities of distribution utilities. The time has come to address incentives and create structures for the efficient provision of electricity services by a diverse range of conventional and distributed energy resources and network assets. In particular, assignment of responsibility for three core functions—market platform, system operator, and network provider—must be carefully considered. In addition, a fourth function, the data hub or data exchange, may become increasingly relevant to unlock competitive markets at the distribution system and retail level.

One of the most challenging and contentious aspects of this new era of restructuring will be taking the steps needed to establish sufficient independence between competitive market activities and critical distribution market platform, system operation, and network planning processes. Just as the independence of transmission networks and bulk system operators was foundational for competitive wholesale markets, independence of distribution system operation and planning is essential to providing a truly neutral trading platform for competition between centralized and decentralized resources, as well as between traditional network investments and solutions to distribution operation challenges that harness DERs (e.g., "non-wires" solutions).

As experience with restructuring in the bulk system has demonstrated, the best solution, from a competitive market perspective, is structural reform that establishes financial independence between the distribution system operator (DSO) and any affiliates in competitive markets, including adjacent wholesale generation and ancillary services markets and competitive retail supply and DER markets within the DSO's service territory.

Experience in bulk power systems and well-established regulatory principles lead to a set of clear recommendations for restructuring at the distribution level, although regulators must also contend with the implications of existing industry structure and the costs of any proposed transition in each context. For example, the best solution from a competitive market perspective is structural reform that establishes financial independence between the DSO and any affiliates in competitive markets, including adjacent wholesale generation and ancillary services markets and competitive retail supply and DER markets within the DSO's service territory. Between the two primary alternatives to achieve structural independence—namely, a combined distribution network owner/system operator (DNO/SO) with financial independence from competitive affiliates and an independent distribution system operator (IDSO)—only the former has so far proven its practical viability at the distribution level. Furthermore, the DNO/SO construct captures the significant economies of scope between system operation and physical network provision. However, two possible variants of the IDSO concept—a sub-transmission operator or extended ISO—may be more practical and more consistent with the existing industry structure in some jurisdictions.

That said, we note that diverse conditions exist in various jurisdictions, and the objective of making the DSO independent must be considered alongside the industry restructuring implications of this strategy in every regulatory jurisdiction and power sector context. As a second-best alternative, various forms of legal and functional independence can be established. These structures will need to be complemented by transparent mechanisms (e.g., auctions or markets) for selecting services where DERs and centralized network services might compete to ensure that no conflicts of interest are exercised. It must also be noted that efforts to establish legal and functional unbundling ultimately proved insufficient at the bulk system level in many jurisdictions (FERC 1999, European Commission 2009). Measures that fall short of making DSOs financially independent may prove more problematic as competitive markets for DERs and for DER-enabled services develop further. Facilitating a level playing field between DERs and conventional approaches to generation and network services is likely to remain challenging if distribution network utilities are vertically integrated into competitive market segments—in that case, significant regulatory oversight will be required.

Effective regulatory oversight for network operation and planning and transparent mechanisms for the provision of distribution system services are critical to prevent conflicts of interest.

6.3 Coordinating Distribution and Bulk Power System Operations

When consumers are passive and power flows across distribution networks are easy to predict, nothing more than a reasonably detailed network model is needed to consider the impacts of distribution networks on bulk power system operations. In contrast, if DERs and flexible demand participate actively in energy- and security-related markets and bidirectional power flows become common, the conventional distinction between the "bulk power system" and the "distribution system" will become increasingly blurry. This makes coordination between both system operators increasingly necessary for the efficient and reliable functioning of the power system.

Different solutions for addressing coordination challenges between the bulk power and distribution system operators are currently being studied by researchers[21] and considered by policy makers and various stakeholders.[22] This section does not aim to resolve these ongoing discussions—instead, it highlights the main coordination challenges for transmission and distribution system operations, regardless of how responsibilities for the three (or four) core functions are assigned.

If DERs and flexible demand actively participate in energy- and security-related markets and if bidirectional power flows become common, the conventional distinction between the "bulk power system" and the "distribution system" will become increasingly blurry. This makes coordination between both system operators increasingly necessary for the efficient and reliable functioning of the power system.

21 See, for instance, the SmartNet and EvolvDSO projects.

22 For instance, the Council of European Energy Regulators (CEER 2016) and several associations of distribution and transmission system operators (CEDEC et al. 2016) are proposing improvements in the coordination between TSOs and DSOs.

6.3.1 Reconsidering boundaries between transmission and distribution system operation

First, it must be noted that boundaries between transmission and distribution systems and the respective domains of bulk power system operators (TSOs or ISOs) and distribution system operators (DSOs) are somewhat arbitrary. The voltage ratings of transmission and high-voltage "sub-transmission" distribution lines often overlap. In addition, the "high-voltage" portions of distribution networks are typically meshed networks, just like transmission systems. "Medium-voltage" distribution networks (also called "primary feeders")[23] are typically also constructed in a meshed fashion (particularly in urban and semi-urban networks), although, to simplify power flows and reduce the severity of short-circuits, they are regularly operated with some circuits "open," which renders them radial in practice.[24] The physical configuration of these networks is important, because electrons care little for jurisdictional boundaries and respect only Kirchhoff's Laws. Power flows across the meshed portion of networks that are considered part of the "distribution system" thus have complex and interactive impacts on power flows, congestion, and system operation decisions in "transmission networks" and bulk power systems.

These impacts already exist today, of course. For this reason, bulk power system operators sometimes incorporate power flow models of adjacent high-voltage distribution networks in their operational decisions and security-related markets.

In a world with greater penetration of DERs, the boundary between system operators may need to be reconsidered. From a technical perspective, the dividing line may be more appropriately drawn at the transition from meshed to radial networks. Agents along a radial network may be treated as a sub-problem in the overall system optimization and represented as a single aggregated point at the head of the radial line (e.g., the substation). This facilitates a single topological interface between system operation problems at the transition from meshed to radial networks. In contrast, separating meshed networks into independent operational problems is technically challenging—as current efforts to coordinate operational decisions and power flows across jurisdictional boundaries between interconnected ISOs or TSOs demonstrate. Finally, of the two candidates for operating meshed networks and associated markets, bulk power system operators (ISOs or TSOs) have much greater experience in this area than DSOs.

Technical considerations, of course, must be taken up alongside jurisdictional, regulatory, and economic considerations as well. Depending on the jurisdiction, changing responsibility for system operation (e.g., to encompass meshed "distribution" networks in bulk power system operations) may require changes in law, regulation, or asset ownership. In other jurisdictions, such changes may result from a natural evolution of steadily expanding and integrating energy markets and power systems. This study therefore takes no final position on the ideal structure for any particular jurisdiction, but rather urges a thoughtful reconsideration of traditional arrangements as part of a broader acknowledgement of the end of the conventional "top-down" paradigm (see **Chapter 3**).

6.3.2 Coordination and information exchange at the interface between distribution and transmission systems

Wherever the boundary between "bulk" and "distribution" systems is drawn, greater coordination between system operators at the distribution and bulk power system levels will be essential as DER penetration increases. Information exchanges between adjacent or embedded DSOs will also become increasingly important. With higher levels of DER penetration, coordination between system operators (TSOs and DSOs) will need to expand in areas such as information exchange, monitoring and analytic capabilities, computation of prices for electricity services, forecasting, scheduling and activation of resources, and management of emergency conditions. Such coordination is of utmost importance for obtaining the full value of services that can be provided by DERs,

23 Voltage ratings for these primary feeder networks are generally 3-36 kV in Europe and 2.2-46 kV in the United States (Eurelectric 2013b; Glover et al. 2011).

24 The meshed construction and radial operation of medium-voltage distribution networks enables changes in network configuration to react to network failures, re-routing power flows and enhancing reliability.

which include (but are not limited to) firm capacity, congestion relief, loss reduction, reactive power, voltage control, and frequency reserves.

6.3.2.1 Energy and power flow forecasts from long term to short term

A higher penetration of DERs will require more accurate forecasts of power withdrawals and injections by load and of DERs embedded in lower voltage networks (Mallet et al. 2014). DSOs are responsible for ensuring that local voltage and thermal constraints across their systems are not violated. As DER penetration increases and demand becomes a more active participant in wholesale and system operator markets, DSOs will require greater information on the schedules and real-time injection or consumption of power by resources and loads connected at distribution voltage. Not only are DSOs in many jurisdictions currently missing this information, in some cases the TSO actually receives the information from distributed generator resource profiles, either directly from distributed generation owners or through aggregators, bypassing the DSO entirely (Mallet et al. 2014). At the planning phase, both system operators may share forecasts and network models, including underlying assumptions and expected distributed-resource connections. Coordinated planning processes will allow market participants, through prices and charges, to react coherently and provide solutions (CEER 2016). Information sharing on planned outages, maintenance, construction, and faults will help system operators coordinate actions to reduce associated costs.

This study envisions a near future of information flow coordinated among three main actors: network users or third-party aggregators, DSOs, and TSOs (or ISOs). Currently, network users are not incentivized to provide energy schedules and system operators independently forecast expected load schedules. In a future with more price-responsive demand and greater penetration of DERs, a hybrid situation may evolve, in which DSO market constructs elicit greater information from participants, some agents may enter into long-term contracts or options, and others simply respond in a distributed fashion to prices and charges. DSOs will then have to compile available information and perform optimal power flow (OPF) calculations to identify likely network constraints. As DERs change load shapes, DSOs should also share with TSOs their forecasts of distribution-level loads and generation in a timely manner. In addition, DSOs would communicate—to network users or their representative aggregators and to the TSO—information about any local active constraint that can restrict the forecasted schedules and/or information about associated congestion price forecasts that would enable flexible demand and DERs to optimize consumption or production and avoid unfavorable prices. The TSO, on the other hand, would communicate to the DSO information about the scheduled dispatch of DERs that provide TSO services, thereby allowing the DSO to modify its OPF computation to cross-check feasibility.

6.3.2.2 Energy schedules at different timeframes

Energy schedules, after positions are taken in different markets, need to be shared between the TSO and DSOs to update forecasts and perform power flow analyses closer to real time. This could be implemented through a market design that expands current day-ahead planning, which includes DERs and any local constraints. After the day-ahead market, the TSO could inform the DSO of final schedules for DERs that participate in the market. In the same way, if the DSO changes DER schedules because of local constraints, DSOs would inform the TSO. Different mechanisms have been proposed to managed information exchange between TSOs and DSOs, specifically in cases where DERs provide services to the TSO that eventually can violate constraints at the distribution level (SmartNet 2016). The mechanism chosen will depend on the assignment of core functions, as discussed in **Section 6.2.4**. Potential changes to schedules due to local constraints may be based on bids from network users or even on the curtailment of injections or withdrawals at certain nodes. In the latter case, compensation schemes may need to be defined. Lack of schedule information may lead to suboptimal dispatch of the system. In the PJM market,[25] for example, most DERs are currently not registered at the wholesale level and their output is not considered when the TSO

25 PJM is an electricity market in the United States that serves all or parts of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia, and the District of Columbia.

runs an optimization for the whole system (PJM 2015). At the same time, a subset of DERs managed by curtailment service providers aggregate different resources to provide economic and emergency demand response to PJM for bulk power system operations (PJM 2014). Although distribution utilities are informed when DERs register with PJM, distribution utilities are not currently informed when these resources are actually scheduled to provide system services. Moreover, the aggregator can at any time change the location of the specific DERs that are providing the required response, so long as deliverability to the system operator is assured. Because this may cause constraints in distribution networks in the future, such disconnects in communication should be avoided wherever possible.

6.3.2.3 Energy prices at the interface between transmission and distribution system operators

Conceptually, it is possible to compute and communicate prices at the TSO and DSO levels that are consistent with one another without resorting to a massive, centralized optimization process. For example, once the TSO computes energy prices for the meshed network, these prices could be sent to the DSO to be incorporated in the computation of prices for the distribution system or when clearing local markets in radial networks (see **Chapter 4** for further details). This decomposition is best facilitated by simultaneously computing prices for the entire meshed portion of the network, before propagating down to radial network branches using a centralized or distributed solution method. Iterative solution methods may be required to reach convergence between DSO and TSO prices.[26] If locational marginal prices (LMPs) are not computed in the distribution system, a second-best solution may be implemented, such as wholesale prices with an estimated distribution loss factor, local market mechanisms, and/or contracts. In any case, a comprehensive system of prices for energy, ancillary services, and network services required by both TSOs and DSOs, together with a well-coordinated series of temporally linked markets, would play an important role in coordinating decisions where conflicts arise between various services. This would help ensure that DERs commit their capabilities to the most valuable combination of services while respecting their various technical constraints.

6.3.2.4 Real-time dispatch and activation of services

In real time, different services need to be activated, final positions need to be determined, and active resources need to be dispatched. DER owners or third-party aggregators would take on responsibility for activation based on dispatch directives, price signals, and penalties for non-fulfillment of previous commitments.

6.3.2.5 Emergency conditions and restoration arrangements

Information on changes in energy schedules and available resources may need to be sent to both operators during emergency situations, when system security is in danger and fast action is needed. In such cases, the TSO may need support from the DSO to reduce or curtail loads or generation connected to the distribution network. In addition, the DSO may have local issues (e.g., line faults) that may be relevant to communicate to the TSO and may require support from resources connected at the TSO level. Advanced network codes may need to incorporate new actions and procedures in emergency situations and establish communication protocols between operators, such as Europe's Network Code on Emergency and Restoration (CEER 2015a; ACER 2015). Restoration arrangements between TSOs and DSOs will enhance the resilience of the system and provide more cost-effective solutions compared to the current situation where DSOs do not intervene (CEER 2016). Data exchange between TSOs and DSOs will require the standardization of formats and protocols for each timeframe so as to reduce the exchange time and increase data usage (CEDEC et al. 2016).

26 Hypothetically, one might do away with the distinction between transmission or meshed power system and distribution or radial power system entirely and compute prices and dispatch decisions using a fully distributed algorithm based on proximal message passing (Kranning et al. 2013). However, this approach remains untested in practice and is not proven to converge to optimality if the problem is not well conditioned (that is, if the problem is non-convex).

6.3.3 Coordinating the procurement of energy services with the use of auctions

As discussed in **Chapter 4**, the combination of economic incentives established by distribution-level LMPs (if employed) and/or peak-coincident network capacity charges can incentivize efficient short-run behavior accounting for the impacts of network users on both short-run losses and network congestion (via LMPs) and anticipated network capacity investments (via peak-coincident network charges). However, relying solely on these economic signals and network users' responses to coordinate long-lived investments in both network capacity (by the DSO) and DERs (by network users) faces several additional challenges.

An economically efficient network planner seeking to maximize social welfare would make investments in network capacity only up to the point where the cost of network expansion equals the benefit derived by network users from the expanded capacity over the economic life of the asset (applying appropriate discount rates). However, network utilities have little knowledge of network users' actual preferences. Individual willingness to pay for network usage can be inferred from historical patterns of network user behavior in response to LMPs and network capacity charges, but these historical observations may provide only an incomplete picture of network user's long-term preferences. The significance of this "observability problem" is exacerbated by the fact that most network investments are long-lived, capital-intensive assets. Once a network investment is made, its costs are almost entirely sunk and the investment is, for all intents and purposes, irreversible. This increases the risks of taking action amidst uncertainty and incomplete information with only network users' historical patterns of behavior to inform network investment decisions.

Network users considering investments in DERs face similar challenges. Many DERs are capital-intensive investments as well, and the economic value of these investments will depend on future network expansion decisions made by network utilities. Thus, both network utilities and network users must make investments given incomplete information that exposes largely sunk costs to significant risks.

Regular auctions for forward network capacity options could solve this coordination challenge and help overcome incomplete information. Such auctions solve the coordination problem by communicating to network users the marginal cost of forthcoming network expansion (or approximation of the marginal cost for discrete investments) and creating incentives for network users to reveal their willingness to pay for forward options to use network capacity.

With sufficient lead-time to make investments based on auction results, the network utility would request demand bids for forward network capacity options contracts for each area of the network that is experiencing congestion or expected to experience congestion in the near-future—i.e., if the network capacity margin has become small. Each bid would reflect a quantity of network capacity (in kW) and a price (in $/kW-yr) reflecting the network user's willingness to pay for the option to use that quantity of capacity during periods of congestion. The option would entitle the user to a payment equal to the difference between the LMP at a reference bus upstream of the congested area and the LMP in the congested zone at any time when the LMP at their zone is higher than the reference bus, similar to a financial transmission right (FTR) in bulk power systems (e.g., see Hogan [1997]). Any actual consumption would pay the full LMP at their node. The option would thus hedge the consumer against increases in nodal prices due to congestion, so long as they consume less than or equal to the option quantity procured, while preserving efficient short-run incentives to reduce consumption if the user's marginal value of consumption is lower than the real-time LMP.[27]

DERs that are able to relieve network congestion by injecting power[28] at the time of network constraints could also bid a contract price ($/kW-yr) and firm capacity quantity (kW) into the options auction as an alternative

27 Equivalently, actual consumption up to the option quantity could be settled at the reference node LMP, with any reductions below this level compensated at the difference between the reference node LMP and the user's LMP. Any consumption above the option quantity would be settled at the user's LMP. This yields the same marginal incentives.

28 This example assumes network congestion is driven by withdrawal (consumption) peaks. Constraints could also be driven by injection (generation) peaks, in which case this auction process is reversed, with bids for network injection capacity options and DERs capable of increasing withdrawals during injection peaks bidding their firm capacity as alternatives to network expansion.

to network expansion. DERs commit to a firm put option which network utilities can exercise at periods of network congestion, up to the contracted firm capacity quantity. Appropriate penalties for non-performance would be assessed. Note that DERs co-located with loads capable of reducing net withdrawals during periods of network congestion can likewise reveal and monetize their value by allowing network users to reduce their requirements for network capacity and lower their options bids accordingly, reducing the costs these users will pay for options contracts.[29]

Network users could submit their own bids, or alternatively, sign up for aggregators who submitted bids on their behalf. Aggregators better able to reflect or predict user preferences would be able to offer lower costs to their customers, as they would face lower fixed costs of options and lower exposure to LMPs. Finally, any users who did not submit bids, either directly or via aggregators, would have a conservative estimate of their demand for firm network capacity added to the left-hand side of the bid stack at a very high willingness to pay. In other words, a very conservative estimate would be used for any users that did not submit bids, with these users then obligated to pay fixed costs reflecting this estimate. This is in effect what utilities do to today absent better information about network users' willingness to pay for network capacity. Users with lower willingness to pay or more accurate information about their preferences have a strong incentive to reveal that preference by submitting more accurate bids.

Finally, bids would be assembled and network users' aggregate demand for network capacity options would be compared to the aggregate supply bids of DERs and the marginal or incremental cost of network expansion, as depicted in **Figure 6.1**. The network utility would clear the auction at the intersection of the aggregate demand for network options and the marginal cost of either network capacity expansion or DER put options, whichever is lower (**Figure 6.2**). Users with cleared bids would agree to pay a fixed cost ($/kW-yr) for the capacity options at the clearing price and would in turn have their consumption hedged against congestion and peak-coincident network capacity charges for the length of the option contract. Users who did not clear would not pay any fixed cost for network capacity and would be fully exposed to short-run prices for all consumption. DER owners who cleared in the supply side of the auction would sign put options with performance penalties and receive a fixed payment ($/kW-yr) at the clearing price. Finally, the network utility would invest in any remaining network capacity expansion needed to accommodate cleared demand for network capacity after accounting for the aggregate ability of cleared DERs to reduce network peaks. **Box 6.1** provides a real-world example of a DER auction that was recently run by Con Edison to defer a costly network upgrade in a high-growth area of its service territory. While this auction was not an options auction such as the one described above, it nonetheless shares many of the same characteristics.

29 Note that this means behind-the-meter DERs that are not providing net injection at times of system congestion but rather reduce a user's net withdrawals would bid on the "demand side" of the auction, rather than the supply side, while those that are able to inject at times of system peak bid on the supply side.

Figure 6.1: Example of Network Capacity Options Auction Demand and Supply Curves

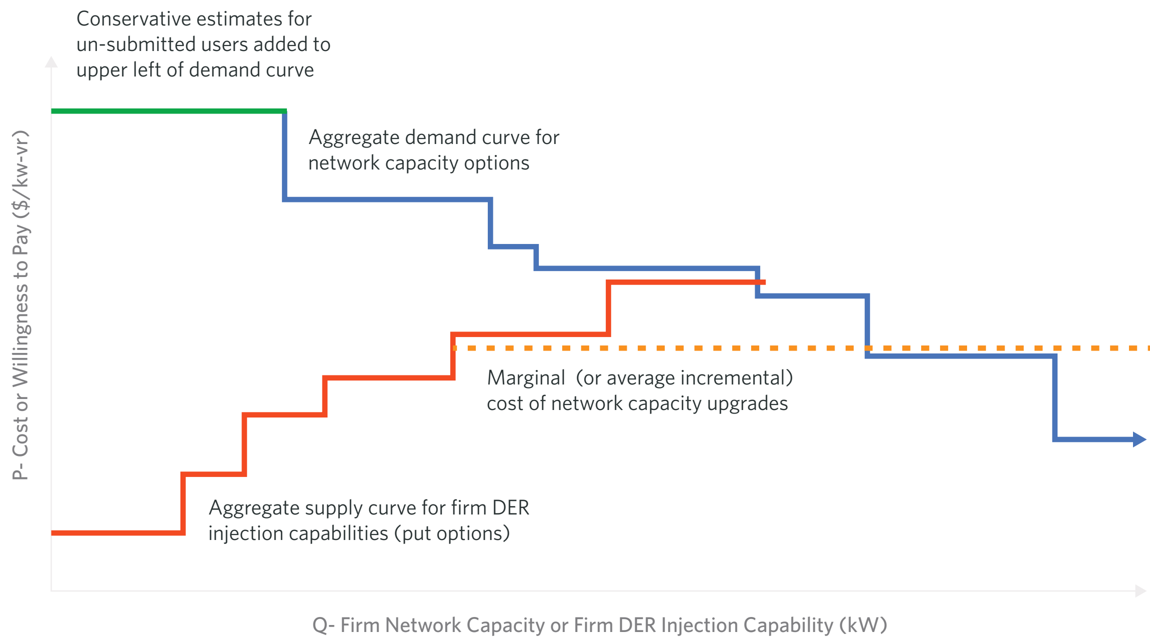
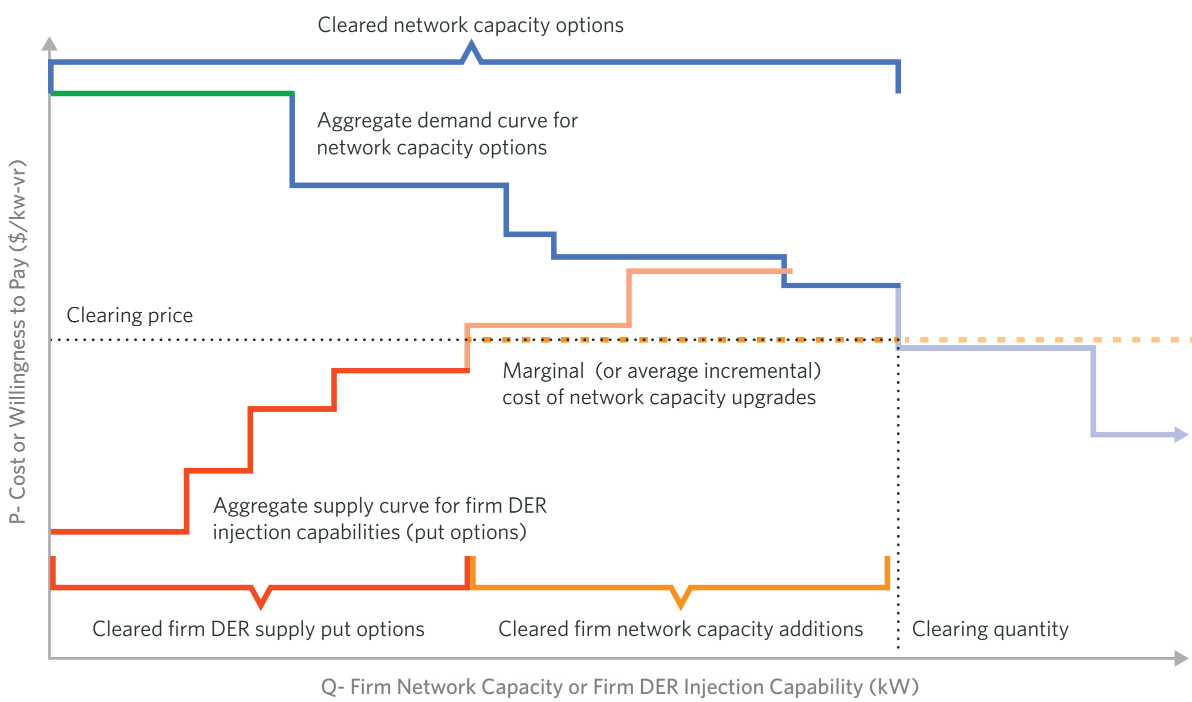


Figure 6.2: Clearing of Network Capacity Options Auction

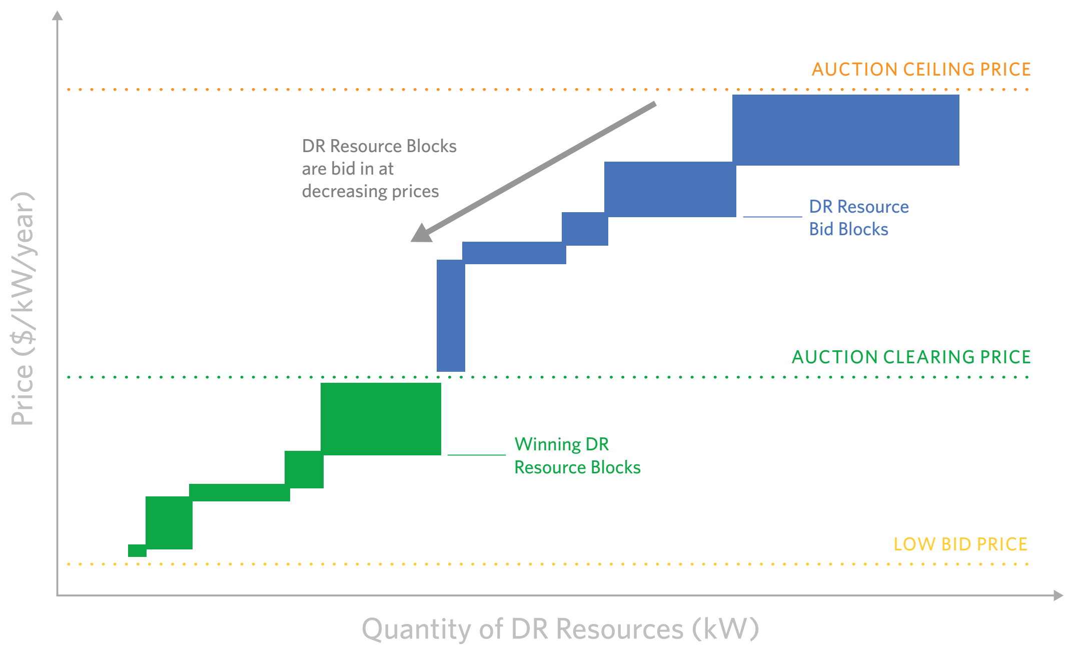


Box 6.1: Brooklyn Queens Demand Management Program Demand Response Auction

Facing growing peak load in parts of Brooklyn and Queens, New York, Consolidated Edison, Inc. (Con Edison) has launched the Brooklyn Queens Demand Management (BQDM) Program. The BQDM Program will procure 52 MW of peak load reduction by summer 2018, with 41 MW of the total load reduction provided by customer-sited solutions such as demand response (DR), energy efficiency, energy storage, fuel cells, solar PV, and combined head and power (NYDPS 2014; Con Edison 2016a). The remaining 11 MW are expected to come from utility-sited solutions such as voltage optimization. Con Edison will use these resources to provide load relief during critical hours on peak summer days, and the deployment of these resources are meant to defer the construction of a $1.2 billion substation upgrade by five years or more.

As part of the BQDM Program, Con Edison ran auctions to competitively procure distributed resources that would deliver load relief between the hours of 4:00 p.m. and 12:00 a.m. (in two separate four-hour blocks) during the summers of 2017 and 2018. The auctions for 2017 and 2018 delivery were held on July 27 and July 28, 2016 and resulted in Con Edison accepting offers for 22 MW of peak hour demand management services from 10 providers including Stem, EnerNOC Inc., Innoventive Power, Direct Energy, Power Efficiency, Demand Energy Networks, Energy Spectrum, and Tarsier (Bloomberg 2016). The BQDM auction clearing prices ranged from $215/kW-year to $988/kW-year (Bloomberg 2016).

Figure 6.3: Illustration of Brooklyn Queens Demand Management Program Auction Mechanics



Graphic source: Con Edison (2016). Reproduced with permission.

Con Edison used a competitive descending clock auction mechanism[30] to procure DR resources under the BQDM Program, as depicted in **Figure 6.3**. Each auction began at a pre-announced ceiling price and auction participants bid peak demand management resources into auction "blocks" that could be between 50 kW and 2,000 kW. Each participant was allowed up to five blocks and up to a total of 10,000 kW per auction. Once each auction was completed, Con Edison established a clearing price that will be paid to all cleared resources (i.e., pay as cleared auction).

Con Edison intends to call a BQDM demand reduction event when it expects that the following day will be a "peak day" in the BQDM Program area. Con Edison defines "peak day" differently for resources providing load relief between 4:00 p.m. and 8:00 p.m., and those providing relief between 8:00 p.m. and 12:00 a.m. For the former (4:00-8:00 p.m.), DR resources are called upon when the anticipated peak electrical load in the BQDM area the following day is no less than 97 percent of the summer peak forecast. For the latter (8:00 p.m. to 12:00 a.m.), DR resources are called upon when the anticipated peak electrical load in the BQDM area the following day is no less than 93 percent of the summer peak forecast. Con Edison expects that there will be an average of three to six calls each year for each product.

30 Each descending clock auction began with a 15-minute timer and any bids placed in the final two minutes of the timer added an additional two minutes to the timer. The auctions ended when the timer ran out or when two hours had passed (whichever occurred first).

6.4 The Value and Role of Aggregation

6.4.1 The emergence of aggregators

As DERs proliferate and opportunities for active or flexible demand grow, "aggregators" of these resources have the potential to help unlock the value of distributed resources and bring them into energy markets at scale. Aggregation is defined here as the act of grouping distinct agents in a power system (i.e., consumers, producers, prosumers, or any mix thereof) to act as a single entity when engaging in power system markets (whether wholesale or retail) or selling services to the system operator(s).[31]

The emergence of a new breed of DER aggregators has sparked debate over the value and role of aggregators in the evolving power system. In Europe's liberalized retail markets, debate is centered around the functioning of retail markets, the ability of retailers[32] to deliver desired levels of consumer engagement and value-added services, and the value or cost of superimposing third-party aggregators over these retailers (Eurelectric 2015; European Commission 2015; Smart Grid Task Force 2015). On the other hand, independent DER aggregators are already highly active in US markets and stakeholders are attempting to design market rules to ensure these aggregators flourish when they provide true value creation as opposed to regulatory arbitrage (CAISO 2015; NYDPS 2014).

Clarifying these debates requires an understanding of the mechanisms by which aggregation creates value. In many cases, aggregators are performing roles today that may not deliver value to power systems but rather reflect

31 This definition builds upon and narrows the definition provided by Ikäheimo, Evens, and Kärkkäinen (2010): *"an aggregator is a company who acts as an intermediary between electricity end-users and DER owners and the power system participants who wish to serve these end-users or exploit the services provided by these DERs."* Ikäheimo et al.'s definition encompasses both aggregation in the context considered in this chapter, as well as the potential role of platform markets. We recognize the existence of other definitions of aggregators; in practice, the definition of an aggregator can be restricted or expanded depending on regulations that define the roles and activities aggregators can perform.

32 Retailers or retail electricity providers (REPs) are aggregators in the sense that they aggregate a number of dispersed consumers (and, at some times, producers) and act as a liaison between these agents and wholesale markets. REPs also comply with power system regulations, perform hedging functions, and undertake other activities on consumers' behalf. Some REPs, such as MP2 Energy in the United States, are performing roles traditionally attributed to third-party aggregators, such as brokering demand response for capacity and ancillary services market participation (MP2 Energy 2016).

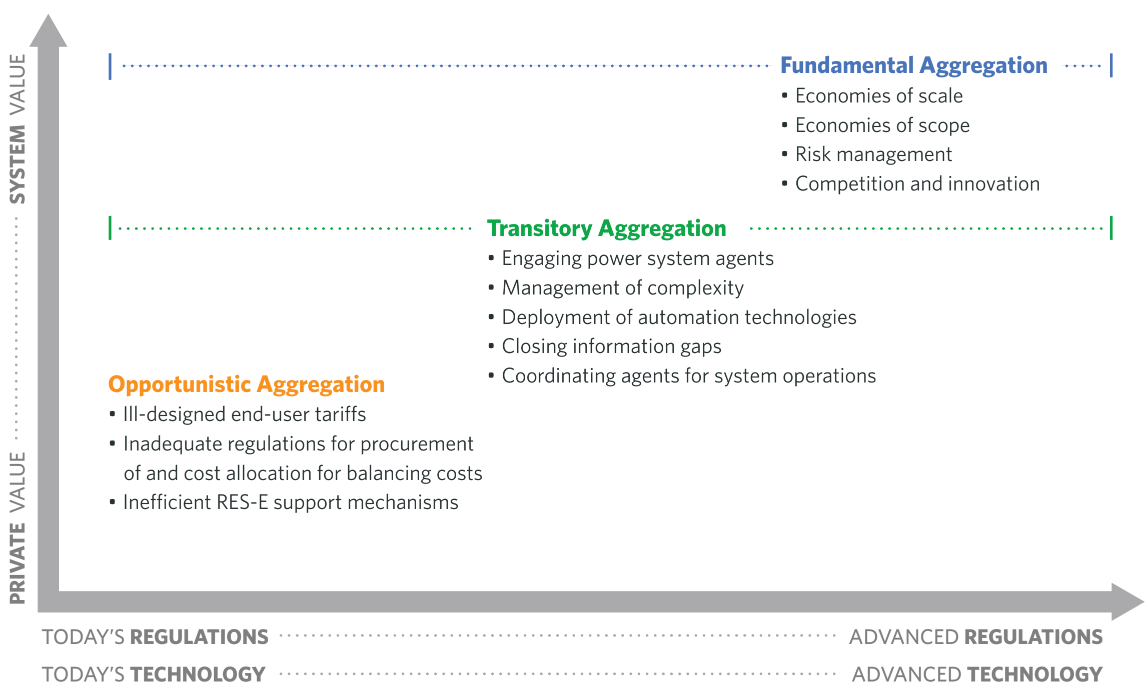
opportunities to arbitrage inadequate regulation. In other cases, aggregation delivers real value, but this value may become less significant in the future, as technological change reduces the costs of information provision, coordination, or transactions. Other activities may deliver enduring value. This section discusses the potential role of aggregation in power systems and the ways in which aggregation can deliver value to both private stakeholders and the power system as a whole. This discussion has implications for pertinent questions, such as: Should the power system accommodate many aggregators or only one centralized aggregator? Who can or should be an aggregator (transmission and distribution system operators, retailers, third parties, etc.)? What market design elements may need to be adapted or adopted to accommodate DERs? What is the best feasible level of unbundling between aggregators and other traditional utilities or retailers?

6.4.2 The present and future value of aggregation

To clarify how and for whom aggregation creates value, we discuss two types of economic value: private value and system value. Aggregation has system value if it increases the social welfare of the power system as a whole. Private value is an increase in the economic welfare of a single agent or subset of agents. Private value creation may or may not align with system value creation. As we will demonstrate, aggregations with private value may create economic value for certain agents at the expense of system-wide economic efficiency. Aggregation may also simply lead to a rent transfer between market actors. In other cases, aggregation helps align private value with opportunities to improve the efficiency of power systems and deliver system value.

We can distinguish three broad categories of aggregation (**Figure 6.4**). First, aggregations with *"fundamental"* value do not depend on the specific regulations, level of

Figure 6.4: Value of Aggregators Based on Technology and Regulatory Contexts



consumer awareness, or technologies in place in the power system, and will be permanent or near permanent in time. Aggregations with *"transitory"* value contribute to the better functioning of the power system under the present and near-future conditions; however, the value of transitory aggregations may wane as technical, managerial, or regulatory conditions improve. Finally, aggregations with only *"opportunistic"* value emerge in response to regulatory or market design "flaws." Due to inherent trade-offs in regulatory principles, there is no single ideal regulatory system. Regulations on system design and operations are inherently plagued by imperfect or asymmetric information, technology constraints, political interference and conflicting regulatory principles, among other flaws. This reality opens the door to different levels of arbitrage. As indicated in the figure, aggregations that create transitory value may exist now (under current regulatory and technological conditions) and in the future (under advanced but "imperfect" regulations and with more advanced technologies). It is difficult to project *a priori* which roles will prove transitory and which will endure, as the significance of economies of scale and scope and the persistence of transaction and coordination costs ultimately depend on uncertain technological developments. The concepts indicated in **Figure 6.4** are nonetheless useful in articulating the potential roles and value of aggregators in modern power systems.

6.4.2.1 Fundamental value of aggregation

Fundamental value stems from factors inherent in the act of aggregation itself. In the context of the power system, aggregation may create fundamental value by capitalizing on economies of scale and scope and by managing uncertainty.

Participation in electricity services markets incurs certain unavoidable costs. First, one must acquire or engage the owner of one or more energy resources (either centralized or distributed resources); second, if these resources are to interact with the market, they must be equipped with some level of information and communications technologies (ICT); third, energy resources and their owners must comply with power system regulations and market rules. Many of these costs include fixed and variable components. The existence of fixed costs may lead to a situation where the average cost of providing a service is higher than the marginal cost. In that case, the average cost of providing the service declines as the quantity of services provided increases. Thus, to the extent that there are fixed costs associated with participating in electricity services markets, there may be value in aggregation via *economies of scale*.[33] For example, fixed costs associated with establishing the ICT infrastructure needed to engage and dispatch DERs may argue for aggregating dispatch responsibilities in larger agents that can harness associated economies of scale.

Furthermore, to the extent that providing multiple services or products entails common technologies, transaction costs,[34] acquisition costs, or knowledge bases, aggregation may create value through *economies of scope*. For example, market participation costs or other transaction costs may mean that it is more efficient for a single aggregator to bundle all required services for or from a customer (e.g., energy, operating reserves, voltage control, etc.), rather than having multiple aggregators that each procure or deliver a single service. Economies of scope and product bundling are present not only for electricity services but also for adjacent sectors that supply heating, gas, energy efficiency solutions, telecommunications, or Internet services. Given the inherent costs of acquiring and engaging a customer, aggregators may realize economies of scope by bundling services and spreading transaction costs across products (Codognet 2004).

Finally, market parties have different risk preferences and capabilities to hedge against risks. A small agent may not be able to hedge against price risks, while hedging products are often available for large agents (through

33 Note that economies of scale depend on the persistence of significant fixed costs for various aggregator services. In some cases, fixed costs may diminish or change in the future, making some activities less valuable over time. As noted previously, which functions ultimately deliver enduring or transitory value depends on the uncertain pace of technological change and its impact on cost structures. This can make it more difficult to distinguish between "fundamental" and "transitory" value *a priori*. In general, activities that exhibit economies of scale because they have a cost structure that involves enduring fixed costs will prove amenable to aggregation that delivers fundamental value.

34 A transaction cost is any cost that an agent has to incur in making an economic transaction. These costs can include search costs (e.g., searching for information) and enforcement costs (e.g., ensuring that a certain amount of energy has actually been provided once contracted for), among others. Like fixed costs, transaction costs may change over time as technology develops. This may render the value delivered by some aggregator functions transitory, rather than fundamental.

contracts for differences, for example). Aggregators may therefore deliver value by *managing uncertainty*—acting as intermediaries between small consumers/producers and volatile markets to provide hedging solutions to market players. For example, in retail markets, retail electricity providers (i.e., demand aggregators) offer stabilized prices to their consumers. The value of this type of risk-hedging service is discussed in Littlechild (2000).

6.4.2.2 Transitory value of aggregation

Aggregators may create value as the power system transitions from current regulations and technologies to a more advanced or idealized future. Temporary value is not inherent to aggregation, but may be unlocked by aggregators. Opportunities for agents in the distribution system to increase system efficiency by engaging with the bulk power system are increasing as ICTs enable loads to become more price-responsive and as DERs are increasingly deployed. However, market complexities, information gaps, lack of engagement, and other biases may prevent the value in these distributed assets from being unlocked. Aggregators may create system value by managing or eliminating these factors.

Aggregators deliver fundamental system value where they harness economies of scale or scope or manage risk for participating agents.

An agent may be capable of providing a service (or set of services) to the system but may lack the information required to do so effectively. For example, consumers often lack information in a number of areas: when system peaks occur, what the prices are for various services they consume, what technologies are available to help them control consumption, what the prices of these technologies are, etc. Additionally, consumers may lack the ability to forecast this information into the future, which is critical for hedging risk or scheduling bids and consumption. Finally, power system operations and planning require the provision of many energy services—these services need to be priced in a manner that incentivizes efficient behavior. However, current power system technology is not capable of calculating and communicating the multiplicity of price signals that are relevant for different agents at theoretically optimal levels of temporal and spatial granularity. An aggregator may be able to intervene to close gaps in information on system operation between the system operator and various agents. Furthermore, one aggregator can gain from economies of scale by processing information from multiple agents, whereas costs would otherwise be multiplied by the number of agents processing this information independently.

It may be difficult to motivate agents to take action if the sums of money that can be gained are small and if the required action is complex. However, when taken in sum, small actions by numerous agents can create significant value for the system. For example, Opower achieves peak energy reductions by communicating when and how energy should be saved. Opower estimates that the average customer in Nevada may save $33 per year (the estimated national average is $28 per customer per year). This equates to just over $2 in savings per month. Without intervention, a customer may not be motivated to take action that would produce such a small sum of savings on his or her own. However, Opower estimates that its service has the potential to avoid over 140 megawatts of generating capacity in Nevada and over 4.5 gigawatts nationally (Opower 2016).

The problem of small financial incentives is exacerbated by the problem of system complexity. Navigating the power system's various rules and codes may be very difficult for an agent that has not historically engaged in power system operations. An aggregator may be able to navigate these complexities on an agent's behalf. Many companies, such as EnerNOC and Comverge, handle complex registration and bidding processes on behalf of the agents they serve, enabling the system to benefit from the services these agents provide and enabling the agents to benefit from previously untapped revenue streams. The value these companies provide stems from engaging an otherwise unengaged customer, but aggregators may, and often do, provide this engagement. Indeed, the mere existence of demand response aggregators in restructured and traditional markets signals that, in many cases, retailers or incumbents are not providing competitive customer engagement services.[35]

Power system operators (ISOs or TSOs) are required to dispatch generation, storage, or demand response resources to match supply and demand at all times. Traditionally, these resources have taken the form of a discrete set of large and easy-to-centralize resources. DERs therefore represent a new challenge for utilities: Coordinating and controlling DERs involves bidirectional communication with an order of magnitude larger number of resources. Aggregators may also create value by coordinating information exchange between various power system actors (Codognet 2009). Whether this value proves transitory or enduring depends on the development of automatic control systems and related technologies.

Finally, if prices and charges for electricity services are not conveyed to end users with sufficient temporal and locational granularity, aggregators may play a transitory role by identifying price-responsive or flexible users in high-priced locations or at high-price times and coordinate their dispatch to reduce system costs and generate revenue for participants. For example, in many US markets, LMPs are calculated at each node in the bulk power system, but loads are settled based on the average price across a zone. In this case, demand response providers could identify high-priced nodes within the zone and coordinate the dispatch of loads at these locations. If the spatial granularity of prices for loads improves, however, this transitory role would be eliminated.

6.4.2.3 Opportunistic value of aggregation

Opportunistic aggregation may emerge as a response to imperfections in market design, regulation, or policy. This form of aggregation occurs when different agents or DERs located at one or more sites aggregate to obtain private value in ways that don't increase the economic efficiency of the system as a whole. Indeed, opportunistic aggregation may work to restrict competition, especially for small agents. We identify three categories of rules that can give rise to opportunistic aggregation: rules related to the procurement of balancing or ancillary services, rules related to the allocation of balancing costs or penalties for non-delivery of committed services to agents, and inefficient locational price signals and/or network charges.

Aggregators may perform many functions that deliver value to power systems today by overcoming information barriers, reducing transaction costs, or coordinating market participants. However, many of these roles may also prove transitory, as improvements in technology (particularly in information and communications technologies and controls) reduce or lower information, transaction, and coordination costs.

For example, the activation of resources committed to provide operating reserves requires available capacity to increase or decrease production or load. Units that provide this availability receive a payment based on the capacity committed or made available to the system operator. In a second step, these units may be required by the system operator to increase or reduce energy production or consumption in real time in response to actual reserve dispatch requirements. In many reserve markets today, a penalty may be applied if a unit that has committed reserve capacity to the system operator is not able to provide energy when called upon. This penalty may exceed the marginal cost of activated reserves. In this situation, inefficiencies can result if aggregators are allowed to net out their reserve delivery requirements across a portfolio of resources—a case of opportunistic aggregation. Consider, for example, an aggregator with a portfolio of two resources, A and B. The marginal costs of providing reserves using these two resources are MC_A and MC_B, respectively, where $MC_B > MC_A$. Consider a case where unit A is committed to provide reserves, the market price of activated reserves is R, and $MC_B > R > MC_A$. That is, it is efficient for resource

35 There may be many factors driving this dearth of action from retail aggregators, including vertical integration with generation (Littlechild 2009). More research is needed in this area.

A to provide reserves at the current clearing price R but not for resource B to do so. However, in a system with a penalty (P) for the non-delivery of committed reserves, the aggregator will dispatch the uneconomical unit B when unit A is not available so long as $(R + P) > MC_B$. Thus, it will be economically rational for the aggregator to dispatch unit B, even if another resource, C, managed by a different market participant, is available with a cost $MC_C < MC_B$. Allowing aggregators to net penalties in this manner can lead to aggregation that improves the private value of DERs but does not improve (and may indeed harm) overall system efficiency. Furthermore, allowing aggregators to net penalties can act as a barrier to entry for smaller market agents. This is one of several possible examples of opportunistic aggregation.[36]

Where aggregation creates only private opportunistic value, regulations, policies, or market rules should generally be modified to align private value with system value.

Where aggregation creates only private opportunistic value, regulations, policies, or market rules should generally be modified to align private value with system value.[37]

6.4.3 An example of aggregation: Charging electric vehicle fleets

As an example of all three forms of aggregation—fundamental, transitory, and opportunistic—consider the case of electric vehicle (EV) charging. Under a fully efficient system of prices and regulated charges for electricity services (see **Chapter 4**), individual EV owners could charge their vehicles in a manner that maximizes both private welfare and system efficiency. In this case, EV aggregators may deliver *fundamental value* by capitalizing on economies of scale associated with the ICT infrastructure needed to manage EV charging. Additionally, EV aggregators may harness economies of scope to provide multiple services to EV owners at lower cost, including charge management, energy supply purchases, and other vehicle-related services, such as regular maintenance, navigation, or other services. Finally, EV aggregators may provide value by managing supply cost volatility and risk for EV owners.

At the same time, wherever prices and charges do not fully reflect time- and location-varying marginal prices for all electricity services, EV aggregators may provide *transitional value* in the near term by aggregating EV fleets and creating appropriate incentives to motivate efficient charging. In the future, smart charging systems embedded in EVs may perform this function automatically without the need for intermediary aggregators. In addition, if transaction costs dictate minimum sizes for participation in wholesale energy markets, capacity markets, or balancing or ancillary service markets, aggregators will be critical to bring potentially cost-effective EV flexibility into these markets. Finally, EV aggregators may provide value by overcoming informational barriers in cases where aggregators are better able to predict EV charging patterns and manage associated uncertainty than system operators.

EV charging may also give rise to *opportunistic aggregation*. For example, many European markets apply dual imbalance penalties, wherein individual agents are penalized for departing from their scheduled consumption or production in *either direction*, even when only one of those directions contributes to overall system imbalances (and thus reserve or balancing market dispatch costs). In such markets, EV aggregators can significantly reduce imbalance penalties by aggregating many individual EV charging locations and bidding a single aggregate charging schedule. This aggregation of imbalance positions is privately valuable, but does not improve overall system efficiency. A single imbalance price,

36 For further discussion of opportunistic aggregation, see Burger et al. (2016).

37 Exceptions include cases where the opportunistic private value created by regulation is explicitly acknowledged and desired as a form of subsidy, or where the cost of regulatory reform exceeds benefits from eliminating inefficiencies and transfers.

reflecting the cost of deviations in the same direction as overall system imbalance[38] would be more efficient and would prevent this type of opportunistic aggregation.

6.5 Business Models for Distributed Energy Resources

Sections 6.2-6.4 explore critical functions and key structural questions in the evolving electric power sector, along with the need for coordinating measures between agents in the sector. This section discusses the diverse competitive agents that are emerging in the US power system, focusing on new business models for harnessing DERs to provide electricity services to system operators, network providers, market participants, and end consumers. Analyzing these business models further reinforces the importance of properly structuring the electricity industry and associated markets in order to (1) minimize the prevalence of DER businesses built on opportunistic, rather than fundamental or transitory, value and (2) create a level playing field for competition between the many business models that are capable of providing the range of electricity services desired by markets, utilities, and end consumers.

6.5.1 Analysis of DER business models

In **Appendix B**, we present the results of a detailed analysis of 144 DER business models in current use—roughly 50 models in each of three technology categories: demand response (DR) and energy management systems (EMSs), electrical and thermal storage, and solar PV. Data for the companies used in this analysis were collected from publicly available news, academic, and industry publications between February 2014 and October 2015. To ensure that our sample was representative of the "universe" of DER business models that exist today, we also sampled from the Cleantech Group's i3 database, a commercial database that contains information on more than 24,000 "clean tech," DER, and sustainability-focused businesses (i3Connect 2016). The i3 database categorizes businesses by their core focus; we used this feature to create three sets of business models—one for each DER technology category—that capture all the businesses in the database: (1) "ground-mounted PV" and "rooftop PV" (which together form the solar PV set), (2) "grid energy storage,"[39] and (3) "demand response." We then drew stratified random samples from each of these three sets, such that the distribution of companies in our final sample was similar to that in the i3 database in terms of the regional location of company headquarters and in terms of companies' founding year. We sampled 50 companies in each of our three DER technology categories. A small number of the sampled companies did not fit our coding criteria, and thus were not included in our final sample.

6.5.2 Summary of DER business models

Figures 6.5 through **6.7** summarize the business models in our sample in terms of the services they provide (**Figure 6.5**), the customer segments they target (**Figure 6.6**), and the revenue streams they leverage (**Figure 6.7**). Each DER technology category (i.e., demand response and energy management systems, electrical and thermal storage, and solar PV) is represented by a different color, and the size of each individual circle represents the number of business models in a given service, customer segment, or revenue stream category. For the sake of efficiency, these summary charts only show categories that contain four or more business models. Detailed analyses of each DER category are provided in **Appendix B**. However, the summary data discussed in this section offer some initial insights into the DER business model landscape.

Figure 6.5 summarizes the electricity services provided by the business models in our sample. The vast majority of these business models provide either a mix of firm capacity and operating reserve[40] products, or energy.

38 That is, if the system overall was "short" and had insufficient supply, agents that were consuming more than their scheduled quantity or producers who were producing less would be penalized at the single imbalance price, since both would be contributing to the overall system imbalance and thus to the costs of dispatching reserve or balance market resources. In this scenario, agents that were consuming less or generating more than scheduled would actually be reducing the overall system imbalance and lowering costs, and thus should not be charged an imbalance penalty (as they would be under dual imbalance price systems).

39 Note that despite the name, this category also includes "behind-the-meter" energy storage companies.

40 Note that operating reserves take different definitions depending on the system.

Figure 6.5: Electricity Services Summary

While operating reserve services are rather lucrative, the markets for these services are very small, typically making up less than 4 percent of total energy costs (Denholm et al. 2013; Rebours et al. 2007; Masiello, Roberts, and Sloan 2014). Intense competition for the provision of operating reserve services could prove challenging for demand response and energy storage technologies. On the other hand, payments for firm capacity can be quite significant (on the order of hundreds of dollars per megawatt-day) depending on desired reserve margins (tens of gigawatts in certain systems) (Crampton, Ockenfels, and Soft 2013; Spees, Newell, and Pfeifenberger 2013). As one might expect, the only business models in our dataset that provide energy in a distributed fashion are those that utilize solar PV. Where a business model is described as providing "operating reserves," the business can provide the range of primary, secondary, or tertiary reserves; other business models may provide only a single type of reserves (notably, secondary reserves).

Figure 6.6: Summary of Customer Segments

CUSTOMER SEGMENT(S)	DEMAND RESPONSE & EMS	ELECTRICAL & THERMAL STORAGE	SOLAR PV
C/I/M & Industrial <-> ISO/TSO/RTO	20	5	
DER Provider		16	6
C/I/M & Industrial	6	4	5
Regulated Utility & ISO/TSO/RTO		11	
Residential		5	6
Residential & C/I/M			9
C/I/M & Industrial <-> Regulated Utility	7		
C/I/M & Industrial & Load-Serving Entity			6
Industrial & Load-Serving Entity			5
Load-Serving Entity & Regulated Utility	4		

Figure 6.6 summarizes the customers targeted by the business models in our sample. Businesses that act as intermediaries between (and therefore service providers to) two agents are represented with a double-sided arrow (<->) connecting the two customer segments. Commercial, institutional, or municipal customers are represented by the abbreviation "C/I/M." The customer segment "DER Provider" indicates that the company is selling its products to businesses that then integrate one or more DERs at customer sites. The "Regulated Utility" customer segment refers to network companies—distribution, transmission, or both. This customer segment also refers to utilities that are vertically integrated such that they also function as the load-serving entity.

Figure 6.7 shows that most business models in our dataset are targeting end users directly. DR and EMS companies primarily target larger commercial and industrial customers, while solar PV and storage companies target more diverse customer segments. A smaller number of companies sell services directly to regulated utilities, independent system operators (ISOs), transmission system operators (TSOs), or regional transmission organizations (RTOs).

Figure 6.7: Summary of Revenue Streams

REVENUE STREAM(S)	DEMAND RESPONSE & EMS	ELECTRICAL & THERMAL STORAGE	SOLAR PV
Subscription Fees & Brokerage Fees	37		
Asset Sale & Lend/Rent/Lease		5	19
Asset Sale & Lend/Rent/Lease & Commodity Sale		11	11
Asset Sale		14	6
Asset Sale & Lend/Rent/Lease & Brokerage Fees		15	
Asset Sale & Commodity Sale			7

Figure 6.7 summarizes the revenue streams leveraged by the business models in our sample. Many companies leverage multiple revenue streams. The structure of the revenue streams often depends on the customer segments that are targeted. For example, many solar PV integrators offer consumers the option to directly purchase their PV system, lease the system, or take out a loan for the system. Notably, the vast majority of DR companies leverage subscription fees or brokerage fees (or a combination thereof) for their services. Such brokerage fees are typically structured as shared savings arrangements or as fees on payments earned in markets. Given the nascence of the storage market and the difficulties in predicting end-user savings, storage business models have so far relied heavily on asset sales or financing.

Using the same data as the data presented in this section, **Appendix B** explores each of the three DER technology categories in much greater detail. For each of the three DER categories, we defined a relatively small set of business model "archetypes;" the archetypes, in turn, encompass many individual business models. These archetypes circumscribe the key value creation and capture components of the most common DER business models that exist today.

6.5.3 Discussion and conclusions

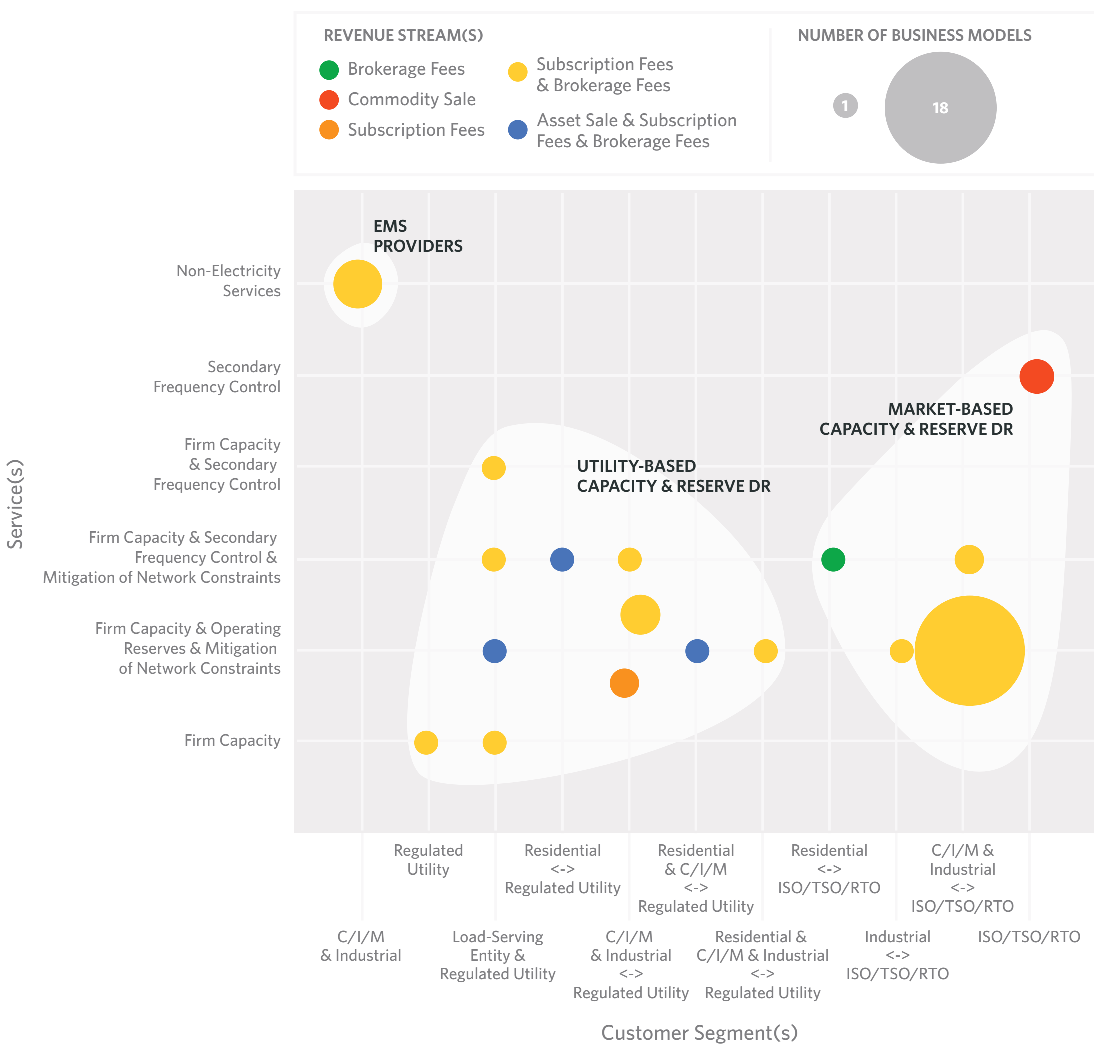
Our analysis here and in **Appendix B** leads to four observations about the DER business landscape that we believe have important implications for electric utilities, DER entrepreneurs, policymakers, and regulators.

6.5.3.1 DER businesses cluster into a few archetypes

Our first observation, after surveying a large number of DER business models, is that these myriad businesses can be clustered into a relatively small set of well-defined archetypes. One might expect, given the thousands of individual DER businesses operating today, that we would find a high degree of diversity in business models. However, our results show that today's DER business models cluster into a relatively small number of groupings that share key value-capture and value-creation components. For example, **Figure 6.8** illustrates the clustering of DR and EMS business models into three distinct categories: DR businesses that provide capacity and reserves to regulated utilities, DR businesses that provide the same services to organized markets and system operators, and EMS businesses that provide energy management services directly to commercial and industrial electricity consumers. **Figure 6.9** shows similar clustering among solar PV and solar-plus-storage business models.

This prevalence of clustering among DER business models suggests that the determinants of business success within a given archetype may include executional capabilities, culture, partnerships, and other activities that are not captured in our framework. Ontological differences alone do not appear to distinguish successful DER companies from their similarly structured rivals. In other words, the archetypal DER models identified in our analysis capture a set of structural elements that appear necessary, but not sufficient, for success.

Figure 6.8: Taxonomy of Demand Response and Energy Management System Business Models



Surveying a large number of DER business models suggests that these myriad businesses can be clustered into a relatively small set of well-defined DER business model archetypes.

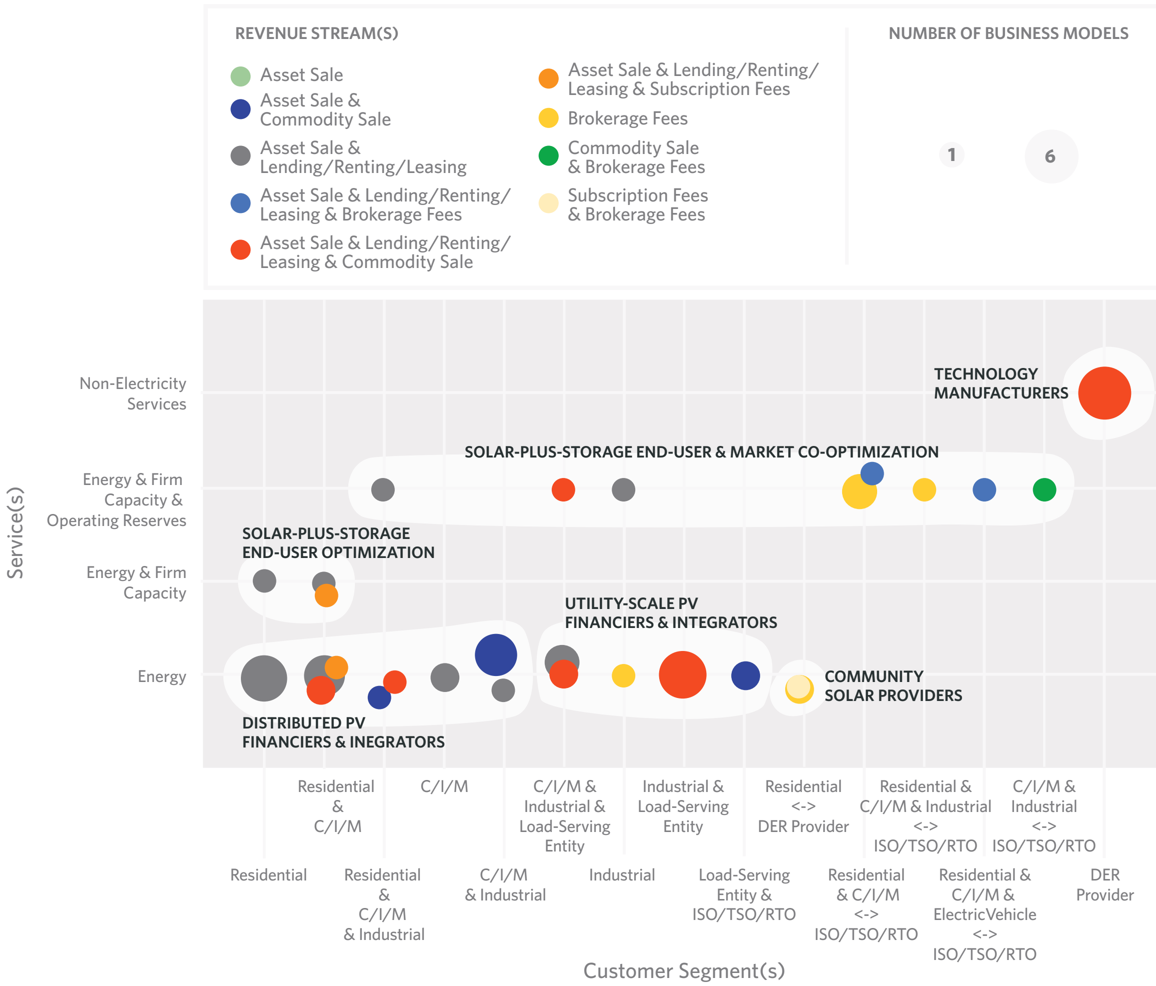
6.5.3.2 Policy and regulation are key drivers of the structure of DER businesses

Our second observation is that current DER business model archetypes appear to be driven more by regulatory and policy factors than by technological factors. The influence of policy and regulation on business structure is arguably most evident with the solar PV business model archetypes, but it applies to all of the technology categories we considered. As **Figure 6.9** illustrates, three solar PV archetypes have emerged (as well as several solar-plus-storage archetypes). The largest cluster consists of distributed PV finance and integration companies, which primarily leverage net metering or feed-in tariff policies and creative financial structures to maximize the value of available tax credits or other incentives and overcome capital barriers to end-user adoption. Utility-scale solar PV finance and integration companies build megawatt-scale solar installations and have largely emerged in response to state renewable portfolio standards in the United States and feed-in tariff policies in Europe. Finally, community solar providers have emerged to capitalize on economies of unit scale or to enable consumers located in areas unsuitable for their own solar PV production to procure solar from community solar projects. The community solar market is still relatively small (tens to hundreds of megawatts in the United States) and geographically restricted to policy-friendly environments. Without key policy and regulatory elements, we would not have seen the widespread emergence of these three solar PV archetypes. Similarly, the majority of business models within the "market-based capacity and reserve DR" archetype depend on capacity remuneration mechanisms that were established by regulators and legislatures.

The influence of policy and regulation on DER business models indicates that many DER businesses today depend to some degree on capturing opportunistic value created by current regulatory or policy regimes. As discussed in **Section 6.4**, this type of value capture accrues to private stakeholders but not to the electricity system as a whole, and thus typically results in economic transfers. If regulation and policy improve over time, opportunistic value capture should be minimized and DER businesses will evolve to derive a greater share of their revenues from the creation of transitory or fundamental value, which delivers real efficiency improvements to electric power systems.

The influence of policy and regulation on DER business models indicates that many DER businesses today depend to some degree on capturing opportunistic value created by current regulatory or policy regimes. If regulation and policy improve over time, opportunistic value capture should be minimized and DER businesses will evolve to derive a greater share of their revenues from the creation of transitory or fundamental value, which delivers real efficiency improvements to electric power systems.

Figure 6.9: Taxonomy of Solar PV and Solar-plus-Storage Business Models



6.5.3.3 DER business models are not static and evolve over time

Indeed, our analysis does not suggest that the DER business model landscape will continue to be defined and driven by regulation and policy indefinitely. Continued cost declines and technological innovation may well lead to markets for DERs that are less defined by policy and regulatory conditions and more by the delivery of fundamental or transitory value. Our analysis also does not suggest that the business model archetypes we have identified are here to stay. On the contrary, the fact that these archetypes are sensitive to regulation and policy environments suggests that they are very likely to change as new regulations and policies are introduced, old ones expire, and existing ones change. Thus, the third key observation to emerge from our analysis is that the business landscape for DERs 10, 15, or 20 years from now will likely look very different than the landscape of today.

6.5.3.4 DER business models compete to provide a limited range of electricity services

A fourth policy-relevant observation centers on the diverse range of DER business models that currently compete for market share in providing a limited set of commodity electricity services (**Figure 6.5**). DER businesses compete within archetypes (e.g., many DR companies compete to provide capacity and ancillary services), but as DER business models and technologies become increasingly mature we can expect increased competition across archetypes. For example, DR and energy storage technologies will likely compete to provide reserves and capacity-based services. We might also expect distributed PV businesses to face competition from utility-scale PV businesses. Because these business models generally provide commodity services, differentiation beyond price may be difficult to realize.

Competition between and among different types of DER businesses reinforces the importance of the structural decisions discussed in previous sections. To create a level playing field for competition between DER businesses and conventional providers of electricity services it will be important to establish sound industry structures and appropriate incentives for the non-discriminatory provision of electricity markets, networks, and system operation.

To create a level playing field for competition between DER businesses and conventional providers of electricity services it will be important to establish sound industry structures and appropriate incentives for the non-discriminatory provision of electricity markets, networks, and system operation.

References

ACER. 2015. *European Network of Transmission System Operators (ENTSO-E) Network Code on Emergency and Restoration*. Agency for the Cooperation of Energy Regulators. March 25, 2015.

Burger, S., J.P. Chaves-Ávila, C. Batlle, and I.J. Pérez-Arriaga. 2016. "The Value of Aggregators in Electricity Systems." *Center for Energy and Environmental Policy Working Paper 2016-001.* Cambridge, MA: Massachusetts Institute of Technology. January 2016.

CAISO. 2015. "Expanded Metering and Telemetry Options Phase 2: Distributed Energy Resource Provider (DERP) Draft Final Proposal." Sacramento, CA: California Independent System Operator.

Caramanis, M., N. Ntakou, W.W. Hogan, A. Chakrobarty, and J. Schoene. 2016. "Co-Optimization of Power and Reserves in Dynamic T&D Power Markets with Nondispatchable Renewable Generation and Distributed Energy Resources." P*roceedings of the IEEE 104(4): 807-836.* April 2016.

CEER. 2015a. *The Future Role of DSOs: A CEER Conclusions Paper.* Brussels, Belgium: Council of European Energy Regulators. July 2015.

CEER. 2015b. *CEER Advice on Customer Data Management for Better Retail Market Functioning.* Report C14-RMF-68-03. Brussels, Belgium: Council of European Energy Regulators. March 2015.

CEER. 2016. *CEER Benchmarking Report on Removing Barriers to Entry for Energy Suppliers in EU Retail Energy Markets.* Report C15-RMF-70-03. Brussels, Belgium: Council of European Energy Regulators.

CEDEC, EDSO, ENTSO-E, EURELECTRIC, and GEODE. 2016. *TSO – DSO Data Management Report*. www.eurelectric.org/media/285585/tso-dso_dm_rep-2016-030-0382-01-e.pdf.

CEER. 2016. "CEER Position Paper on the Future DSO and TSO Relationship." Brussels, Belgium: Council of European Energy Regulators. Accessed September 22.

SmartNet. 2016. "Towards (Close) Harmony: Options for Enhanced TSO-DSO Coordination." *SmartNet*. July 18. smartnet-project.eu/blog/towards-close-harmony-options-enhanced-tso-dso-coordination/.

Codognet, M-K. 2004. "The Shipper as the Architect of Contractual Relations in Access to Natural Gas Networks." *Proceedings of the 8th Annual ISNIE Conference.* Tuscon, AZ: The International Society for New Institutional Economics: 23.

Corneli, S. and S. Kihm. 2015. "Electric Industry Structure and Regulatory Responses in a High Distributed Energy Resources Future." *Future Electric Regulation Report No. 1.* Berkeley, CA: Lawrence Berkeley National Laboratory. November 2015.

CPUC. 2014. Order Instituting Rulemaking Regarding Policies, Procedures and Rules for Development of Distribution Resources Plans Pursuant to Public Utilities Code Section 769. Sacramento, CA: State of California Public Utilities Commission. August 2014.

Cramton, P., A. Ockenfels, and S. Stoft. 2013. "Capacity Market Fundamentals." *Economics of Energy and Environmental Policy 2:1–21*. doi:10.5547/2160-5890.2.2.2.

Cossent, R., L. Olmos, T. Gómez, C. Mateo, and P. Frias. 2011. "Mitigating the Impact of Distributed Generation Advanced Response Options." *International Transactions on Electrical Energy Systems 21(6):* 1869–1888.

De Martini, P. and L. Kristov. 2015. "Distribution Systems in a High Distributed Energy Resources Future." *Future Electric Utility Regulation Report No. 2.* Berkeley, CA: Lawrence Berkeley National Laboratory. October 2015.

Denholm, P., J. Jorgenson, T. Jenkin, D. Palchak, B. Kirby, and M. O'Malley. 2013. *The Value of Energy Storage for Grid Applications.* Golden, CO: National Renewable Energy Laboratory.

Eurelectric. 2013a. *Active Distribution System Management: A key tool for the smooth integration of distributed generation*. Brussels, Belgium: Eurelectric. February 2013.

Eurelectric. 2013b. *Power Distribution in Europe: Facts and Figures*. Brussels, Belgium: Eurelectric. February 2013.

Eurelectric. 2015. *Designing Fair and Equitable Market Rules for Demand Response Aggregation.* Brussels, Belgium: Eurelectric. March 2015.

European Commission. 2003. Directive 2003/54/EC Concerning Common Rules for the Internal Market in Electricity and Repealing Directive 96/92/EC. Brussels, Belgium: European Commission. June 2003.

European Commission. 2009. Directive 2009/72/EC Concerning Common Rules for the Internal Market in Electricity and Repealing Directive 2003/54/EC. Brussels, Belgium: European Commission. July 2009.

European Commission. 2010. "Interpretative Note on Directive 2009/72/EC Concerning Common Rules for the Internal Market in Electricity and Directive 2009/73/EC Concerning Common Rules for the Internal Market in Natural Gas: the Unbundling Regime." Commission Staff Working Paper. Brussels, Belgium: European Commission. January 2010.

European Commission. 2015. "Delivering a New Deal for Energy Consumers." Brussels, Belgium.

FERC. 1996. Order No. 888, Final Rule. Promoting Wholesale Competition Through Open Access Non-discriminatory Transmission Services by Public Utilities (Docket No. RM95-8-000). Washington DC: US Federal Energy Regulatory Commission. April 24, 1996.

FERC. 1999. Order No. 2000, Final Rule. Regulation Transmission Organizations (Docket No. RM99-20-000). Washington DC: US Federal Energy Regulatory Commission. December 20, 1999.

Friedrichsen, N. 2015. "Governing Smart Grids: The Case for an Independent System Operator." E*uropean Journal of Law and Economics 39(3): 553-572*. June 2015. First online July 2012.

Glover, J.D., M.S. Sarma, and T.J. Overbye. 2011. "Power Distribution." In *Power System Analysis and Design*, Chaper 14. 5th Edition. Cengage Learning. January 2011.

Hawaii PUC. 2014. The Commission's Inclinations on the Future of Hawaii's Electric Utilities: Aligning the Utility Business Model with Customer Interests and Public Policy Goals. Honolulu, HI: State of Hawaii Public Utility Commission. April 2014.

Hedman, W., M.C. Ferris, R.P. O'Neill, E.B. Fisher, and S.S. Oren. 2010. "Co-optimization of Generation Unit Commitment and Transmission Switching with N-1 Reliability." *IEEE Transactions on Power Systems* 25(2): 1052-1063. May 2010.

Hedman, W., S.S. Oren, and R.P. O'Neill. 2011. "Optimal Transmission Switching: Economic Efficiency and Market Implications." *Journal of Regulatory Economics* 40(2): 111-140. October 2011.

Hogan, W.W. 1997. "Transmission rights, through the looking glass: a false dichotomy between active and passive trading." *The Electricity Journal* 10(8): 86-91.

i3Connect. 2016. Cleantech Group. i3connect.com. Accessed March 1, 2016.

Ikäheimo, J., C. Evens, and S. Kärkkäinen. 2010. *DER Aggregator Business: the Finnish Case*.

Joskow, P.L. and R. Schmalensee. 1983. *Markets for Power: An Analysis of Electric Utility Deregulation*. Cambridge, MA: MIT Press. November 1983.

Jenkins, J.D. and I.J. Pérez-Arriaga. 2017. "Improved Regulatory Approaches for the Remuneration of Electricity Distribution Utilities with High Penetrations of Distributed Energy Resources. *The Energy Journal* 38(3). doi:10.5547/01956574.38.3.jjen.

Kranning, M., E. Chu, J. Lavaei, and S. Boyd. 2013. "Dynamic Network Energy Management via Proximal Message Passing." *Foundations and Trends in Optimization* 1(2): 70-122.

Littlechild, S.C. 2000. "Why We Need Electricity Retailers: A Reply to Joskow on Wholesale Spot Price Pass-through." Cambridge, UK.

Littlechild, S.C. 2009. "Retail Competition in Electricity Markets — Expectations, Outcomes and Economics." *Energy Policy* 37:759–63. doi:10.1016/j.enpol.2008.07.025.

Masiello, R.D., B. Roberts, and T. Sloan. 2014. "Business Models for Deploying and Operating Energy Storage and Risk Mitigation Aspects." *Proceedings of the IEEE* 102:1052–64. doi:10.1109/JPROC.2014.2326810.

Mallet, P., P. Granström, P. Hallberg, G. Lorenz, and P. Mandatova. 2014. "Power to the People. European Perspective on the Future of Electric Distribution." *IEEE Power and Energy Magazine*: 51-64. doi: 10.1109/MPE.2013.2294512.

Mansur, E.T. and M.W. White. 2012. "Market Organization and Efficiency in Electricity Markets." Working paper. January 13, 2012.

Millán, J. 2006. *Entre el Estado y el Mercado. Tres décadas de reformas en el sector eléctrico de América Latina* (in Spanish). Banco Interamericano de Desarrollo, 2006. ISBN: 1597820288.

MP2 Energy. 2016. First-Class MPACT Energy Management. www.mp2energy.com/demand-response. Accessed March 1, 2016.

NYDPS. 2014. Case 14-M-0101 Proceeding on Motion of the Commission in Regard to Reforming the Energy Vision: DPS Staff Report and Proposal. Albany, NY: State of New York Public Service Commission, Department of Public Service.

NYDPS. 2015a. Order Adopting Regulatory Policy Framework and Implementation Plan. Case 14-M-0101: Reforming the Energy Vision. Albany, NY: State of New York Public Service Commission, Department of Public Service. February 2015.

NYDPS. 2015b. "Staff White Paper on Ratemaking and Utility Business Models." Albany, NY: State of New York Public Service Commission, Department of Public Service. July 2015.

Opower. 2016. Capacity Savings Potential of Behavioral Demand Response. www.opower.com/bdrpotential/#us. Accessed March 1, 2016.

Parsons, J., C. Colbert, J. Larrieu, T. Martin, and M.E. Mastrangelo. 2015. "Financial Arbitrage and Efficient Dispatch in Wholesale Electricity Markets." Center for Energy and Environmental Policy Research, Working Paper 2015-002. Cambridge, MA: Massachusetts Institute of Technology. February 2015.

Pérez-Arriaga, I. 2011. "Managing Large Scale Penetration of Intermittent Renewables." MIT Energy Initiative Symposium. Cambridge, MA: Massachusetts Institute of Technology.

Pérez-Arriaga, I., S. Ruester, S. Scwenen, C. Batlle, and J-M Glachant. 2013. *From Distribution Networks to Smart Distribution Systems: Rethinking the Regulation of European Electricity DSOs.* THINK Project, Topic 12. Florence, Italy: European University Institute.

PJM. 2014. Retail Electricity Consumer Opportunities for Demand Response in PJM's Wholesale Markets.

PJM. 2015. Demand Response Fact Sheet. June 29, 2015.

Rebours, Y.G., D.S. Kirschen, M. Trotignon, S. Rossignol. 2007. "A Survey of Frequency and Voltage Control Ancillary Services - Part II: Economic Features." *IEEE Transactions on Power Systems* 22:358–66. doi:10.1109/TPWRS.2006.888965.

Ruiz, P.A., E.A. Goldis, A.M. Rudkevich, M.C. Caramanis, C.R. Philbrick, and J.M. Foster. 2016. "Security-constrained transmission topology Control MILP Formulation Using Sensitivity Factors." *IEEE Transactions on Power Systems*, doi: 10.1109/TPWRS.2016.2577689.

Slot, T., H. Dijk, R. Haffner, and A. van der Walle. 2015. *Options for Future European Electricity System Operators*. Brussels, Belgium: European Commission. December 2015.

Smart, D.C.C. 2015. Building a Smart Metering Network for Great Britain. London: Smart Data Communications Company, Ltd. October 2015.

Smart Grid Task Force. 2015. *Regulatory Recommendations for the Deployment of Flexibility*. Report of Expert Group 3 to the European Commission. Brussels, Belgium.

SmartNet. 2016. "Towards (Close) Harmony: Options for Enhanced TSO-DSO Coordination." SmartNet. July 18.

Spees, K., S. Newell, and J.P. Pfeifenberger. 2013. "Capacity Markets - Lessons Learned from the First Decade." *Economics of Energy and Environmental Policy 2*. doi:10.5547/2160-5890.2.2.1.

Tabors, R., G. Parker, P. Centolella, and M. Caramanis. 2016. *White Paper on Developing Competitive Electricity Markets and Pricing Structures*. Tabors Caramnis and Rudkevich, Inc., prepared for the New York State Energy Research and Development Authority and New York State Department of Public Service. April 2016.

Tirole, J. 1988. *The Theory of Industrial Organization*. Cambridge, MA: MIT Press. August 1988.

U.S. DOE. 2015. Data Privacy and the Smart Grid: A Voluntary Code of Conduct (VCC). Washington DC: U.S. Department of Energy. January 2015.

Wellinghoff, J., J. Tong, and J. Hu. 2015. *The 51st State of Welhuton: Market Structures for a Smarter, More Efficient Grid*. 51st State Initiative. Solar Electric Industry Association. February 27, 2015.

Wilson, R. and B. Biewald. 2013. *Best Practices in Electric Utility Integrated Resource Planning: Examples of State Regulations and Recent Utility Plans*. Cambridge, MA: Synapse Energy Economics and Regulatory Assistance Project. June 2013.

PART 2: A FRAMEWORK FOR AN EFFICIENT AND EVOLVING POWER SYSTEM

07

The Re-evolution of Short- and Long-Term Electricity Market Design

7.1 Toward an Electricity Market for All

One of the key objectives for power system restructuring was the establishment of wholesale markets that allowed competition in the generation of electricity. Wholesale markets replaced vertically integrated utilities with an open-access model, whereby various generating companies could compete to meet a share of electricity demand.

This transformation required replacing centralized decisions with competitive processes, and creating multiple markets with different time scales. Today's wholesale electricity markets comprise a series of sequential trading opportunities — from long-term markets that are cleared years in advance, to very short-term balancing markets (namely in the European context) and so-called real-time markets in the United States — with the ultimate goal of meeting electricity demand in an efficient, secure, and environmentally acceptable way. These wholesale market mechanisms allow market agents to manage their risk exposure and support the chain of decisions necessary to match electricity production and consumption in the most economically efficient manner. This supply–demand balance, however, is not the result of interactions between market agents alone. Regulators, system operators, and market designers have historically intervened to enhance investment signals via capacity mechanisms and/or technology-specific incentives (e.g., so-called clean energy technology support mechanisms). Such interventions have been used, for example, to ensure the availability of sufficient resources to cope with contingencies that could endanger the reliable operation of the power system.

Given the complexity of electric power systems, electricity wholesale market design remains a topic of heated discussion, even after decades of experience. Design decisions for markets and investment support mechanisms are further complicated by a host of new challenges, including adapting to technological developments with respect to distributed energy resources (DERs) and information and communications technologies (ICTs). DERs present specific challenges to market design due to their small size (compared with centralized resources) and because they are often affected by local distribution network constraints that are not typically considered in wholesale markets. Moreover, some DERs — such as electric vehicles, batteries, and demand response — represent a new class of technologies. Each has its own set of operating constraints and, in most cases, little experience with market integration. Most DERs have not played a significant role in electricity markets previously, but now have the potential to expand very quickly — which means that electricity market design needs to be ready ahead of time.

The goal of this chapter is to address the challenges that currently inhibit efficient integration of all types of resources — centralized and distributed, conventional and emerging — into electricity markets. Achieving this requires a series of new or updated market mechanisms and platforms to generate price signals that will drive efficient operational and investment decisions. An in-depth review of existing subsidy mechanisms and a transition to more market-compatible schemes for better integrating clean energy technologies are also needed. The reimagining of support mechanisms is all the more critical now that some renewable and distributed technologies are reaching competitive levels — comparable to those of more conventional generating technologies — in some contexts.

This chapter is divided into two parts. The first (**Section 7.2**) focuses on the design challenges associated with short-term market mechanisms (i.e., markets that range from day-ahead to real time) and analyzes options for efficiently integrating market trading and system operation. This section also examines reserve and balancing markets, which are intimately related to short-term markets. The second part (**Section 7.3**) reviews needed changes to long-term capacity mechanisms and clean energy support mechanisms. Specifically, the focus of this section is on redesigning these mechanisms to enhance their compatibility with existing markets and minimize associated market distortions.

7.2 Leveling the Playing Field: Short-term Markets for an Efficient Integration of the New Diversity of Resources

The short-term market mechanisms examined in this section are the so-called energy markets: day-ahead and intraday markets. Prices in these markets are frequently used as a reference for many other transactions, as well as for the reserves and imbalance markets, which operate closer to real time and are thus closely linked to operational security conditions. For reference, we rely on the US and European electricity market models, which represent two different but increasingly converging design archetypes. The most prominent difference between these two models is found in the division of responsibilities between the power exchange and the system operator. However, it is not possible to attribute all differences to this foundational decision. The next sections describe several specific market design elements while highlighting differences and similarities.

7.2.1 Energy (day-ahead and intraday) market mechanisms

Energy (day-ahead and intraday) mechanisms are a sequence of short-term market mechanisms that operate up to one day prior to the actual delivery of electricity. They are designed to allow market agents and the system operator to match supply and demand in a secure and efficient manner. This section focuses on organizing markets that are compatible with system operation requirements while also producing more accurate market signals. It then introduces the concept of market "granularity" to describe the temporal and geographical resolution achieved by market products and price signals. The section concludes with a discussion of the elements of auction design. For each of these topics we focus on changes that are necessary to achieve an efficient co-integration of both conventional generation technologies and DERs.

7.2.1.1 Short-term markets

Design of an efficient short-term market sequence

The sequence of short-term markets present in today's electric power sector was strongly conditioned on the operational procedures of centralized generating plants prior to restructuring. These procedures, in turn, were tailored to the characteristics of demand profiles and generating technologies at that time. Once it was generally determined that one day in advance of real time was the optimal timing to forecast (with reasonable accuracy) demand consumption, and to schedule thermal plant commitments, the vast majority[1] of electricity systems implemented day-ahead markets.

Several markets are needed between the time of electricity delivery and the day before so as to incorporate the different planning horizons of all available supply and demand resources.

Today, power systems and the factors that condition optimal dispatch procedures are changing rapidly. Price responsive demand, intermittent energy resources, greater constraints on the operation of thermal generators,[2] and other trends mean that day-ahead may no longer be the optimal length of time for key commitment decisions. Intraday markets generate price signals between the day ahead and the time of electricity delivery and are necessary for market agents to efficiently manage increasingly variable generation and/or demand patterns.

For the same reason, it may seem desirable to extend short-term markets to their extreme and permit trading until the very last second before the delivery of electricity. This, however, is not possible in practice. Coordination between markets and system operation becomes more challenging as markets move closer to real time. Therefore, it is commonly accepted that a time window is necessary between market closure and physical dispatch in order for the system operator to carry out reliability procedures.

In most European power systems, this time window begins after the commonly named "last gate closure," and occurs between a few hours to a few minutes before real time, depending on the jurisdiction. The transition between market and system operation has traditionally been quite clear due to the differentiation of functions between power exchanges (PX) and transmission system operators (TSOs); two different entities that typically perform these tasks separately but in close coordination. The last gate closure has two implications: It is the final opportunity for market agents to submit updated bids on the PX to balance their market positions, and it marks the point after which only the system operator can adjust production and consumption schedules. The trend of late has been to move last gate closure as close to real time as practically possible.[3]

In the United States, the division between trading and system operation is less defined. A single entity, the independent system operator (ISO), manages the day-ahead market and unilaterally operates the power system thereafter, rescheduling generation units as it deems necessary. In this context, a last gate closure in the European sense cannot be clearly defined. For most agents, the last bidding opportunity is the day-ahead market, although, under certain conditions, some resources are allowed to re-bid later. Following the day-ahead market, the ISO continues to adapt generation schedules through various mechanisms, the latest

1 In a handful of systems characterized by large hydro-reservoir capacity (namely Brazil and to a lesser extent Norway), the commitment of generating units is primarily determined on a longer-term basis than the day ahead.

2 Fluctuations in renewable energy production require more varying output from thermal generators, pushing them closer to their operating limits. Also, some power systems have recently increased their reliance on natural gas due to low gas prices. This has led to increased congestion in gas pipelines, which can cause variations in short-term gas prices that need to be translated to electricity markets.

3 An issue presently under discussion in some jurisdictions is whether there is a need for the TSO to intervene earlier in time, a question that is discussed later in this chapter.

being the "real-time market,"[4] which provides dispatch instructions a few minutes before delivery in five-minute intervals. These dispatch instructions are partly based on bids (updated if allowed), but also on the ISO's forecast of load and renewable energy production.

In practice, this implies that — in the United States — different resources are effectively subject to different last gate closures, on grounds that some participants, such as consumers and renewable producers, are more efficiently integrated through ISO forecasting rather than through active market participation. While this may be the case, especially when the engagement of consumers and the maturity level of renewable technologies are low, ample evidence suggests that renewable generators and demand response providers can significantly improve their forecasts and overall performance when faced with cost-reflective price signals.[5] If it can be shown that there is value in transferring forecasting/scheduling responsibility from some market participants to the ISO, the ISO should perform this task as a mere intermediary, passing all costs incurred for forecasting and balancing of deviations on to market participants. This way, market agents could determine which arrangement is truly more efficient and would have the choice to transfer their forecasting responsibility to the ISO — or not.

Efficient intraday prices and market designs

Market designs affect price signals produced in the intraday timeframe, (i.e., between the day-ahead and real-time markets) and result in different balances between electricity trading and system operation.

In Europe, where intraday markets play a significant role, current discussions are focused on how to implement these markets most efficiently. Two market designs have generally prevailed: continuous intraday markets and intraday auctions that occur at prescribed times. Both of these designs allow generators to make adjustments to their energy schedules, which helps to improve the efficiency and accuracy of the initial day-ahead market.

Continuous markets allow market agents to submit bids and offers to modify schedules at any time before last gate closure. Each bid and each offer consists of a quantity of energy with a price (a "price–quantity pair"). Continuous markets are cleared on a first-come, first-served basis — in essence, they arrange a bilateral transaction whenever a buy order (bid) is matched with a sell order (offer). Conversely, auction-based intraday market sessions resemble day-ahead auctions, albeit with several market sessions throughout the day. The auctions are centrally cleared with marginal prices. While continuous markets provide greater flexibility to market agents, as transactions can be made at any time, intraday auctions have several advantages from the perspective of efficient price formation. Auction-based intraday markets improve liquidity,[6] in part by concentrating all transactions that reflect the accumulated events that occurred since the previous session. Thus, the prices that result from these auctions provide a more reliable reference price compared to prices in an illiquid continuous market, which might not reflect system marginal cost at a given time. Liquidity is of particular relevance in markets that are dominated by a small number of large companies, as is the case in many restructured power systems. In such markets, low liquidity is particularly challenging for new and small market entrants, such as DERs and small aggregators. Small market participants would further benefit from intraday auctions, since they usually lack continuous 24/7 trading desks. Moreover, intraday auctions have the advantage of implicitly pricing transmission capacity. In a continuous market, by contrast, transmission capacity is allocated at zero cost (in other words, the bid–offer spread is zero) on a first-come, first-served basis.

4 The "real-time market" involves a number of processes that go beyond what it is commonly understood as an exchange. In most cases, US real-time markets provide dispatch instructions before real time, but prices are computed in separate *ex post* processes. Also, as described by Helman et al. (2008): "In the real-time market, the system operators manage congestion on a minute-to-minute basis in part through auction prices and in part through non-market operating decisions."

5 For example, wind forecast errors have gradually decreased due to the exposure of these generators to imbalance prices (Herrero et al. 2016).

6 Market liquidity can be defined as a market's ability to allow assets to be bought and sold at stable prices. Liquidity can be quantified in different ways. For example Hageman and Weber (2015) develop a trading-volume-based analysis of liquidity in European intraday markets that also highlights an important consideration: To properly assess liquidity in different power systems, trading volumes must be compared with a benchmark expected volume, which does not depend on the market design but on the characteristics of each system. For example, in a system with high renewable energy production, the expected trading volume in intraday markets is also high due to forecast errors. If this system has higher trading volumes than a mostly thermal system, the difference cannot be attributed to market design.

Considering the advantages and drawbacks of both alternatives, we believe that a series of intraday auctions is superior to continuous trading. Furthermore, even in systems where the regulatory choice is to use, or continue to use, continuous trading, it is possible and potentially advantageous to combine continuous trading with intraday auctions in a hybrid design. This approach was recently implemented in Germany.[7]

Where intraday markets are implemented, such as in Europe, improvements in market liquidity are especially critical for DERs and other small market participants. A series of discrete, auction-based intraday market sessions can contribute to this end.

Determining the frequency of intraday auctions presents an interesting trade-off. Running frequent auctions can lower the cost of making dispatch corrections, since system operation becomes more inflexible as it approaches real time. For example, consider a case where the output of a renewable energy source is lower than anticipated in the day-ahead forecast and other generators are required to make up the resulting deficit. If the renewable energy generator is presented with frequent opportunities to correct its forecast, it will have an economic incentive to do so as early as possible, when many generators are still able to respond. On the other hand, if the renewable generator cannot make its correction until only a few minutes ahead of time, only fast-start units or expensive flexible generators will be able to compensate for the deficit. That said, if auctions are held too frequently, liquidity in each auction is reduced. Thus, there is also value in grouping all deviations before a correction is made. Note that this is the same argument discussed previously regarding continuous trading.

7 As discussed in Neuhoff et al. (2016), the addition of an intraday auction to the German continuous intraday market increased liquidity and led to greater market depth (revealing market participants' capacity/flexibility) and reduced price volatility.

Electricity markets in the United States use a two-settlement system (a day-ahead settlement and a real-time settlement) without any intraday settlement periods.[8] US ISOs do, however, supplement the day-ahead market with several subsequent processes aimed at committing additional units when considered necessary. These intraday processes can provide commitment instructions to generators, but they do not produce financially binding prices at the time they are executed, and hence do not have an immediate effect on the settlement of electricity production. Participants who want to change their forward financial positions during the day may have to rely on the bilateral market. Afterwards, the ISO uses the real-time market settlement to allocate aggregated re-scheduling costs among market agents who deviated from the day-ahead program. Under the two-settlement system, market agents must pay the same amount for every megawatt-hour (MWh) deviation with respect to the day-ahead market, irrespective of when the deviation is corrected. This approach does not allow market settlements to reflect the different costs of dispatch corrections at different times (for example, a close-to-real-time correction would require very flexible resources, and would therefore be more costly than a correction that is made hours ahead).

US markets should be enhanced by complementing intraday rescheduling mechanisms with corresponding price signals. This would provide agents with incentives to improve their forecasts and short-term performance.

The lack of intraday settlements in US electricity markets is very much related to another design choice discussed previously: ISOs' use of their own load and renewable production forecasts as the basis for intraday and real-time decisions, as opposed to relying on updated bids. For

8 The California ISO features intraday markets, although with a limited scope. The hour-ahead scheduling process is binding for imports, and the fifteen-minute market is binding for some renewable resources.

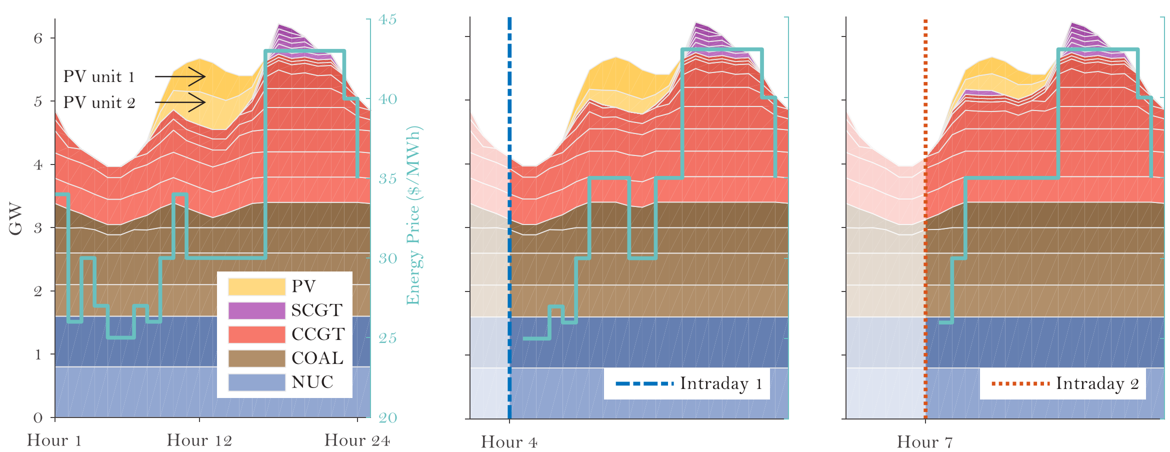
Box 7.1: Enhancing Intraday Price Signals in US ISO Markets — Results of a Modeling Exercise

(Herrero et al. 2016)

ISOs could benefit from receiving forecast updates directly from producers, since producers can better account for local conditions, but only to the extent that producers are appropriately rewarded for the value of accurate forecasting. Although receiving forecast corrections as soon as possible has clear economic value, under the two-settlement system used in ISO markets, this value is not fully disclosed and costs are not properly allocated. To improve the efficiency of the market, each intraday commitment process should be accompanied by its own intraday settlement. This leads to the proposed multi-settlement system, which permits allocation of intraday costs according to cost causality principles and creates efficient signals for market agents to improve forecast accuracy.

In our modeling exercise, we compare outcomes for two renewable energy units that deviate from their forecasts at the same time to illustrate the incentives produced by intraday settlements versus the two-settlement system (**Figure 7.1**). Unit 1 corrects its forecast later (**Figure 7.1c**), while Unit 2 performs a better forecast and is able to make a correction sooner (**Figure 7.1b**). The higher cost of the later correction is captured by intraday prices. However, if the two-settlement system is used, these two deviations are lumped together and both units pay the real-time price for the full difference between the day-ahead forecast and the real-time deviation. In this way, the two-settlement system loses the more precise intraday signal.

Figure 7.1: Sequential Dispatches as PV Output Is Corrected and Corresponding Intraday Prices



(a) Day-ahead dispatch (b) Intraday 1 dispatch (c) Intraday 2/final dispatch intraday prices to effectively capture new information, participants should be allowed to revise offers and bids during the operating day. We do not, however, suggest that the current US short-term market sequence should evolve to resemble the EU's intraday markets. Rather, as discussed in a modeling exercise by Herrero et al. (2016), we argue in favor of an alternative settlement system that produces intraday price signals (similar to the price signals from European intraday markets), but that is compatible and consistent with the current (centralized) organization of ISO markets. (The modeling exercise by Herrero et al. is described in **Box 7.1**)

Other design elements of these markets, such as bidding formats and the definition of settlement periods, are equally important; they are analyzed separately in subsequent sections.

7.2.1.2 Granularity of electricity products

Electricity products can exhibit granularity in two dimensions: temporal and spatial. Temporal granularity, for example, distinguishes electricity that is traded in annual baseload products in long-term markets from electricity that is traded in hourly blocks in day-ahead markets. Similarly, spatial granularity varies for electricity that is priced at the system level, in an aggregated zone, or at a particular node. This section investigates the concept of temporal and spatial granularity separate from the perspective of dispatch instructions and price signals.

Dispatch instructions with high temporal granularity

Electricity markets necessarily divide continuous time into discrete segments. For example, day-ahead markets typically break electricity production and consumption into hourly blocks. Real-time markets in the United States use five-minute dispatch instructions, while many EU-organized intraday markets still use hourly products. (Germany, which uses fifteen-minute products, is one of the exceptions.) While the use of long-duration electricity products has, to some degree, always been inefficient, the integration of DERs reinforces the need for shorter products. Many new DER technologies are especially well suited for generating, consuming, or shifting consumption of electricity in short intervals. Examples include small electricity storage systems or automated (smart) demand management systems, such as systems that control the cycling of air conditioning units.

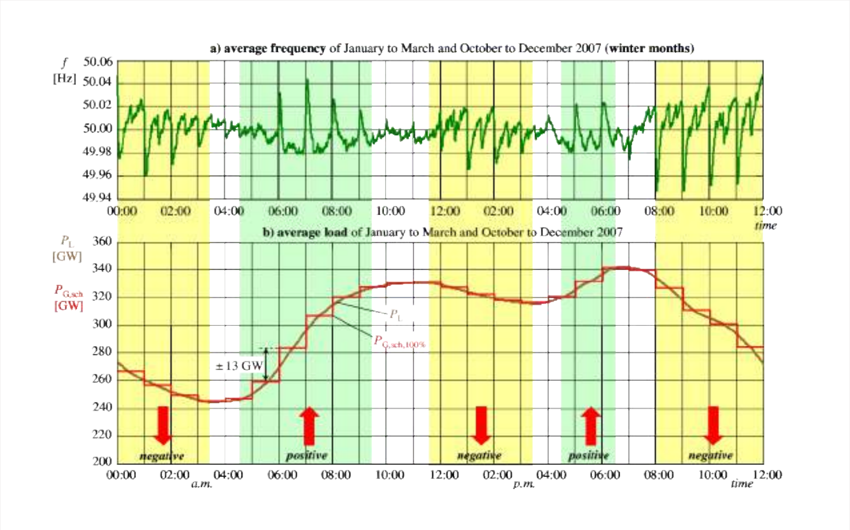
Electricity markets, particularly those in the European Union, should use more granular (i.e., shorter duration) electricity products when close to real time.

The advantages of shorter-duration electricity products are more significant as markets move closer to real time. For example, the efficiency gains of implementing a five-minute product in a day-ahead market are less significant than if the same product is implemented in a shorter-term intraday market — in fact, the five-minute product could even be detrimental in the day-ahead market if it overly complicates bidding and clearing processes. Hence, market sequences should introduce greater temporal granularity as markets approach real time, while also considering trade-offs in terms of other important market design elements and technical challenges (e.g., metering issues, transaction costs, etc.).

Box 7.2: One Effect of Low Temporal Granularity

Among the most visible effects of breaking electricity production into hourly blocks is the deviation of system frequency at the transition from one hour to the next. This leads to an inefficient use of operating reserves that could be avoided with more continuous representations of demand — for example, via more granular dispatch intervals.[9]

Figure 7.2: Frequency Excursions in the European Power System



Source: Weissbach and Welfonder (2009)

9 Another alternative to deal with this inefficiency would be to convert constant hourly schedules into continuous piecewise linear schedules.

Box 7.3: US Federal Energy Regulatory Commission Order 825

Although it applies only to generation connected at the transmission level and not to distributed resources, US Federal Energy Regulatory Commission (FERC) Order 825, which was issued on June 16, 2016, aims to improve the alignment of prices and dispatch instructions:

"The Federal Energy Regulatory Commission is revising its regulations to address certain practices that fail to compensate resources at prices that reflect the value of the service resources provide to the system, thereby distorting price signals, and in certain instances, creating a disincentive for resources to respond to dispatch signals. We require that each regional transmission organization and independent system operator align settlement and dispatch intervals by: (1) settling energy transactions in its real-time markets at the same time interval it dispatches energy; (2) settling operating reserves transactions in its real-time markets at the same time interval it prices operating reserves; and (3) settling intertie transactions in the same time interval it schedules intertie transactions." (FERC 2016)

Alignment of prices and dispatch instructions

More granular dispatch instructions alone do not guarantee a more efficient outcome if price signals are not aligned with the dispatch instructions. We focus here separately on the temporal and spatial resolution of prices. Cases of low resolution commonly occur in both dimensions and can lead to inefficiencies.

Temporal resolution: In some cases, energy is measured and settled in relatively long intervals for reasons of simplicity, even if dispatch instructions are given at shorter intervals. This is the case in some US ISOs, for instance, where five-minute prices are averaged to hourly prices for real-time market settlement; it is also the case for balancing markets in Europe, which use longer settlement intervals than the granular timescale of reserves deployment (imbalance settlement is analyzed in more detail in **Section 7.2.2.2**).[10] In the vast majority of power systems, settlement intervals are even longer for consumers at the distribution level (e.g., monthly), mainly due to constraints imposed by metering equipment. Long settlement periods for retail customers provide inefficient signals that fail to incentivize either smart demand management or the use of other DERs that could lower or shift net demand (see **Chapter 4**).

Many power system operators, concerned with finding ways to attract flexible generating resources, have proposed the implementation of "flexibility products" that would allow them to preferentially procure electricity from the most flexible resources. However, "flexibility" is just a concept — it is not really a service, and its value cannot be decoupled from the electricity price by implementing a separate product. To reflect the value of flexibility for the power system, the granularity of electricity prices should be aligned with dispatch instructions and reflected in reserve product design. At the same time, metering equipment and communication protocols should be updated to allow for settlement at the same resolution. Energy settlements with higher granularity can show the increased value of electricity in some periods over others, while a longer settlement period dilutes this signal and does not reward flexible resources for their higher value.

10 European system operators use different length settlement periods: fifteen minutes (Germany, Belgium, the Netherlands), thirty minutes (France, UK), and one hour (the Nordic countries, Spain, Portugal).

Spatial resolution: Prices can also inform the value of energy at different locations in the system. Two different approaches to pricing can be used to reflect spatial granularity.

Locational marginal pricing, or nodal pricing, is the prevailing paradigm in US wholesale electricity markets. Energy is priced differently at each bus of the transmission network. When price differences exist between two connected nodes, those differences reflect the effect of congestion and network losses that may make electricity injection in one bus more valuable than in the other.

Flexibility is not a service, it is a characteristic of energy resources and its value should be reflected in granular energy prices.

Zonal pricing, the prevailing approach in Europe, is essentially a simplification of nodal pricing. The power system is artificially divided into zones within which little congestion is expected to occur. As a result, prices reflect only the previously expected, most relevant cases of transmission congestion between zones. Within zonal pricing there are different degrees of spatial resolution. For example, some EU member states only feature one price zone (as in Germany or France) — this represents the lowest possible spatial resolution. Others consider a larger number (for example, Italy has 10 price zones).

Stakeholders and policymakers have long debated the merits of nodal and zonal pricing depending on the characteristics of the power system. The advantages and disadvantages of each approach typically revolve around a number of issues, including, among others, short- and long-term efficiency, implementation costs, computation burden, hedging complexity, long-term markets liquidity, compatibility with avoiding

DERs increase the need for nodal pricing in wholesale markets, while simultaneously making it necessary to expose DERs to price signals with the same temporal and spatial resolution.

geographical consumer discrimination, and institutional and governance issues. Different analyses have shown how the gains from implementing nodal pricing — in the context of congested systems — clearly outweigh implementation costs.[11] At the same time, zonal pricing is usually deemed appropriate in strongly meshed systems with easy-to-forecast power flows because it simplifies zones identification. The arguments for zonal pricing often rest on the need to create larger bidding zones so as to reinforce liquidity and competition, and induce investment signals that are more stable for generators and less discriminatory toward demand resources. However, many of the points often made in favor of zonal pricing are, to say the least, highly debatable. For example, in nodal markets with long-term hedging, market liquidity is not necessarily affected — financial transmission rights (FTRs) and hubs[12] play a key role here. On the other hand, nodal prices can be compatible with whatever geographical resolution is used in end-consumer prices (in other words, retail customers do not necessarily need to face nodal prices in nodal markets).

The fact is that large shares of dispersed variable generation are already causing congestion problems to increase, for reasons that are sometimes localized and sometimes structural. Growing congestion between, for example, northern and southern Germany makes power flows less predictable and makes it more difficult to define effective zones (**Figure 7.3**). Also, increased DER participation will improve liquidity in electricity markets, but requires ever more localized investment signals.

We recommend a shift toward providing more detailed and robust spatial resolution at the wholesale level, and in this respect nodal prices represent the ideal framework if workable. Regarding workability, clearly, the most important hurdles for implementing nodal prices stem from institutional and governance issues. Market integration within the same country (the case in the United States), is challenging because of the need to coordinate different states with different institutions (including different ISOs) and regulations. But the challenge is an order of magnitude greater when it comes to integrating markets across different countries, as in the case of the European internal energy market.

From wholesale to retail granularity

As mentioned above, a nodal market could be compatible with average (zonal) pricing for low-voltage consumers. Specifically, nodal prices are often averaged, both in the United States and Europe, to compute retail tariffs with lower (effectively, zonal) resolution. However, price averaging at the distribution level will be increasingly inefficient (as we have already pointed out, this is also true for temporal averaging), as low-voltage consumers start responding to price signals, both in the short and the long term.

Note that consumers can be exposed to market prices through tariffs, or through direct market participation. If consumers become active market participants, the logic used to clear markets would have to account for the division of the market into two granularities (nodal prices for generation and averaged zonal prices for consumers). This would create additional implementation challenges and inefficiencies. Clearing a market where participants face different prices gives rise to arbitrage opportunities and is computationally complex.[13] Computational complexity is also a source of inefficiency to the extent that it limits the introduction of other necessary market design elements (the next section, for example, discusses the need for more complex bidding formats). Ultimately, as the emergence of DERs erodes the traditional separation between generation and demand, and between transmission and distribution, exposing all market

11 See, for instance, Neuhoff and Boyd (2011).

12 In US nodal markets, trading hubs that use an average price across a set of nodes have provided adequate liquidity. This is a suitable trade-off to avoid liquidity problems as a result of moving toward a higher level of spatial granularity.

13 The Italian power exchange is an example of a market where market participants face different prices. Generation receives zonal prices in the day-ahead market, while for political reasons the same weighted average price (known as "PUN" for Prezzo Unico Nazionale or "single national price" in Italian) is used for all demand (including large consumers actively bidding in the market). The market clearing algorithm explicitly considers this policy, which results in a very complex problem given that the algorithm must search for prices that ensure supply–demand equilibrium, while also making the PUN equal to the weighted average of the zonal prices and satisfying transmission constraints, in an iterative fashion, with no mathematical guarantee that a solution exists.

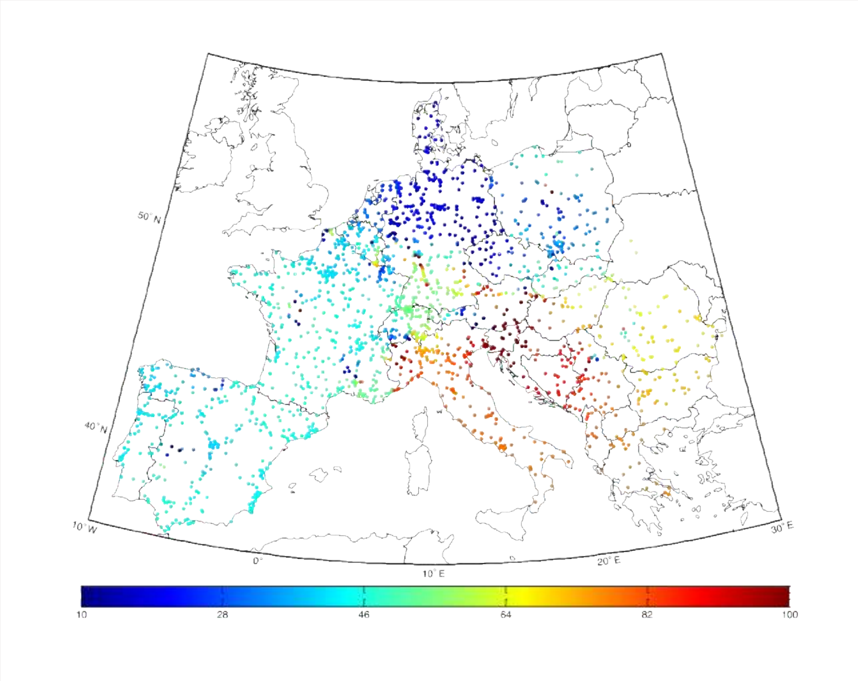
Box 7.4: Zonal vs. Nodal Pricing as Affected by Wind Penetration

Neuhoff et al. (2013) conduct a numerical assessment of the effects of nodal pricing in Europe under different wind penetration scenarios. The results suggest that cost savings of €0.8–€2.0 billion per year and an increase in the use of international transmission capacity of up to 34 percent can be achieved with nodal pricing. Two main reasons are given for the inefficiency of zonal pricing:

First, zonal pricing should reflect transmission constraints between countries but, in many cases, these constraints are not actually associated with lines between countries — rather, they reflect congestion within a country. This occurs because TSOs sometimes limit international flows to avoid domestic congestion.

Second, nodal pricing simulations illustrate that congestion and price patterns would vary considerably between different scenarios for wind penetration. Therefore, approaches aimed at defining price zones within countries are not suitable to address internal congestion, as the zones would either have to vary depending on system conditions (which would be impractical for contracting purposes) or be small (and thus be essentially equivalent to nodal pricing).

Figure 7.3: Nodal Pricing for a High Wind Penetration Scenario (shown as euros per megawatt-hour in the colored scale)



Source: Neuhoff et al. (2013)

participants to price signals with the same temporal and spatial resolution will become unavoidable (at least with respect to the transmission network, although the extension of such price signals to the distribution level would increase efficiency too, as discussed in **Chapter 4**).

7.2.1.3 Auction design

Fitting bidding formats to new resources

While electricity is often defined as a commodity (in the sense that one MWh of electricity is indistinguishable from another), experience has shown that electricity markets are more complex than markets for any other commodity. In addition to the constraints that are imposed by the network, electricity producers face various operational constraints (e.g., start-up, minimum power output, ramping constraints) that have a significant impact on the cost and thus the operation of generators. New resources (including demand-side resources) will introduce their own operational constraints, which will also need to be addressed — further complicating matters. These constraints are usually most relevant in short-term markets (i.e., day-ahead, intraday, and real-time markets) and therefore must be considered in the design of electricity auctions.

The goal of a market exchange is to facilitate trading to help participants manage their risks and efficiently match available supply with demand. In this respect, a key element for electricity markets is the design of bidding formats that allow market agents to include information about their costs and operating constraints. US and EU markets follow two different approaches to solve this problem. In the United States, costs (e.g., variable and start up) and operation constraints are represented in bidding formats in a straightforward way, using essentially the same unit commitment and dispatch optimization models that were used prior to restructuring. ISOs require generators to submit multi-part offers that represent detailed operational (and opportunity) costs (e.g., the California ISO uses warm, intermediate, and cold start-up offers), and also the technical constraints that apply to different generating units (e.g., offers in the Midcontinent ISO include generators' minimum run times and down times).

The European approach is based on the concept of a simple auction, where market agents indicate their willingness to buy or sell electricity using price–quantity pairs. The EU market, nonetheless, has quickly evolved toward increasingly complex bidding formats, which do not specifically capture the characteristics of any particular type of unit but are instead formulated in a general way with the goal of providing a level playing field to all market participants. The key information provided in the bid is a price to buy or sell electricity, but the bid can also include implicit information about the conditions under which an agent is willing to buy or sell (e.g., agents can present block bids and other complex conditions). These rather limited bidding formats can be handled reasonably well by large generators that can use their ability to manage large portfolios to compensate for a lack of complexity, while at the same time benefiting from disclosing the minimum amount of information about their operating cost structures. However, limiting the amount of information contained in bids particularly hinders the participation of small producers, makes it more difficult to monitor market power, and does not improve the efficiency of the resulting dispatch.

A hypothetical example is useful for comparing the US and EU approaches. This example involves a thermal power plant that is bidding in a day-ahead auction. Designing its bid is straightforward in an ISO market. The generator must submit a multi-part bid containing relevant information on its operating constraints and opportunity costs, and the auction will reveal the most efficient use of all contributing resources. Bidding in a European market, the same unit could submit a block order that made its electricity production conditional on being committed for a given number of hours and at a given average price (incorporating start-up cost). It could also use a profile block order to express that its order is only valid if a minimum ratio of the bid is executed (representing its minimum power output constraint). Incorporating operating costs and constraints in this bidding format requires making assumptions about what the most likely operating regime will be on the following day, based on prior knowledge and the ability to predict future market results. In the European Union, this predictability has decreased due to uncertainty introduced by the growing presence of intermittent energy sources. Moreover, the number of block bids submitted to the market has increased as agents create complex combinations of different types of bids to better represent their constraints, thereby increasing the computational complexity of the market clearing problem. An alternative that is currently being considered is a so-called thermal order, which is identical to the multi-part bid that is used in ISO markets in the United States. Implementing this approach could address the problem of computational complexity, since a single thermal order can replace multiple block bids. The thermal order approach has implications for clearing and pricing rules, which are discussed in the next section.

New bidding formats will be necessary to accommodate new market players. In Europe, a turn toward multi-part bids (similar to those in the United States) could better manage the increasing complexity of the current bidding system.

In the United States, Europe, and worldwide, electricity auctions have been designed for traditional resources, and current bidding formats are not ideal for some new technologies such as small-scale storage and demand response. One approach to deal with this problem in the European Union is to continue to increase the complexity of existing bidding formats, but there are questions about the scalability of this solution. A preferable alternative would be to adopt the ISO approach and create customized bidding formats for each type of resource. In ISO markets, where multi-part bidding formats are already in place, it would be necessary to design new bidding formats (or significantly enhance existing ones) to accommodate new resources.

Market clearing and pricing

Differences in bidding formats between the US and EU markets can be attributed to a fundamental difference in the models that are used for market clearing and pricing in the two regions. The US model for market clearing and pricing is an optimal unit commitment and dispatch model (or market welfare maximizing model). It uses locational marginal prices (LMPs) derived from the marginal cost of the system and side payments to clear the market. In contrast, the goal of the market-clearing algorithm in Europe is to find a uniform market-clearing price (the same price for all resources in the same location[14]) at which, for each time interval, the quantity of accepted generation offers equals the quantity of accepted consumption bids.

If only simple bids (price–quantity pairs or variable cost offers) were used,[15] both market clearing approaches would produce the same results. In this simplified setting, the US-style market welfare maximizing dispatch can also be obtained through an EU-style uniform price auction. In this case, uniform prices not only support the welfare maximizing solution, they also incentivize generators to bid their true costs into the auction and send an efficient signal that supports optimal long-term investment decisions.

When electricity auctions include complex bidding formats (as is necessary, for the reasons we have discussed),[16] the US and EU market clearing models begin to diverge. In this case, it is not possible in general to support the welfare maximizing solution with a uniform price due to the fact that some generators that are needed for optimal dispatch may not recover all of their costs at the marginal price. In the United States, for example, a generator could be setting the market price and still not recover its start-up cost, because the marginal price is only affected by the variable cost component of its offer. In this instance, the generator receives a side payment,[17] called uplift or "make-whole" payment in this context. Complex bidding formats in Europe have quite different consequences: The auction results deviate from the welfare maximizing solution

14 Uniform pricing is compatible with nodal pricing; it simply implies that side payments cannot be used.

15 Assuming that resources had convex cost structures, otherwise simple bids cannot adequately represent operating costs.

16 More precisely, when the market-clearing problem is non-convex.

17 We continue to use the term "side payment" throughout this section — while avoiding the use of "uplift" — because, in most jurisdictions, the term "uplift" is broader and also includes other system charges such as different operating reserve charges.

Box 7.5: Bidding a Storage Resource in Electricity Markets

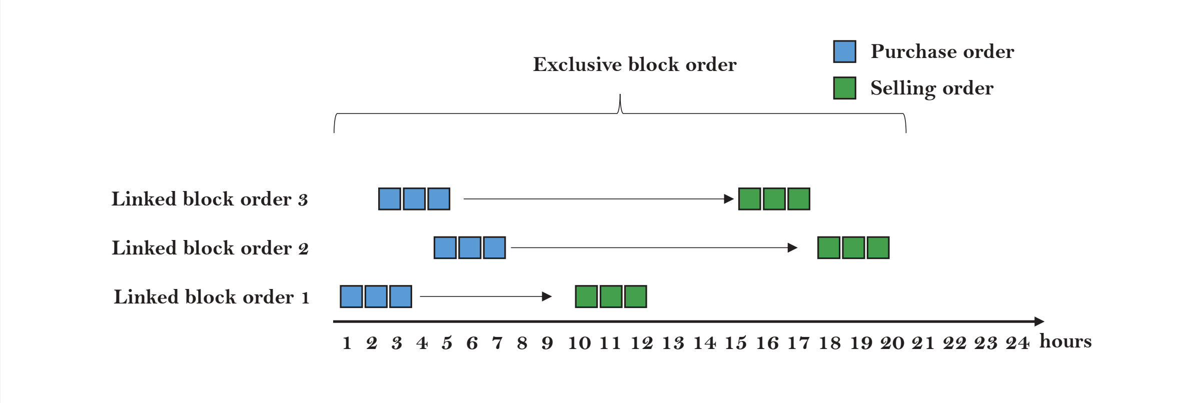
(Usera et al. 2016)

Significant limitations to optimally bidding an energy storage system (ESS) exist both in the US market (with some exceptions) and in the EU market. In many cases, bidding formats require the participant to anticipate specifically in which periods the ESS should be producing electricity and in which periods it should be consuming electricity, and to place separate bids as a generator and as a consumer at different times of the day according to these predictions.

However, it is possible to devise more suitable bidding formats. The California ISO already provides a bidding format designed for ESS participants (also called non-generating resources or NGRs), which can account for capacity limits, ramp rate constraints, maximum and minimum energy constraints, and state-of-charge constraints subject to charge/discharge efficiency rates. This is an extension of the multi-part bid approach to new resources, and its main challenge is the added computational complexity of market models.

In Europe, the best bidding format available would be a combination of linked and exclusive block orders (which are not available in all power exchanges). Using linked block orders, the ESS participant can connect its production and consumption bids to ensure, for example, that the storage system is not called to discharge before it has been charged. This still requires anticipating optimal periods for consumption and production, but by using exclusive block orders the agent can submit multiple, linked block-order pairs and have only the most profitable pair accepted (**Figure 7.4**). This combination of bidding formats still does not ensure the most efficient use of the resource but it is a considerable improvement over other, simpler bidding formats.

Figure 7.4: Linked and Exclusive Block Order Usage for a Storage Asset



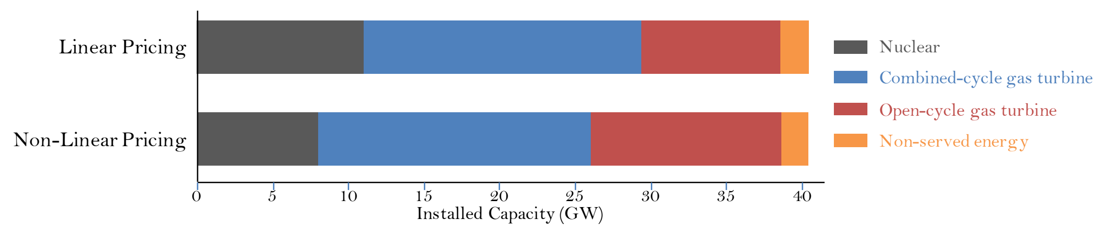
Source: Usera et al. (2016)

Box 7.6: The Impact of Pricing Rules on Investment Incentives

(Herrero et al. 2015)

This example makes use of a long-term capacity expansion model wherein generation investment decisions are based on market remuneration. For a realistic, thermally-based electricity system with a high penetration of renewable energy sources, the model determines the resulting generation mix in the case of traditional marginal cost pricing (i.e., a non-linear or discriminatory pricing scheme) and in the case of a simplified version of so-called convex-hull pricing (i.e., a linear or non-discriminatory pricing rule). **Figure 7.5** shows the generation mixes that result from applying the two different pricing rules.

Figure 7.5: Impact of Pricing Rules on Investment Decisions in a Simulated System



Source: Herrero et al. 2015

Three important conclusions can be drawn from this simulation. First, the way in which non-convexities are reflected in the uniform price can have a significant impact on the prices that drive market agents' investment decisions. Second, the linear pricing rule appears to promote a more efficient energy mix (see details in the referenced paper). And third, contrary to what a superficial analysis might suggest, a linear pricing rule does not necessarily produce higher energy prices than a non-linear pricing rule; in fact, a linear pricing rule can lower the average energy price since it attracts generation technologies with lower variable costs. However, these results also suggest that the gains resulting from different investment decisions under different pricing rules are small and easily swamped by other considerations.

because some bids are rejected as needed to ensure that all accepted bids are profitable at the uniform price (in other words, an offer will never be accepted if doing so requires a side payment).

The advantages of the US ISO approach are clear: The market clearing mechanism is such that it makes the most efficient use of all resources. However, the use of side payments distorts price signals and could lead to inefficient investment and operating decisions. This problem has been acknowledged for a long time, and ISOs have implemented slight variations to the traditional marginal cost pricing to minimize its impact (FERC 2014). **Box 7.6** summarizes a modeling exercise that assessed the extent to which long-term investment incentives can be affected by differences in pricing rules.

The demand side of electricity markets should not bear the full cost of side payments. EU and US markets should improve the balance between optimal dispatching and uniform pricing.

Innovations in US pricing rules will improve price signals and reduce the need for side payments, but those side payments that remain will be increasingly challenging to allocate (O'Neill et al. 2016). Currently in the United States, side payments are charged to demand, which is the least distortionary approach given that demand is relatively price inelastic. This choice also comes from the perception of side payments as generation costs, since they originate from complex generation offers. However, as soon as demand resources are allowed to use complex bidding formats (e.g., to reflect some of the usual constraints that could also apply to demand resources, such as a minimum consumption level that would prevent a demand bid from being partially accepted, or a minimum run time), some demand bids would equally require side payments. Demand resource providers can also experience revenue insufficiency if their bids are accepted in the market at a certain price and they are later charged additional side payments.

Therefore, allocation rules must be updated so that side payments are not borne exclusively by either the demand or generation part of the market, and the computation of these allocation decisions must be incorporated in the market-clearing problem.

EU markets, on the other hand, avoid side-payment problems at the expense of short-term dispatch efficiency, a trade-off that should be carefully reconsidered. An additional problem in EU markets is computational complexity. The EU market-clearing model, which includes uniform pricing and revenue sufficiency constraints, is intrinsically more difficult to solve than an optimal unit commitment and dispatch model. These two problems are aggravated by the increasing complexity of bidding formats and the entry of DERs. Thus it is necessary to find new solutions that bring EU markets closer to optimal dispatching. Some such alternatives are at an early stage of discussion (PCR 2016), but European legislation is ambiguous on the possibility of introducing discriminatory side payments (Madani et al. 2016).

7.2.2 Reserves and Balancing Markets

To maintain reliable operation, power systems must have the ability to respond to sudden changes in generation or consumption in order to constantly balance both. Once the day-ahead market is cleared, reliability is achieved through the use of different types of reserves that can be competitively procured by the system operator. This section reviews the changes that should be taken into consideration when redesigning the reserve products that are traded, as well as the market mechanisms for acquiring these resources.

7.2.2.1 Reserves markets

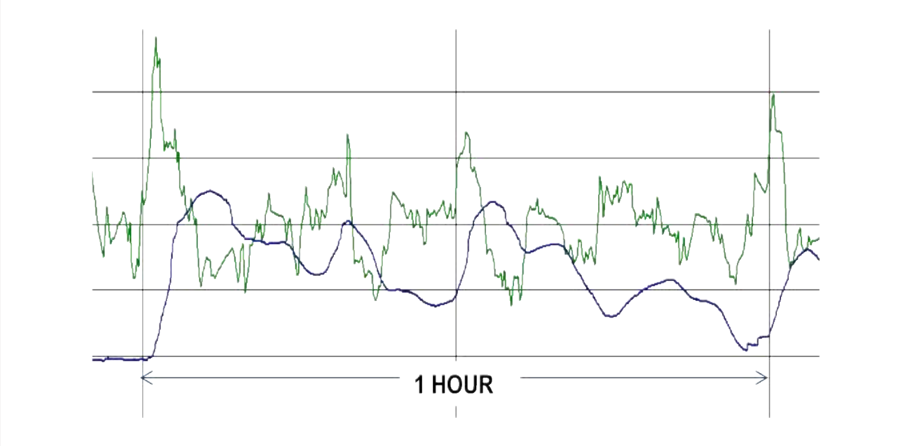
Redefining reserve products

The reserve products currently purchased by system operators are tailored to the characteristics of conventional generating technologies. The transformation of the power system changes reserves requirements, not

Box 7.7: New Regulation Requirements in PJM

PJM (the Pennsylvania-New Jersey-Maryland Interconnection) offers two secondary reserve sub-products: Reg A is the traditional regulation product, while Reg D is a newer product that is intended for faster-responding resources. Both products are jointly procured based on their relative contribution to a total regulation reserve requirement. Joint procurement is important to maintain competition and to compare different resources.

Figure 7.6: Regulation Signals for Two Secondary Reserve Products in the PJM System



Note: Regulation signals are shown for Reg A (blue) and Reg D (green). Source: PJM (2013)

The figure shows the main difference between the two products: The Reg A signal includes a smoothing (low-pass) filter to better fit ramping requirements to the capabilities of traditional resources, while Reg D uses a faster regulation signal (responding to higher frequency perturbations). The Reg D product was designed to take advantage of small storage (battery) resources, which have virtually no ramping constraints but must contend with a different constraint: namely limited energy availability. To overcome this barrier, the control system for Reg D is implemented in such a way that the net energy required is zero over a short period of time (95 percent of the time the controller converges in less than 15 minutes). This ensures that storage resources will not be required to sustain a response for a long time and will be able to maintain a relatively stable state of charge. Importantly, Reg D is not exclusively available for batteries — many hydro resources are also qualified to provide this service.

only because some new technologies are more variable than conventional technologies, but also because these new resources may offer a new source of flexibility. For example, distributed solar PV inverters could be used to support primary frequency control, but they will do so only to the extent that this service is appropriately rewarded. To fully exploit this potential flexibility, it is important to remove inefficient barriers in current product definitions and to create innovative reserve products that value different capabilities (see example in **Box 7.7**).

However, there is a delicate balance between defining reserve products that are better suited to a variety of new technologies, and creating new products for each individual technology. The objective of new reserve products should always be to provide value to the power system; new products should not be exclusively tailored to the capabilities of a particular technology. If new products are adapted to facilitate the use of a particular technology, as has been done in the past for conventional technologies, these adaptations should also contribute value to the electricity system. Product adaptations to support a particular technology (e.g., for policy purposes), on the other hand, are not advised, as they would inhibit effective competition.

Innovative reserve products should be implemented to fully exploit the flexibility of all resources. Inefficient barriers to using DERs for reserve provision should be removed.

Aggregating resources can help maintain the balance between developing refined reserve products and developing too many reserve products. An aggregator can combine the responses of different devices to meet the requirements of the system operator, thereby exploiting new sources of flexibility without overly complicating reserve product definitions.

Some established simplifications in reserve product definitions have become barriers to DER participation and should be reconsidered. For instance, most power systems have combined products for upward and downward reserves (especially for secondary/frequency restoration reserves) that could easily be replaced by two separate products. Separately procuring upward and downward reserves not only results in a more efficient use of resources (by not over-procuring one type of reserve only because the other is needed), it also allows for the participation of resources that are best suited to provide only one of these services. For example, wind or solar generation can be readily curtailed to provide downward reserves, while various forms of controllable demand can easily be curtailed to provide upward reserves.

The requirement for minimum bid sizes is another barrier that can be removed. While the cost of metering and control equipment will continue to provide economies of scale for large resources, markets should not impose artificial barriers to entry for small resources. If minimum bid sizes are unnecessary for the operation of the market, they should not be enforced. At the same time, DER aggregation should be facilitated to help DERs realize economies of scale, incur lower transaction costs, mitigate risk, and contribute to a more efficient system.

Efficient pricing of reserves

Implementing a reserve market is necessary but not sufficient to efficiently price reserves. The definition of reserve requirements and reserve settlement mechanisms significantly affects the remuneration of reserve services and is an important determinant of an efficient pricing system, for reasons we discuss in this section.

In some cases, reserve products are not procured competitively. Typically, providing primary frequency control is obligatory (for generating technologies that are capable of providing it) and in the vast majority of cases, generators are not remunerated for doing so effectively. Indeed, it might be unnecessary to implement a market for this service if it is assumed that the inertia in the electric power system is sufficiently abundant. Such a market might even have a detrimental effect if its relative lack of liquidity creates the potential for abuses of market power. But as the generation mix changes and market structures evolve, an adequate remuneration scheme for primary frequency control might be necessary.[18] For instance, if DERs are deployed without any incentive to provide primary regulation and/or inertial response, resources capable of providing primary regulation might become scarce (in that case, a proper pricing mechanism for this service would reflect the value that storage could have for the system).

Except in those systems with significant hydro-reservoir capacity — where again, due to the abundance of the product, the price of reserves is assumed to be zero — secondary and tertiary reserves are competitively procured, although a variety of market designs exist. In some systems, pay-as-bid rules have been implemented with the ostensible objective of mitigating market power and reducing price volatility. However, pay-as-bid pricing may not be the best or unique solution when a market poses market power concerns; other solutions include incentivizing competition or resorting to a regulated mechanism. If possible, a market-based marginal pricing approach is always preferable, as it most accurately reflects the value of flexibility and encourages long-term investments that are commensurate with system needs.

18 Since December 1, 2007, German transmission system operators (TSOs) have been meeting their primary control reserve (PCR) requirements through shared calls for tenders. On 2011, the minimum lot size was set at plus or minus 1 MW (as compared to the previous minimum of 5 MW). See: www.regelleistung.net.

Another feature that differentiates power systems is the extent to which reserve markets are coordinated with energy markets. The values (prices) of reserves and energy are mutually dependent. The fact that power plants can simultaneously offer more than one product links the value of reserves with the opportunity cost of providing energy, and vice versa. In this regard, a market design that favors the discovery of this interdependent value produces more efficient outcomes. A co-optimized design wherein the reserves market is jointly cleared with the energy market, as it is often the case in ISO markets, is theoretically the first-best approach. When this is not possible, whether for reasons of computational complexity or governance structure (in Europe, energy markets are organized by power exchanges and reserves markets by transmission system operators), it is important to improve the coordination of energy and reserves markets. This can be done by connecting their timelines, so that price variations in one market are reflected in the other.

Remuneration of reserves must be integrated with or connected to energy markets to reflect the value of reserves and account for the performance of different resources in providing reserves. This integration is particularly important during scarcity events.

The connection between energy and reserves is an essential component of designing markets that adequately reflect scarcity and drive efficient investment decisions. Properly defined reserve requirements can help reveal the value of system flexibility, an increasingly important consideration in electricity systems. In most markets, the quantity of reserves required is fixed, regardless of the cost of those reserves. However, these reserve requirements should reflect the changing value of reserves to the system, and should surely be periodically reexamined (see further discussion in **Box 7.8**). Assessing this value is difficult, and any attempt to do so entails significant intervention into markets. Therefore, the long-term benefits of flexible reserve requirements should be weighed against the costs of such intervention.

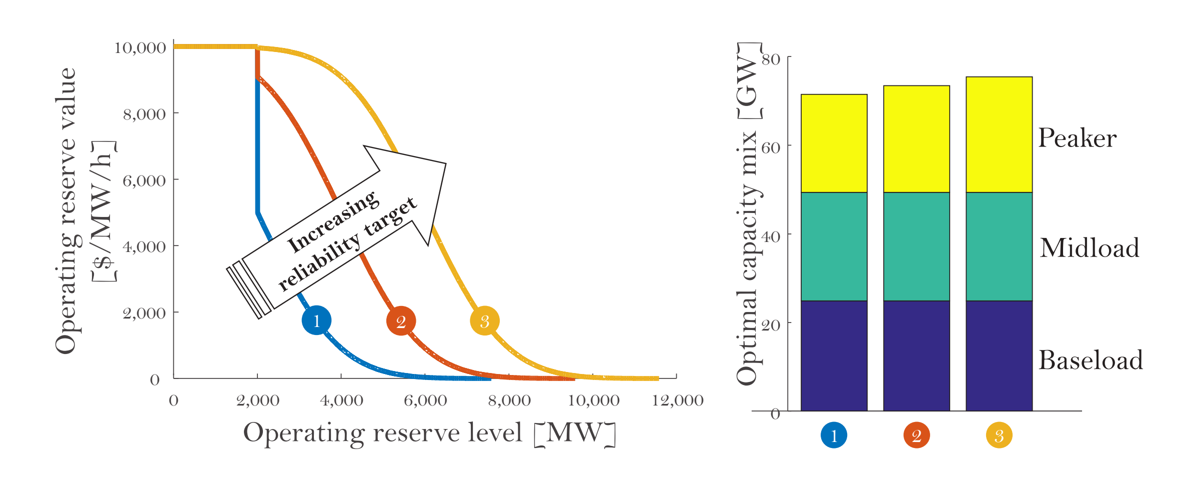
Operating reserves require the provision of both system capacity (MW) and the use of balancing energy (MWh). In many cases, capacity and balancing energy are jointly procured for various services. In such cases, only resources that are selling capacity can provide real-time balancing energy (typically via a pro-rata allocation). This is detrimental for technologies that cannot guarantee availability when reserves are procured (often, a day or more ahead) but that could be efficiently deployed in real time. As an example, a fleet of electric vehicles may not be able to commit much capacity in advance (its available capacity is more uncertain than that of a conventional generator) but it can provide balancing energy in real time. In cases like this, joint procurement of capacity and energy can be a barrier for the participation of some DERs. This barrier could be overcome by organizing large DER aggregations, which can provide more firm reserve commitments. This approach, however, still would not guarantee the most efficient use of energy from remaining resources. For this reason, separating the procurement of reserve capacity and energy is recommended. If the two products are separated, capacity can be procured ahead of time as a kind of "insurance" to provide the certainty and reliability needed for system operation — thereby still providing incentives for DER aggregation — while the decision about which resources will provide energy in real time can be postponed for maximum efficiency.

Box 7.8: The Impact of the Operating Reserves Demand Curve on the Configuration of the Future Capacity Mix

Stoft (2003) highlighted the impact of increased short-term reserves contracting on long-term system reliability. Hogan (2013) proposed implementing an operating reserves demand curve (ORDC). The ORDC would be integrated into economic dispatch and would internalize the value of operating reserves into standard electricity markets. Defining the ORDC is complex since it requires estimating the value of lost load and the probability of scarcity events for a given reserve level, and defining the connection between different reserve products. While determinations about the probability of scarcity events should be informed by rigorous technical simulations, the final decision ultimately involves policy judgment. Moreover, these determinations can potentially be adjusted to induce investments, with varying implications for system reliability.

Figure 7.7 illustrates the impacts of implementing an ORDC through a simple case study. Using a realistic electricity system, the study shows that as the reliability target increases (thereby shifting the ORDC outward), the long-run installed capacity of peaking units increases. The implementation of the ORDC also serves as a risk hedge for these scarcity-oriented units, as it increases the frequency with which these units are remunerated in the short term.

Figure 7.7: Illustrative ORDCs and Optimal Capacity Mix in a Sample Case



A simplified version of the ORDC was implemented in ERCOT, the system that serves most of the state of Texas, in 2014. Additionally, most ISOs implement some form of scarcity pricing, and discussions are currently underway to redefine flexible reserve requirements.[19]

19 We recommend the implementation of an ORDC, as it improves the market signal for agents in the short and long terms. However, as we discuss in **Section 7.3.1.1**, implementing the ORDC by itself is unlikely to solve reliability problems (when detected as such), as the key reason behind the alleged need to implement capacity mechanisms in today's markets happens to be the large regulatory and planning uncertainty that investors face.

7.2.2.2 Imbalance settlements

This section analyzes three market design elements that are related to imbalance settlements: (1) the possibility to exempt some resources from balancing responsibilities; (2) the aggregation of resources for imbalance settlement purposes; and (3) the imbalance settlement pricing scheme. The design of each of these elements has implications for the efficiency of the others.

Balancing responsibility and exemptions

Balancing responsibility refers to the extent to which market agents are financially responsible for their deviations (from their physically declared programs) in real-time operation. Balancing responsibility is enforced so that resources have incentives to help maintain the physical balance of the electricity system and, secondarily, to maintain the financial balance of the electricity system.

In an effort to support renewable energy technologies, many countries permit renewable producers partial or full exemption from balancing responsibility (in other words, the costs associated with renewable energy imbalances are socialized). In other cases, renewable energy resources are "balance responsible" but are subject to different, more favorable charges than other resources. The impact of these exceptions is relatively small with low penetrations of renewable resources. However, as these technologies become widespread, differential treatment leads to market inefficiency and should be eliminated so that all resources face similar market risks. In systems where renewable energy resources are exposed to these signals, the ability to forecast their future production has significantly improved (see, for example, Herrero et al. [2016]). The same reasoning applies to DERs, which, at the low-voltage level, usually do not have the same balancing responsibilities as centralized resources.

Renewable energy technologies and DERs should be exposed to the same balancing requirements as other technologies.

Maintaining special treatment for particular technologies has negative consequences for overall system costs. It dilutes price signals that could incentivize investment in technologies with load-following capabilities. Moreover, it inhibits the ability of renewable energy resources and DERs to update their production/consumption schedules in close-to-real time (see the previous discussion on intraday markets).

Imbalance pricing

In addition to determining balance responsibilities, the prices or charges that are applied to imbalances must also be determined. Using the European terminology, there are two major approaches to imbalance pricing:

- *Single imbalance pricing.* A single reserve energy price at a given location (zone/node) and time is used to settle deviations from the market program. Negative deviations pay for positive deviations and the result is a zero-sum mechanism for the system operator.
- *Dual imbalance pricing.* Different prices are used to settle positive and negative imbalances. In this case, positive deviations are not paid the full value of the injected energy, and negative deviations pay for more than their cost. This results in a non-zero-sum mechanism.

In theory, single imbalance pricing provides optimal incentives, since it means that a balancing responsible party (1) only profits if it deviates from its market program in favor of system requirements (whether the deviation is up or down) and (2) pays exactly the marginal cost of its deviation when it deviates against system requirements. In practice, however, system operators may prefer to avoid an incentive-based system and opt for a more command-and-control type of mechanism.[20] In this case, dual imbalance pricing may be used to penalize all deviations from the market program to create incentives for agents to each keep a balanced program. In other words, dual imbalance pricing promotes more conservative behavior and decreases the risk of suffering high imbalances, but does so at the cost of distorted market signals (thus giving rise to associated inefficiencies).

20 As discussed above, theoretically efficient single imbalance price signals may be distorted by other design elements, such as a lack of sufficient price granularity.

Allowing aggregation for imbalance settlement coupled with dual imbalance pricing imposes an unnecessary barrier to entry for smaller participants (particularly DERs).

Imbalances are measured and settled at different levels of aggregation depending on the system. In some power systems, imbalances are measured on a plant-by-plant basis, while in others market agents are allowed to aggregate various resources for imbalance settlement purposes. In Europe, for example, these aggregations are called "balance responsible parties" (BRPs).

Note that, depending on the imbalance pricing scheme, the level of aggregation has additional consequences. If the aggregation of various resources is allowed under dual imbalance pricing, large portfolios of power plants have a significant advantage over smaller BRPs due to the netting of internal deviations, which gives the plants a competitive advantage without necessarily contributing to system value. Additionally, the presence of large BRPs reduces balancing market liquidity (even, in some cases, opening the door to potential balancing capacity withholding). This can act as a barrier to entry for smaller DERs. As previously discussed, dual imbalance pricing should be avoided, but if it is not, all BRPs should have similar sizes. In the traditional context, this calls for an imbalance settlement defined on a unit-by-unit basis so that all resources can compete on a level playing field. In the more complex context of DERs, a maximum size (comparable to that of a conventional power plant) for DER aggregators should be set for imbalance settlement purposes. Again, these complexities could be avoided with a single imbalance pricing scheme.

Liquidity in reserve and balancing markets

Finally, in very short-term markets, liquidity is typically not sufficient to efficiently cope with real-time adjustment needs. This lack of liquidity can lead to volatile balancing costs that could unduly penalize the business activities of new (and typically small) agents, particularly small aggregations of DERs.

When this is the case, regulators should consider the convenience of introducing additional mechanisms to enhance liquidity and competition. There are two major alternatives in this respect (Batlle et al. 2007): (1) allowing long-term contracting of (tertiary) reserves or (2) introducing market makers in the balancing market.[21]

7.3 Balancing Long-term Policy-Oriented Planning Objectives, Uncertainty, and Market Shortsightedness

This section discusses the criteria that should be taken into consideration when designing the regulatory mechanisms that are used to promote investments in generation capacity, starting with an assessment of capacity mechanisms and then giving special attention to their technology-specific form — that is, support mechanisms for clean energy technologies.

7.3.1 Dealing with long-term uncertainty: Capacity mechanisms

We next analyze how the regulatory design of capacity mechanisms should evolve to improve their performance in a scenario that anticipates significant penetration of large-scale renewables and distributed energy resources. Thus, we discuss only those aspects of capacity mechanism design that explicitly have to do with the integration of these technologies.[22]

21 Market makers must maintain continuous two-sided orders (bid and offer) within a predefined spread (i.e., the difference between the buy and sell price). Market makers ensure that the market bid–offer spread stays within these limits, by permanently providing a buyer for every sell order and a seller for every buy order in the predefined range. In this way market makers function as a sort of permanent liquidity provider.

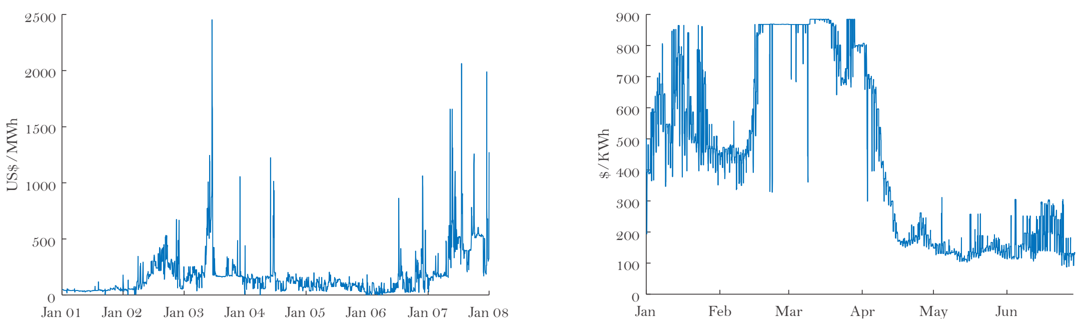
22 A more general review of the different design elements of capacity mechanisms can be found for example in FERC (2013), EFET (2013), or Batlle et al. (2015).

Box 7.9: Wind Contribution to Security of Supply in Systems with Large Storage Capabilities

In most thermal systems, scarcity conditions, though infrequent, can arise from a deficit of installed capacity that translates into scant reserves during a few dispersed hours of very cold or very hot weather. This is the situation that prevails, for example, in the eastern United States, in areas such as New York or New England, where capacity mechanisms have been instituted. Resulting market prices reflect this potential for scarcity conditions, as shown in the series of real-time prices in the New York ISO during the polar vertex event of 2013–2014. The left graph in **Figure 7.8** shows prices during this period of extreme cold weather that reflect the scarcity of installed capacity, not of energy. In such cases, intermittency results in a low capacity credit for wind or solar photovoltaic (see also footnote 23 later in this section).

In hydro-dominated systems with large storage capacity, as is the case in Colombia (with more than 60 percent hydro), rationing takes place due to lack of available energy, not capacity. This is also clearly reflected in day-ahead market prices, as shown in the right graph of Figure 7.8. Day-ahead market prices are very flat, since the regulation capability that comes with hydro resources gives much flexibility to the unit commitment process. Thus, in this context, the intermittency of wind and solar is not an issue.

Figure 7.8: Spot Market Prices for Electricity in the NYISO and Colombian Systems

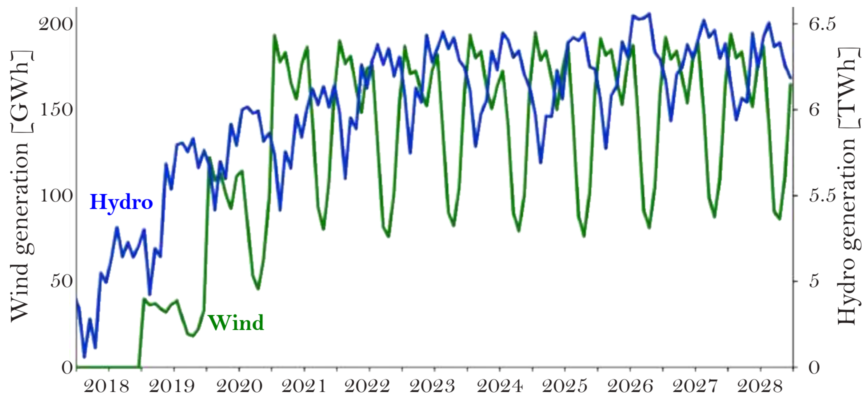


NYISO RTM prices, first week of January 2014

Colombian exchange prices, first semester 2016

In some systems it can also be the case that the pattern of wind production is largely complementary with that of key generation resources. Again, this is the case in Colombia, as illustrated in **Figure 7.9**, which was provided by Colombia's Mining and Energy Planning Unit (UPME 2015) in its detailed planning reports.

Figure 7.9: Future Complementarity of Wind and Hydroelectric Resources in the Colombian System



7.3.1.1 "Disruptive" technologies and generation reliability

There is a widespread misconception that the penetration of renewable energy technologies in power systems is one of the key reasons to implement a capacity mechanism; see, for example, Artelys (2015). From that perspective, the deployment of renewables, which has been driven, so far, by regulatory support, is behind the need for capacity mechanisms because renewables depress prices in the short-term electricity market. In addition, "the intermittent character of renewables creates uncertainty regarding the frequency of price spikes that help conventional technologies to recoup their investment costs" (European Commission 2016). Actually, it is uncertainty about the long-term penetration of renewables (due to policy uncertainty and also to learning curves for renewable technologies) that hampers future investments — not increased price volatility or the transitory reduction of average prices due to the quick deployment of these technologies (see, for example, Henriot and Glachant [2015]).

Another classical misconception is that some DER technologies do not represent a valuable resource for security of supply. In the first place, most DERs can make some contribution to system reliability. In addition, in some power systems that are reliant on hydroelectricity and that have some storage capabilities, renewable sources may also offer a complementary availability with respect to the main energy resource; see, for example, Gouveia et al. (2014) and the description of the Colombian case in **Box 7.9**. The time needed to deploy renewables may also be significantly shorter than for other technologies — this is the case in Brazil, where the main capacity alternative is large-scale hydro, which involves long construction times (Barroso and Batlle 2011).

In energy-constrained systems, as in Colombia, Brazil, Norway, or New Zealand, it is lack of energy, not capacity, that causes rationing. In these contexts, if renewables are able to deliver — on average — their expected contribution in the medium term (e.g., within the week), they save water in the reservoirs, regardless of their daily or hourly production schedule. Their participation in a capacity mechanism is therefore largely beneficial.

In systems with significant medium-term storage capabilities, the intermittency of renewable technologies does not impede the renewables from contributing substantially to the reliability of the system.

In capacity-constrained systems, as in Great Britain, Germany, Italy, or (in the United States) PJM or New England, scarcity conditions arise because there is not enough installed capacity available to meet load at every given moment during extreme weather events. In these systems, the intermittency of renewable production is an issue.[23]

7.3.1.2 Interaction of capacity mechanisms with other support mechanisms

The debate about whether or not to isolate developing technologies from market signals, and to what extent, has been intense. However, the discussion has been limited to the interaction of these new technologies with short-term markets. The discussion now needs to be extended to include increasingly widespread capacity mechanisms. First, capacity markets should allow "non-conventional" energy resources of all types (renewable generation, demand response, storage, etc., whether distributed or not) to participate in any kind of capacity mechanism. At the same time, to the best of their capabilities, these resources should be subject to the same conditions as any other technology that participates in the capacity mechanism. As discussed above, these conditions should be defined to ensure that resources contribute to system reliability.[24]

Nevertheless, regulators in many countries have so far decided that renewable and demand response resources, which already benefit from other kinds of incentives (such as grants, feed-in tariffs, targeted programs, etc.), are not eligible for remuneration through the capacity mechanism. This has been the case, for example, in Great Britain's Capacity Market (DECC 2014). The regulatory practice of excluding DERs from these markets misses the key point we are making here: that these technologies need to be exposed to all market signals in order to maximize the benefits they provide. Thus, renewables and demand response should be integrated in the wholesale market as much as possible, with all associated consequences.[25]

"Non-conventional" technologies (renewables, storage, demand response, etc.), whether distributed or not, should be integrated in capacity mechanisms on an equal footing with other technologies. If coherently designed, this integration is fully compatible with other support mechanisms for these technologies.

Capacity mechanisms not only provide an investment incentive, they are also intended to provide market incentives for making committed resources available during scarcity conditions, with the objective of optimizing the efficiency of the system. If the design of the capacity mechanism is robust and the reliability product is technology-neutral, all technologies that participate in system reliability are acknowledged and exposed to an economic signal that prompts them to be available when the system most needs them. Non-conventional technologies should not be favored, either by being exempted from penalties or by receiving over-generous recognition for their firm energy/capacity contributions.

Mechanisms to support renewable technologies are not incompatible with this approach. As discussed in **Section 7.3.2.2**, the ideal renewable support scheme is one in which revenues obtained by renewable plants, from all markets in which they participate, including energy and capacity markets, are deducted from the subsidy they require to support their deployment.

7.3.1.3 Capacity mechanism design elements and distributed energy resources

Applying the same rules to any technology that bids in an auction does not necessarily guarantee that all technologies actually compete on a level playing field.

23 In 2017/2018, the PJM capacity market (known as "RPM" for Reliability Pricing Model) cleared 803 MW of wind power plants and 116 MW of solar resources in its base residual auction (PJM 2014). On average, wind and solar units were assigned a 13 percent and 38 percent capacity factor, respectively, during the qualification phase prior to the auction.

24 This is already taking place in Brazil, as analyzed in Mastropietro et al. (2014), where generation capacity mechanisms are a component of the income from renewables. Renewable technologies started being cleared in specific renewable auctions, but they ended up being selected also in conventional long-term auctions for new electricity supply. As discussed in the next subsection, however, this convergence is not yet complete and some rules still need to be harmonized.

25 As highlighted by Finon and Roques (2012), "a more comprehensive approach based on a strong public governance has been proposed to deal with both objectives of decarbonisation and system reliability through the same policy instrument which would be a market-wide capacity forward auctioning."

In practice, capacity mechanisms cannot be "technology neutral." Any single capacity or reliability product, however it is designed, will be better suited to some technologies than to others.

The numerous design elements that need to be defined when implementing a capacity mechanism heavily condition which technologies will be better adapted to the mechanism. These design elements include, among many others, physical guarantees, financial commitments, contract duration, lead-time, penalties for non-compliance, etc. Different resources call for different contract conditions. For instance, a seven-year contract might work for a gas turbine, but may not be suitable for a large hydro project.

A combination of reliability products

One well-known option for solving the problem that capacity mechanisms inevitably favor some technologies over others is to define different products in the mechanism.[26] Obviously this differentiation comes at a cost: The regulator needs to define a methodology to compare apples and oranges, since each technology actually offers different products that are hard to compare. (For example, what is better suited for a certain system? A 10-year contract for 1000 MW of gas turbines or a 20-year contract for 1000 MW of large hydro?) As a result, the methodology needed to clear the auction ends up being rather complex, and efficiency becomes a matter of debate. Although there is no perfect solution for this problem, some good practices should be considered.

With the advent of new technologies that can contribute to system reliability (e.g., biomass, geothermal, wind, solar, demand response, storage, etc.), existing challenges related to defining capacity products are exacerbated. As a result, multiple controversies arise. Regarding the duration of reliability contracts, for example, in Great Britain's capacity market, different contract lengths are available to new generators and demand response (Orme 2015). While new generators can opt for multi-year contracts (up to fifteen years for projects above a capital expenditure threshold of £255/kW), only one-year contracts are available for demand response (since it is deemed to be less capital intensive and can change more rapidly — for example, by moving to another location). From the perspective of investors, the longer the duration of the contract, the better. Although it makes sense to try to adapt contract duration to the different capital structures of different technologies, we suggest that regulators should seek to avoid extremely lengthy contracts as much as reasonable. For example, due to the effect of discounting, future income has less and less impact on decision-making at the time of the auction, depending on how far out this income occurs from the present. Therefore, contract durations of thirty years are seldom justified with the high discount rates used by the generation sector in South America, for instance.

Recently, the most controversial set of issues relates to the specific design of the firm capacity product (and particularly to the obligations embedded in the product). Many examples can be cited here. Walton (2016) points out that units selling the so-called Capacity Performance product in PJM are expected to be online year-round, meaning that solar generators or demand response programs focused on summer cooling loads would struggle to bid this product.[27]

In many cases, the obvious solution to these matters is aggregation, which can either be performed by the system operator or be simply left to market agents. In the first case, the regulator can define tailor-made products — for example, defining seasonal products (such as summer and winter) and then defining the most efficient way to clear a combined auction (some agents might bid for only one of the products while others might bid for both). Alternatively, the regulator can design a product that implies a maximum number of hours of response for each provider and then manage different providers in the best way possible when scarcity is detected.

26 When the first capacity mechanisms were implemented back in the 1990s, particularly in thermal systems (such as those in the eastern United States), this problem of technology bias was initially not perceived as such. Regulators determined that system expansion was going to be led by just one or at most two types of generation technologies with similar financial and cost structures (i.e., gas- and coal-fired plants). On the contrary, Latin American systems faced, from the start, different alternatives: namely hydro and thermal plants. As a result, most capacity mechanism designs included reliability products tailored to different generation technologies, such as ten- to thirty-year contract terms and three- to seven-year project lead times (Batlle et al. 2015).

27 Other mechanisms, such as the one implemented in the New York ISO, hold winter and summer auctions separately.

The market price as the efficient indicator of critical periods

In any case, our recommendation is to take the market price as the indicator of a critical period in which firm capacity is needed (in other words, use market price to identify scarcity or rationing conditions). The spot, balancing, or real-time market price should be used as the critical period indicator and scarcity conditions should be defined as characterizing the period of time during which the reference market price exceeds a predetermined strike price.

A proper and transparent definition of the critical period indicator that takes as a reference the price of a sufficiently liquid short-term market provides good incentives for enhancing flexibility.

As discussed in Batlle et al. (2015), this approach should be increasingly valid in a future scenario that features increased elasticity of demand. In fact, as long as the amount of completely inelastic demand in the market (i.e., the demand that bids at the price cap) declines, it will become more and more difficult to define, in the long term, the demand that "must" be served. Consequently, it will also become more and more difficult to identify near-rationing conditions using only a comparison between peak demand and generation available. Rather, any time that prices rise above the marginal cost of the most costly generators can be considered a period of capacity limitation, in which resources committed in capacity markets should be available.

This critical period indicator obviously assumes the presence of a liquid reference short-term market in the system. The selection of the reference market also affects the kind of scarcity conditions that are covered, and that the regulator wants to have covered, by the capacity mechanism. On the one hand, day-ahead markets are only capable of capturing emergency situations related to the combination of high loads (as on days of peak winter demand) and reduced supply (due to fuel constraints or a dry year that limits hydro production) — that is, pure reliability issues. On the other hand, intraday and balancing markets are also subject to price fluctuations caused by sudden events (such as the outage of a nuclear plant or, in systems with high renewable penetration, a decline in intermittent generation due to bad forecasting). These situations can provoke temporary generation scarcity even if total load is far from the peak and do so over a time horizon that extends beyond the period covered by ancillary services. In this case, the system faces firmness and flexibility issues. However, the selection of resources should also take into account the signal that the capacity mechanism is providing to the generation mix. While all units are more or less technically capable of producing if notified one day ahead, certain technologies (including baseload technologies, such as coal power plants) would not be able to take part in the balancing market because ramp constraints prevent them from responding on such short notice. Therefore, a capacity mechanism that uses the balancing market as the reference market provides agents with a signal that discourages the installation of new baseload units. This may not be the objective or it may be precisely the objective, if the aim is to enhance the installation of so-called flexible units. An alternative would be to define different products in the mechanism, but using a combination of reliability products creates problems that have been discussed above, since it unequivocally implies that the regulator determines how much capacity is to be installed of each sort of technology or group of technologies. Again, the best approach is system-dependent and will need to account for different trade-offs.

How to avoid mixing objectives

It is common in regulatory debates to find arguments against certain rules for capacity mechanisms that set barriers to the promotion of certain technologies. In fact, rules should be defined in such a way as to provide market agents with incentives to plan, and later operate, their investments to maximize system reliability — not to promote one technology or another. Thus, the capacity mechanism needs to require availability just when the system needs it, which is why the market price is the best indicator of scarcity.

Capacity mechanisms should be exclusively aimed at enhancing long-term system reliability. Promoting specific technologies, if desired for policy reasons, should be achieved through separate support mechanisms.

While regulators should avoid pursuing positive discrimination — that is, favoring certain technologies (other mechanisms can be used for this purpose, as we discuss in the next section) — regulators should, at the same time, avoid unjustified negative discrimination against particular technologies. Along these lines (and as noted previously), a design element that needs careful consideration is the methodology used to calculate the firm energy/capacity that resources are granted with the capacity payment or allowed to trade through the capacity mechanism. In order to coherently account for the potential complementarities of different technologies, the most appropriate way for the regulator to determine the value of firm energy/capacity (sometimes also called "capacity certificate" or "capacity credit") should be through an integrated simulation of future system operation. This would allow the regulator to calculate the actual contribution from renewable resources during expected scarcity periods. Traditionally, especially in hydrothermal systems that have implemented capacity mechanisms, the problem has been that these calculations fail to properly account for the actual contribution of renewables. For example, in the Colombian case discussed in **Box 7.9**, only the minimum instantaneous production expected from wind generating units is considered. This leads to an administratively determined capacity credit or "firm energy" value of 6 percent of installed capacity even though, due to complementarity between large hydro storage capabilities and wind, the firm energy value of wind resources in this system should be closer to their average energy production.

7.3.1.4 Capacity mechanism cost allocation in end-user tariffs as a demand response incentive

As argued throughout this study, a proper end-user tariff design provides the regulatory cornerstone for an efficient future evolution of the power system. This also holds true in the case of capacity mechanisms.

A well-designed capacity mechanism should guarantee consumers that their demand will be supplied in the future (or that an equivalent compensation will be paid according to the targeted level of supply security). This service has a cost that is usually included in the tariff as a specific item, although until now there has not been a specific methodology to properly allocate this cost.

This issue turns out to be crucial since, without an adequate cost allocation methodology, inefficient arbitrage opportunities quickly arise for end users.[28] A specific component should therefore be added to electricity tariffs, in such a way that end consumers pay for the costs of the capacity mechanism as a function of their consumption during those periods when the system is close to a scarcity condition.

The cost of capacity mechanisms should be properly allocated in end-users' tariffs, as a function of their consumption in periods of system stress, to enhance demand response and to avoid inefficient arbitrage opportunities.

Once the tariff regime is properly designed, if demand is to participate in the capacity mechanism, two different approaches can be followed. The first option is to allow demand to opt out from the capacity mechanism. Its security of supply will not be guaranteed (specific procedures should be put in place to guarantee that these consumers can be effectively disconnected or

28 This can be easily illustrated with a real-world example: A university in (west) Cambridge, Massachusetts, participates in the forward capacity market of ISO New England by providing demand response on critical days. As a result, the university responds to ISO instructions on abnormally hot summer days by sending an email to its employees telling them that they are allowed to stay home because the university will not be running its AC systems. The consequence of this reaction might be that the university's employees stay home and turn on their own air conditioning units. This could lead to larger net consumption. While the university is being sent the right price signals, its employees (via their ill-designed volumetric tariffs) clearly are not.

heavily penalized), but this demand will not be required to pay the corresponding tariff item. The second option involves forcing the whole system demand to be covered by the capacity mechanism, while at the same time allowing consumers to sell demand-response products in the capacity market. With this approach, demand response competes with other reliability providers in the supply curve of the capacity mechanism (while with the previous approach demand response is simply removed from the demand curve) and is subject to the same contract provisions, including penalties for under delivery. In this case, it is essential to properly define how to calculate the customer baseline (Chao 2011), in order to assess the performance of the demand resource during scarcity conditions.

7.3.2 Support schemes for clean energy technologies[29]

The interference of support mechanisms with market signals today

From early efforts to implement support mechanisms, the debate about whether or not to isolate renewables from market price risk has been intense. In most systems, it was assumed that the inefficiencies derived from reduced wholesale market exposure were outweighed by the benefits of reducing investment risk for these emerging technologies. As renewable energy technologies mature and reach higher levels of market penetration, however, the system impacts of insulating renewables from market signals are becoming more difficult to ignore, as they lead to both short- and long-term distortions. First, some production-based subsidies can lead to negative spot prices as renewable generators with very low variable costs (wind and solar) have an incentive to keep running even when prices are negative in order to access support payments. Also, reserve costs can increase unnecessarily if renewables are not exposed to balancing and regulation market signals. As discussed in **Section 7.2**, this is because the absence of market signals gives these generators no incentive to improve their production forecasts, which in turn results in the need to commit larger quantities of reserves. Finally, ill-designed subsidies for clean energy technologies can lead to long-term distortions in the composition of the generation fleet (see **Box 7.10**).

The market penetration and maturity of renewable technologies have grown to the point that these technologies can now be reasonably expected to manage market risks like other, more conventional generators.

Promoting the integration of clean energy technologies in markets, while at the same time trying to reduce risks for investors in these technologies to just the level that they can properly manage, calls for careful consideration of design elements in support mechanisms for renewable technologies, as we discuss next.

The key to a proper analysis of renewable technology support schemes: Focusing on design elements instead of labels

Because labels for traditional renewable support policies (e.g., feed-in tariffs, feed-in premiums, auctions, etc.) are rarely applied in a standard or uniform manner and are commonly subject to misinterpretation, this study focuses on the design elements of these policies instead of resorting to traditional classifications.[30] We begin by highlighting three high-level decisions in the design of any type of support mechanism:

- *Desired degree of technical and operational integration.* There are typically specific technical requirements for the grid connection of renewable technologies, such as effect on reactive control (IEA-RETD 2015). The most reasonable approach is to ensure that, to the extent technically and economically feasible, renewable generation complies with the same grid standards as conventional generation.

29 The discussion included in this section is further developed in Huntington, Rodilla, and Batlle (2016).

30 For example, in Europe the debate was around whether to use feed-in tariffs, tradable green certificates, or auction-based mechanisms (Butler and Neuhoff 2008). In the United States, the debate continues today, with recent discussions centered on power purchase agreements versus long-term fixed price certificates (Harris 2015). However, confusion about the actual meaning of these labels is becoming an increasingly recognized problem in the literature (Couture et al. 2015).

- *Total support to be provided.* There are different ways to decide the amount of total support to be provided (in other words, the amount spent on the support mechanism), ranging from purely price-based to purely quantity-based mechanisms. Quantity-based mechanisms can take advantage of competitive pressure and technological improvements. Thus, auctions are increasingly being designed with a requirement for a given quantity of renewable capacity as a way to reveal more accurate price information by leveraging competition between renewable technology investors. In some cases, auctions have resulted in prices that are close to or even below the prices of conventional technologies (Maurer and Barroso 2011).[31] However, this approach also has drawbacks. Besides the fact that small-scale projects could hardly compete on prices with large-scale ones yet, quantity-based mechanisms might involve higher entry costs for small investors (domestic customers, farmers, etc.), unless sufficiently mature aggregators are available. If the regulator wants to specifically support certain smaller-scale investors, the solution can be to apply different types of mechanisms to different sizes of projects or classes of technology (Couture et al. 2015).

- *Structure or format of remuneration.* This design decision involves the incentive that renewable generators expect to receive depending on the characteristics of the plant itself, operational decisions made by plant operators, or exogenous market factors. There are two main formats for remuneration, namely production-based and capacity-based.

The remainder of this section focuses on design elements associated with this latter issue — the format of remuneration — and in particular on the market compatibility of different design alternatives. We first discuss the design of production-based support mechanisms. **Section 7.3.2.2** analyzes capacity-based support mechanisms. Section **7.3.2.3** concludes with major highlights and recommendations.

In a nutshell, as we discuss below, we conclude that as long as support is needed, a well-designed and implemented capacity-based support mechanism is preferable if the goal is true market integration. A capacity-based mechanism can be refined to reduce investor risk by adjusting periodic compensation (per-installed-MW) *ex post* based on market conditions (e.g., annual compensation can be revised upwards to guarantee a minimum level of income each year when prices are low, or revised downward when prices are high). However, even if such compensation is annually averaged in an effort to disconnect the subsidy from short-term price signals, an *ex post* compensation based on the plant's actual market outcome will, in the end, result in a kind of "production-based" incentive and will therefore cause market distortions. A suitable trade-off for implementing per-installed-MW compensation would be to relate it to the market performance of a benchmark reference plant. The reference plant should exemplify the most efficiently dispatched plant of each kind, so that it creates an incentive for actual plants to orient their siting and planning decisions to enhance overall market efficiency. Market rents would then be calculated for this reference plant and the appropriate level of support would be determined by estimating the additional compensation needed for the reference plant to break even. Providing this incentive to all supported plants would give investors more certainty, while at the same time not distorting market signals.

7.3.2.1 Production-based mechanisms

In production-based remuneration schemes, payments are based on a generator's actual production (e.g., MWh) and, in some cases, also on market prices. These mechanisms have several key design features:

- Though there are many variations, the payment mechanism can be classified based on whether remuneration is (1) not tied to any market (in other words, renewable generators are paid on a per-kilowatt-hour basis regardless of the market price at the time), (2) tied to the energy market (via fixed or sliding premiums), or (3) tied to a separate market (i.e., quota obligations and tradable compliance instruments, such as renewable portfolio standards and associated renewable energy credits).

- The quantity of production that is supported by the remuneration scheme may be limited (for example, per-MWh payments may be available only up to a benchmark limit, after which any additional production only receives the market price).

31 Auctions for long-term supply contracts held in August 2016 in Chile resulted in a significant amount of wind and solar PV, which have displaced other technologies. As reported in the Wall Street Journal (Dube 2016): "Renewable-energy companies (...) won many of the contracts to supply 12,430 gigawatt hours per year of electricity over 20 years starting in 2021. The average cost of the contracts is $47.59 per megawatt hour, which was 63% lower than the average during the previous administration's last auction, Energy Minister Máximo Pacheco said."

Most production-based incentive schemes provide renewable generators with incentives to produce even when market prices are negative. This can turn out to be highly inefficient for the system. Negative prices have been a major concern in systems where high levels of renewable generation are being supported with production-based mechanisms (Götz et al. 2014).

Alternatives exist to avoid this inefficiency even if the choice is to use production-based schemes. For example, if the production incentive is based on an increasing function of spot prices, the problem of negative prices disappears.[32] (In that case, however, the pure marginal price would still not be the guiding signal for short-term operation.) Other alternatives for minimizing the inefficiency that results from negative prices include capping total support payments in those periods with negative prices or directly banning negative bidding by generation that receives regulatory support.

Production-based schemes and market compatibility

Production-based schemes always suffer from the fundamental tension between creating revenue certainty (which at this stage should not be seen as such a particular problem, since the long-term risk position of investments in renewable projects does not essentially differ from that of investments in conventional generation) and incentives for high-quality projects on one hand, and limiting distortions on the other.

Paying for production alters market price signals, leading to changes in generators' operating behavior. This has a negative impact on the economic efficiency of the system.

Production-based schemes create incentives to roughly maximize total energy generation. However, as illustrated in **Box 7.10**, projects with the highest capacity factors do not necessarily produce the highest-value energy.

32 For example, a percentage of the market price would result in a penalty for producing during periods of negative prices.

7.3.2.2 Capacity-based mechanisms

Capacity-based subsidies are intended to support investments on a per-installed-MW basis. In principle, capacity-based subsidies avoid market distortions by decoupling payment and performance. These subsidies can be implemented following different criteria (for example, the per-installed-MW fixed subsidy can be proportional to cost), depending on the generation technology and scale being targeted. The remainder of this section focuses on how to design a capacity-based mechanism for a technology that is subject to wholesale market signals.

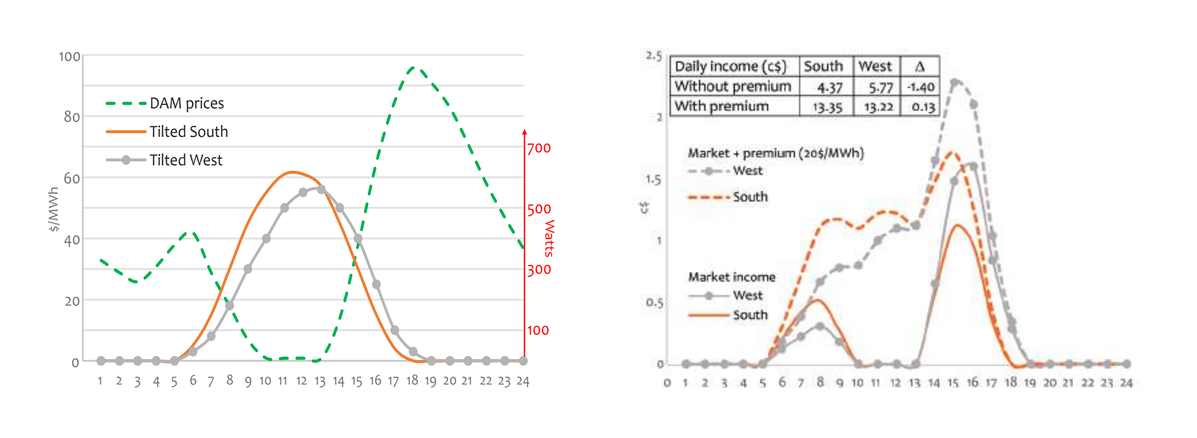
In cases where the supported technology perceives market signals (and there is therefore an associated market remuneration), these capacity- or investment-based subsidies are intended to cover (on a per-installed-MW basis) the difference between a plant's upfront investment cost and the present value of any expected market revenues after subtracting operating costs, such that the project is likely to be profitable and thus attractive to investors.

The cornerstone of these subsidization mechanisms is an estimate of how much a project will earn or need to pay by operating in various segments of the market (e.g., potentially including markets for capacity and other services). In practice, this estimate can be based on an *ex ante* forecast of future market revenues for a given renewable power plant over the course of its life, or conversely, compensation can be calculated *ex post*, based on market results. The problem with *ex post* compensation based on a plant's actual market performance is that compensation becomes unavoidably "production-based"—and therefore causes distortions. An alternative that avoids this problem is to define the *ex post* payment using a benchmark plant. The benchmark plant is a reference plant (synthetic or real) that would ideally reflect an efficient and well-managed installation. Market rents would be calculated for this reference plant and additional support would be estimated based on the amount determined to be needed to make that plant break

Box 7.10: Production-Based Subsidies and System Adaptation

Solar and wind developers are able to make decisions at the time of investment that influence the degree to which they align production and demand profiles. They have no incentive to optimize this alignment, however, as long as the emphasis is purely on total generation. We illustrate this effect using a quantitative example below. On the left we show the simulated energy output of 1 kW of solar PV capacity in Los Angeles, California depending on the orientation and tilt of the solar installation (EIA 2014) and hourly prices in the California power system (CAISO), simulated using the same model discussed in Chapter 8 of *The Future of Solar Energy* (MIT 2015). The right-hand graph shows hourly market income for the solar panel and income when a $20/MWh premium is added. The table includes daily income for each orientation. It shows that the energy-based subsidy leads to less efficient tilt because it results in higher income for the panel tilted toward the south.

Figure 7.10: Day-Ahead Market Prices As Well As Market and Premium Income for Solar Production for Different Panel Orientations



even. This would provide greater investment certainty without at the same time distorting short-term market signals. Furthermore, using real plants creates yardstick competitive pressure for renewable power plants to outperform their competitors: Since all plants receive the same support payment, those plants that perform better (because they are more efficient or have better forecasting abilities or a better sales strategy, etc.) will earn higher returns overall.

Avoiding the problem of paying for nothing

If a low-cost, low-performance facility (which typically is not the intended target of the support mechanism) turns out to be profitable with the support provided, then clearly additional rules will be necessary to achieve an efficient result. This can be tackled via minimum performance requirements (e.g., minimum operating hours) or minimum standards for component efficiency or system design (although setting such standards is challenging in light of rapidly evolving technology).

Capacity-based subsidies avoid market distortions by decoupling payment and performance. Using reference plants can mitigate the risk of payments flowing to low-quality projects.

Auction-based mechanisms are needed to bid the percentage of the reference plant compensation

Note that the use of a reference plant is compatible with different mechanisms to determine final remuneration, including price- or quantity-based mechanisms (such as auctions). For example, once the reference plant is established, an auction could be used for price discovery, where bids could consist of declaring the percentage of breakeven compensation needed by the reference plant (for example, 80 percent or 110 percent).

Reference plants can be defined using different criteria, such as location and size, and they can change over time to reflect technology changes.

7.3.2.3 Conclusions and recommendations on the design of support mechanisms for clean energy technologies

If the policy objective is to support a particular renewable technology, the approach we advocate would function as follows: Any new plant brought into operation would first be assigned to a particular reference facility that corresponds to the same technology type (wind, solar, etc.) and similar project size.

A reference support payment methodology is calculated for each reference facility based on the difference between the reference facility's market revenues and the annuity of its investment cost, plus an adder to ensure a reasonable rate of return. If possible, reference plants should be based on actual installed plants (e.g., the median 10 percent by market revenue or investment cost to better reflect actual bidding behavior).

Efficiency considerations argue in favor of auctioned capacity-based support payments that make use of reference facilities to estimate market revenues

Auctions, if workable, should be used in the procurement process to secure the most competitive projects.[33] Plants would bid a percentage of the reference support payment for a number of years. The conditions of this support should remain fixed over the duration of the contract to provide investors with greater revenue security. The use of regular auctions will result in updated reference investment costs that reflect changes in technology cost and performance, market conditions, and access to resource-rich locations for new projects (the best sites should gradually fill up).

The solution proposed allows for a prompt convergence with capacity mechanisms and creates a natural competitive pressure to optimize investment costs, siting, and production.

The capacity-based support scheme outlined above has several notable features. First, it is highly compatible with competitive electricity markets compared with a production-based scheme because it severs the link between production and payment: Only market signals are left to dictate operating decisions. This compatibility has to do not only with the short-term energy market but also with capacity mechanisms. Indeed, support for clean energy technologies should already be starting to converge with these mechanisms as the competitiveness of renewable technologies is approaching (and in some cases surpassing) that of conventional generation technologies (Mastropietro et al. 2014). Second, the approach we propose guards against the traditional risks of capacity-based schemes by combining performance requirements with natural incentives: Generators can increase their profits by beating the reference facility (e.g., through superior bidding strategies, better forecasting/fewer imbalances, superior location for wind/solar projects, etc.). This creates a natural competitive pressure to optimize production and siting decisions. Third, the use

33 As discussed above, in the absence of aggregation, price-based mechanisms might be preferred for small-scale projects, while quantity-based mechanisms are better suited for larger projects.

Box 7.11: Merging Capacity Mechanisms and (Capacity-Based) Renewable Support Mechanisms

Renewable technology support mechanisms are just a subset of capacity mechanisms. They aim to enhance investment in generation, albeit in very specific types of generation — only renewable technologies are eligible. This means that as the learning curves and maturity of renewable technologies continue to improve, and as there is an increasing need to coordinate renewable technology investments with investments in other, complementary, alternatives (e.g., flexible gas-fired plants or demand response or storage technologies), there is no reason to maintain separate mechanisms.

At the same time, auctioned capacity-based support payments should be the preferred alternative to support renewables (and other immature technologies). One of the reasons to favor this approach is that it would allow for prompt convergence with capacity mechanisms.

In systems that have chosen to implement capacity mechanisms and pursue a renewable penetration objective, there is, at this stage, a straightforward way to make these two mechanisms converge: Add a constraint in the capacity auction in such a way that a minimum amount of renewable investment would be guaranteed.

A reliability product should be defined, aimed at acknowledging the actual contribution of each technology to the future reliability of the system (for example, such a product would reward actual production at times when the market price reveals that the system has reached a near rationing event). All technologies in place should be exposed to the same commitment. If, for some reason, there are higher-level policy objectives to assure a minimum penetration of one technology in particular, a descending-clock auction could be easily designed to clear this amount. In that case, two different prices would arise: one for the promoted technology (higher, if the constraint is activated) and another price for the rest.

This solution presents numerous advantages: Once the auction is cleared, it exposes all technologies to the same incentives. More importantly, it guarantees that once the learning curve of the promoted technology improves sufficiently to allow the technology to fully compete with the others, this fact is naturally revealed by the auction, as the constraint ends up being inactive.

of auctions introduces competition into the procurement process, ensuring that investment costs are not inflated in an attempt to secure larger support payments.

Properly designing all these sorts of schemes is fairly complex, especially compared with simpler production-based mechanisms that have dominated support schemes in many parts of the world for decades. But this highlights an important point: Delivering subsidies effectively is challenging. Simple and straightforward support payments too often lead to market distortions and inefficient outcomes. The support schemes of the future will necessarily have to be more sophisticated to spur investment without obscuring the benefits of competitive markets.

References

Artelys. 2015. *France-Germany Study: Energy Transition and Capacity Mechanisms.* Report for the Union Française de l'Electricité and the German Association of Energy and Water Industries. ufe-electricite.fr/.

Batlle, C., C. Vázquez, and J. Barquín. 2007. "The Balancing Market Maker: A Tool to Enhance Liquidity in the French Balancing Mechanism." *The Electricity Journal* 20 (6): 76-86. July 2007.

Batlle, C., P. Mastropietro, P. Rodilla, and J.I. Pérez Arriaga. 2015. "The System Adequacy Problem: Lessons Learned from the American Continent." In *Capacity Mechanisms in the EU Energy Market: Law, Policy, and Economics:* 119-139. Oxford, England: Oxford University Press.

Barroso, L. A. and C. Batlle. 2011. "Review of Support Schemes for Renewable Energy Sources in South America." *IAEE Third Quarter* 2011: 27-31. International Association for Energy Economics.

Bird, L., M. Milligan, and D. Lew. 2013. *Integrating Variable Renewable Energy: Challenges and Solutions.* National Renewable Energy Laboratory. NREL/TP-6A20-60451. September 2013. www.nrel.gov.

Boot, P. and B. van Bree. 2010. *A Zero Carbon European Power System in 2050: Proposal for a Policy Package.* ECN-E — 10-041. Petten, Netherlands: Energy Research Centre of the Netherlands. April 2010.

Butler, L. and K. Neuhoff. 2008. "Comparison of Feed-in Tariff, Quota and Auction Mechanisms to Support Wind Power Development." *Renewable Energy* (33): 1854–1867.

Cochran, J., L. Bird, J. Heeter, and D. Arent. 2014. *Integrating Variable Renewable Energy in Electric Power Markets: Best Practices from International Experience.* National Renewable Energy Laboratory. NREL/TP-6A00-53732. April 2012. www.nrel.gov.

Couture, T.D., D. Jacobs, W. Rickerson, and V. Healey. 2015. *The Next Generation of Renewable Energy Policies: How Rapid Change Is Breaking Down Conventional Policy Categories.* National Renewable Energy Laboratory. NREL/TP-7A40-63149. February 2015. www.nrel.gov.

DECC. 2013. *Electricity Market Reform: Capacity Market — Detailed Design Proposals.* Cm 8637. Department of Energy & Climate Change. June 2013. www.gov.uk.

DECC. 2014. *Implementing Electricity Market Reform (ERM) — Finalised Policy Positions for Implementation of EMR.* Department of Energy & Climate Change. June 2014. www.gov.uk.

EFET. 2013. *Design Principles for Capacity Mechanisms.* European Federation of Energy Traders. February 2013. www.efet.org.

EIA. 2014. "Solar Photovoltaic Output Depends on Orientation, Tilt, and Tracking." *Today in Energy.* U.S. Energy Information Administration. November 19, 2014. www.eia.gov.

European Commission. 2013. *European Commission Guidance for the Design of Renewables Support Schemes.* Commission Staff Working Document 439 final. Brussels. November 2013. ec.europa.eu/.

European Commission. 2014. "Guidelines on State Aid for Environmental Protection and Energy 2014-2020." *Official Journal of the European Union* (2014/C 200/01). June 2014. eur-lex.europa.eu/.

European Commission. 2016. *Interim Report of the Sector Inquiry on Capacity Mechanisms. Report from the Commission.* Commission Staff Working Document 119 final. Brussels, Belgium. April 13, 2016. ec.europa.eu/.

FERC. 2013. *Centralized Capacity Market Design Elements.* Commission Staff Report AD13 7 000. US Federal Energy Regulatory Commission. August 23, 2013. www.ferc.gov.

FERC. 2014. *Operator-Initiated Commitments in RTO and ISO Markets.* Staff Analysis in Docket No. AD14-14-000. US Federal Energy Regulatory Commission. December 2014. www.ferc.gov.

FERC. 2016. *Settlement Intervals and Shortage Pricing in Markets Operated by Regional Transmission Organizations and Independent System Operators.* Order No. 825, 155 FERC 61,276. US Federal Energy Regulatory Commission. June 16, 2016. www.ferc.gov.

Finon, D. and F. Roques. 2012. *European Electricity Market Reforms: The "Visible Hand" of Public Coordination.* September 2012.

Gottstein, M. and L. Schwartz. 2010. "The Role of Forward Capacity Markets in Increasing Demand-Side and Other Low-Carbon Resources: Experience and Prospects." *Roadmap 2050*. The Regulatory Assistance Project. May 2010.

Götz, P., J. Henkel, T. Lenck, and K. Lenz. 2014. "Negative Electricity Prices: Causes and Effects. An Analysis of Recent Developments and a Proposal for a Flexibility Law." Agora Energiewende. 049/07-A-2014/EN. August 2014. www.agora-energiewende.de.

Gouveia, J. P., L. Dias, I. Martins, and J. Seixas. 2014. "Effects of Renewables Penetration on the Security of Portuguese Electricity Supply." *Applied Energy* 123: 438-447.

Hagemann, S. and C. Weber. 2015. "Trading Volumes in Intraday Markets - Theoretical Reference Model and Empirical Observations in Selected European Markets." EWL Working Paper No. [03/15]. University of Duisburg-Essen. April 2015. www.ewl.wiwi.uni-due.de/en/.

Harris, D.M. 2015. *Large-Scale Renewable Energy Development in New York: Options and Assessment.* NYSERDA Report 15-12. New York State Energy Research and Development Authority. June 1, 2015. www.nyserda.ny.gov.

Helm, D. 2010. "Market Reform: Rationale, Options and Implementation." Policy Paper. October 8, 2010.

Helman, U., B.F. Hobbs, and R.P. O'Neill. 2008. "Chapter 5 - The Design of US Wholesale Energy and Ancillary Service Auction Markets: Theory and Practice." In Sioshansi, F.P. (Ed.), *Competitive Electricity Markets.* Elsevier Global Energy Policy and Economics Series: 179–243.

Henriot, A. and J.M. Glachant. 2015. *Electricity Markets and RES Integration — Key Challenges and Possible Solutions*. D5.1 Report. IEE project towards 2030-dialogue (work package 5). Intelligent Energy - Europe. ALTENER. March 2015. www.towards2030.eu.

Herrero, I., P. Rodilla, and C. Batlle. 2015. "Electricity Market-Clearing Prices and Investment Incentives: The Role of Pricing Rules." *Energy Economics* 47: 42-51. January 2015.

Herrero, I., P. Rodilla, and C. Batlle. 2016. "Enhancing Intraday Price Signals in U.S. ISO Markets for a Better Integration of Variable Energy Resources." MIT Energy Initiative Utility of the Future Working Paper MITEI-WP-2016-05. May 2016. energy.mit.edu/uof.

Huntington, S.C., P. Rodilla, and C. Batlle. 2016. "Revisiting Support Policies for RES-E Adulthood: Towards Market Compatible Schemes." MIT Utility of the Future Working Paper MITEI-WP-2016-07. September 2016. energy.mit.edu/uof.

IEA-RETD. 2016. "RE TRANSITION — Transitioning to Policy Frameworks for Cost-Competitive Renewables." IET — International Energy Transition GmbH, International Energy Agency Technology Collaboration Programme for Renewable Energy Technology Deployment (IEA-RETD). Utrecht, the Netherlands. iea-retd.org/.

Madani, M., M. Van Vyve, A. Marien, M. Maenhoudt, P. Luickx, and A. Tirez. 2016. "Non-convexities in European Day-ahead Electricity Markets: Belgium as a Case Study." 13th International Conference on the European Energy Markets. 1–5. doi:10.1109/EEM.2016.7521190.

Mastropietro, P., C. Batlle, L.A. Barroso, and P. Rodilla. 2014. "Electricity Auctions in South America: Towards Convergence of System Adequacy and RES-E Support." *Renewable and Sustainable Energy Reviews* 40: 375-385.

Maurer, L. and L. Barroso. 2011. *Electricity Auctions: An Overview of Efficient Practices.* World Bank Publications. ISBN 978 0 8213 8822 8.

MIT. 2015. "Integration of Solar Generation in Wholesale Electricity Markets." *The Future of Solar Energy*. Cambridge, MA: Massachusetts Institute of Technology. ISBN (978-0-928008-9-8). May 2015. energy.mit.edu/publication/future-solar-energy/.

Neuhoff, K., R. Boyd. 2011. "International Experiences of Nodal Pricing Implementation." Berlin, Germany: Climate Policy Initiative. July 2011. www.climatepolicyinitiative.org.

Neuhoff, K., J. Barquin, J.W. Bialek, R. Boyd, C.J. Dent, F. Echavarren, T. Grau, C. von Hirschhausen, B.F. Hobbs, F. Kunz, C. Nabe, G. Papaefthymiou, C. Weber, and H. Weigt. 2013. "Renewable Electric Energy Integration: Quantifying the Value of Design of Markets for International Transmission Capacity." *Energy Economics* 40: 760–772.

Neuhoff, K., N. Ritter, A. Salah-Abou-El-Enien, and P. Vassilopoulos. 2016. "Intraday Markets for Power: Discretizing the Continuous Trading?" DIW Discussion Paper 1544. doi:10.2139/ssrn.2723902.

O'Neill, R., A. Castillo, B. Eldridge, R.B. Hytowitz. 2016. "Dual Pricing Algorithm in ISO Markets." *IEEE Transactions on Power Systems* PP(99): 1–1. doi:10.1109/TPWRS.2016.2614891.

Orme, B. 2016. "Incapacitated: Why the Capacity Market for Electricity Generation Is Not Working, and How to Reform It." Institute for Public Policy Research. March 2016. www.ippr.or.

PCR. 2016. Main Presentation in Price Coupling of Regions Stakeholder Forum. Brussels, Belgium. www.apxgroup.com/wp-content/uploads/Euphemia-Stakeholder-Forum-11-Jan-2016.pdf.

PJM. 2013. PJM "Ancillary Services" online training course. Pennsylvania New Jersey Maryland Interconnection. www.pjm.com/training/course-catalog.aspx.

PJM. 2014. 2017/2018 *RPM Base Residual Auction Results.* PJM DOCS No. 794597. Pennsylvania New Jersey Maryland Interconnection. May 23, 2014. www.pjm.com.

SEDC. 2014. *Mapping Demand Response in Europe Today Tracking Compliance with Article 15.8 of the Energy Efficiency Directive.* Smart Energy Demand Coalition. April 2014. www.smartenergydemand.eu.

Stoft, S. 2003. "The Demand for Operating Reserves: Key to Price Spikes and Investment." *IEEE Transactions on Power Systems* 18: 470–477.

UPME. *Unidad de Planeación Minero Energética, 2015. Plan de expansión de referencia: Generación — Transmisión, 2015-2029* (in Spanish). www1.upme.gov.co/planes.

Usera I., P. Rodilla, and C. Batlle. 2016. "Regulatory Framework for Energy Storage Systems." MIT Energy Initiative Utility of the Future Working Paper.

Walton, R. 2016. "New PJM Capacity Rules Challenge Demand Response Aggregators." *Utility Dive Newsletter.* June 8, 2016. www.utilitydive.com.

Weissbach, T. and E. Welfonder. 2009. "High Frequency Deviations within the European Power System: Origins and Proposals for Improvement." In 2009 IEEE/PES Power Systems Conference and Exposition.

PART 3: INSIGHTS ON THE ECONOMICS OF DISTRIBUTED ENERGY RESOURCES AND THE COMPETITION BETWEEN CENTRALIZED AND DISTRIBUTED RESOURCES

08

Understanding the Value of Distributed Energy Resources

8.1 Introduction

Preceding chapters of this study have outlined a framework for proactive reform designed to enable the efficient evolution of the power system over the next decade and beyond. The goal of this framework is to facilitate the cost-effective integration of both distributed and centralized resources. The framework includes four elements: (1) a comprehensive system of prices and regulated charges for electricity services that sends sufficiently granular signals about the temporal and locational value of resources (**Chapter 4**); (2) improved regulation for distribution utilities that incentivizes integration of cost-effective alternatives to conventional network investments, rewards utilities for cost savings and performance improvements, and encourages ongoing innovation (**Chapter 5**); (3) reconsideration of the structure of the electricity sector, with a focus on the role of distribution utilities and on minimizing potential conflicts of interest (**Chapter 6**); and (4) upgrades to wholesale electricity and ancillary service markets to remove unnecessary barriers to the participation of distributed resources (**Chapter 7**).

This framework is designed to create a level playing field for all resources, centralized and distributed, to contribute to the cost-effective provision of electricity services. It should also ensure that resources are properly compensated and charged for the services they provide and consume, revealing when, where, and how these resources deliver value or drive costs in electric power systems. With this framework in place, this chapter focuses on understanding the value of distributed energy resources (DERs) and providing insights about the factors that are most likely to determine the portfolio of cost-effective resources, both centralized and distributed, in different power systems.

Section 8.2 revisits the taxonomy of electricity services introduced in **Chapter 2** and focuses on the ways in which the value of some of these services changes depending on where these services are provided or consumed within the power system. This focus on "locational value" is key to understanding when, where, and how DERs can provide potentially significant additional value relative to larger-scale centralized resources. **Section 8.3** discusses the potential for unlocking contributions from existing resources, including flexible and price-responsive demand and existing power electronics, such as smart inverters. For resources that can be deployed at multiple scales, including solar photovoltaics (PV) and energy storage devices, trade-offs between locational value and economies of unit scale determine the optimal scale and location of deployment; we discuss these important trade-offs in **Section 8.4**.

8.2 Locational Value and Distributed Energy Resources

As discussed in **Chapter 2**, the range of electricity services provided by different energy resources can be organized into a set of services with either locational or non-locational value (**Table 8.1**). The value of some electricity services differs substantially depending on the location where the service is provided or consumed within the power system. These differences in value emerge from physical characteristics of electricity networks, including losses, capacity limits of network components, and voltage limits at network nodes. We refer to services that exhibit this quality as having "locational value."

The value of some electricity services differs substantially depending on the location where the service is provided or consumed within the power system. These differences in value emerge from physical characteristics of electricity networks, including losses, capacity limits of network components, and voltage limits at network nodes. We refer to services that exhibit this quality as having "locational value."

Table 8.1: Classification of Values

	LOCATIONAL	NONLOCATIONAL
POWER SYSTEM VALUES	• Energy • Network capacity margin • Power quality • Reliability and resiliency • Black-start	• Firm generation capacity • Operating reserves[1] • Price hedging
OTHER VALUES	• Land value/impacts • Employment • Premium values[2]	• CO_2 emissions mitigation • Energy security

The value of other services is nonlocational — that is, the value of the service does not change based on where it is delivered in the power system. For example, operating reserves are deployed to contain frequency deviations that emerge as a result of imbalances between supply and demand due to forecast errors or unexpected failures of power plants or transmission lines. Except at very short time scales, frequency is consistent across an entire synchronized interconnected system. As a result, the value of frequency regulation is typically consistent across that entire system. Similarly, the value of mitigating carbon dioxide emissions depends on the social cost of carbon (EPA 2016) and does not change if the resource driving emissions reductions is distributed or centralized.

DERs can provide a range of services, including services with locational and nonlocational value. Due to their distributed nature, DERs can be sited and operated in areas of the power system where the services they provide are most valuable. Understanding the specific

1 The value of firm capacity and operating reserves may differ by location when frequent network constraints segment electricity networks and prevent delivery of capacity or reserves to constrained locations. This leads to the designation of zonal requirements for capacity and reserves in some jurisdictions. However, even in such cases, the value for these services are uniform within these zones and thus remain constant across wide geographic areas. As such, these services can be considered non-locational at least within broad segments of a given interconnected transmission system.

2 "Premium value" is a catch-all term that refers to the value that DER owners may derive from factors that are inherent to DERs but that are not directly associated with the electricity services that a given DER provides. For example, some consumers may derive value from independently producing their own power or aligning their electricity production with personal values (such as environmental attributes). In such cases, these "premium values" are private—that is, they accrue only to the individual(s) who own the DER(s). Accordingly, these premium values should not be internalized by public policies or electricity tariff design.

services that have locational value is thus critical to understanding the ways that DERs can create value and is the focus of the remainder of this section. As we demonstrate, the value of certain services can differ by orders of magnitude even within the same wholesale market or transmission system, and even within a given distribution network. Just as there is no single value for the services that DERs provide, there is no single value for DERs. Indeed, the value that these resources can provide is extremely context-dependent. Additionally, this value can decline quickly as more resources are deployed or activated to supply a given service in a particular location and time. To accurately value the services provided by DERs, prices, regulated charges, and other incentives must therefore reflect the marginal value of these services to the greatest extent practical (see **Chapter 4.6**).

8.2.1 Magnitude and distribution of locational values

In this section, we describe the magnitude and variation of locational value associated with electrical energy, network capacity margins and network investment deferral opportunities, reliability, and other values. Note that this section proceeds from the perspective of power system value—that is, it focuses on the true marginal value or cost of these services and takes a system-wide perspective. In practice, the monetized or private value of these services for any particular agent may differ from the power system value, as services are compensated and remunerated with a variety of market mechanisms, regulated tariffs and charges, and policy incentives. These mechanisms and charges or tariffs rarely reflect with perfect accuracy the marginal value to the system of each service at each point in time and each location in the system (see **Chapters 4** and **7** for further discussion). Nonetheless, these services provide value to the power system, regardless of how that value is monetized for the various agents who provide the services, and we take the perspective of this system value throughout this section.

Due to their distributed nature, DERs can be sited and operated in areas of the power system where the services they provide are most valuable. Understanding the specific services that have locational value is thus critical to understanding the ways that DERs can create value.

8.2.1.1 Electrical energy: Locational value due to constraints and losses

Due to the impact of network losses and constraints on the delivery of electricity, the value of electrical energy consumption or injection varies at different points in the power system. As discussed in **Chapter 4**, this variation in the locational value of electrical energy is well-explained by the theory of spot pricing of electricity (e.g., Schweppe et al. 1988), which yields locational marginal prices (LMPs) at each node in the power system.[3]

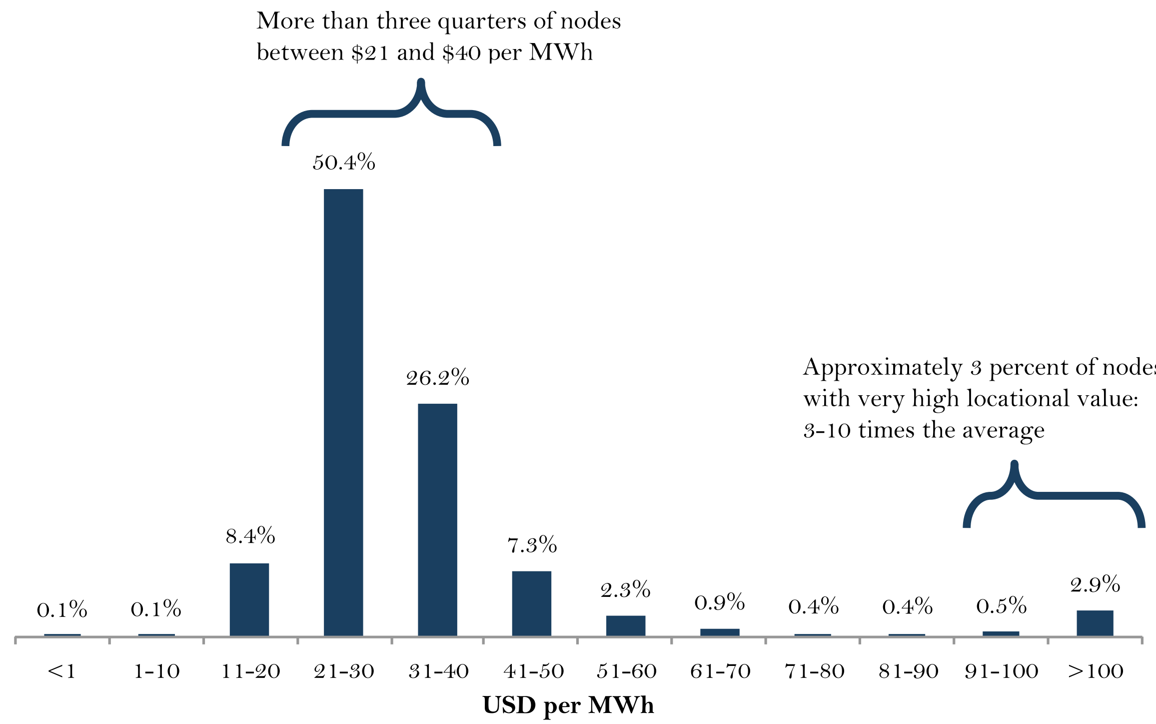
DERs therefore have the potential to create significant value by supplying energy at locations where networks are frequently constrained and/or transmission and distribution losses are large. When averaged across an entire year, the variation in locational value for energy due to losses and congestion at the transmission level generally clusters within a range extending about $10 per megawatt-hour (MWh) above or below an average value in most markets. As **Figure 8.1** illustrates, the average annual LMP at more than three-quarters of the nodes in the PJM transmission network ranged between $21 and $40 per MWh in 2015. At 3 percent of the nodes in PJM, however, the locational value of energy is three to ten times larger than the average value. Resources located at these nodes could therefore create much greater value by selling energy than they could by selling energy at an average node.

The highest values in the distribution of average wholesale LMPs are driven almost entirely by network constraints.[4] Constraints on power flows in transmission and distribution networks (commonly referred to as congestions) can occur as a result of line or transformer thermal capacity limits, node voltage constraints, or stability limits. These constraints result in different prices for energy delivery on either side of the binding constraint, reflecting the different marginal costs of supplying the next unit of electrical energy at each location. LMPs can rise significantly at certain times and locations where constraints preclude access to lower-cost generators and force reliance on high-cost power plants or demand response to meet marginal demand. The impact of network constraints on the locational value of energy thus varies over time as well. To capture higher locational value due to network constraints, DERs must be able to operate both *where* and *when* constraints are binding.

3 As discussed in the introduction to this section, while LMPs accurately reflect the marginal locational value of electrical energy at each location and each point in time, these prices are not uniformly used to compensate or charge producers and consumers of electricity in practice. For example, European markets use zonal locational prices that lack nodal resolution, and nodal prices are not at present employed at the distribution system level. However, whether market designs or tariff choices accurately reflect this value or not, the marginal value (or cost) of electrical energy to the power system is accurately reflected in LMPs. Accordingly, we use these prices to discuss the power system value of electrical energy. Private value for any given agent will reflect how these services are monetized in any given power system.

4 Absent network constraints, differences in LMPs are driven only by resistive losses. The impact of marginal transmission losses on wholesale LMPs is typically a few percent of the cost of the marginal generator. The impact of distribution losses can be larger, as resistive losses are greater in lower voltage networks, as discussed below.

Figure 8.1: Distribution of 2015 Average Nodal LMPs in PJM

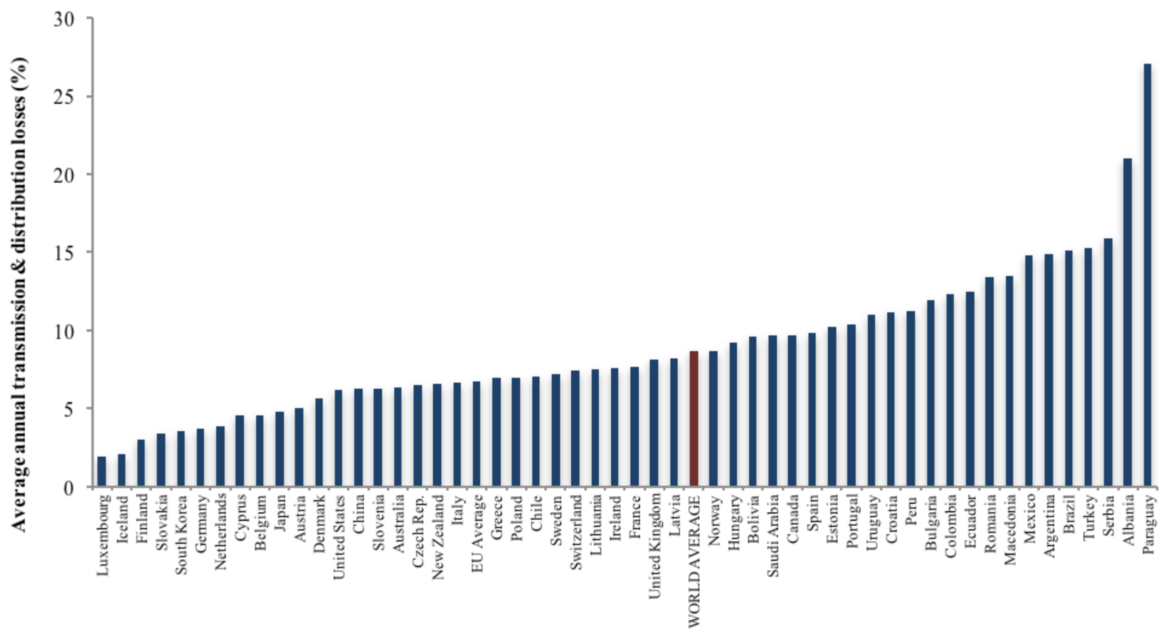


Losses occur due to resistance in transmission and distribution lines, transformers, and other electricity network infrastructure. Losses also contribute to the locational value of energy, although they generally have a smaller effect than network constraints. Total transmission and distribution losses (including technical and nontechnical losses) averaged 6.2 percent in the United States and 6.7 percent in the European Union in 2014, although average losses can range from as low as a few percent up to double digits in some countries (**Figure 8.2**).[5] The bulk of these losses occurs in distribution networks, which operate at much lower voltages than transmission networks (increasing the current associated with a given power flow) and are more resistive (increasing losses associated with a given current).

To capture locational value due to network constraints, DERs must be able to operate both *where* and *when* constraints are binding.

5 Losses reported in **Figure 8.2** include both technical losses due to resistance and nontechnical losses from theft or other unmetered consumption. In some countries, nontechnical losses account for a significant share of total losses—hence the larger values in this figure may reflect significant nontechnical losses. Locational energy value is affected only by marginal technical losses. Nontechnical losses should thus be disregarded when considering the locational value of energy.

Figure 8.2: Average Annual 2014 Transmission and Distribution Losses in Select Countries



Source: EnerData (2015)

Total resistive losses rise quadratically with power flow as described by **Equation (1)**:

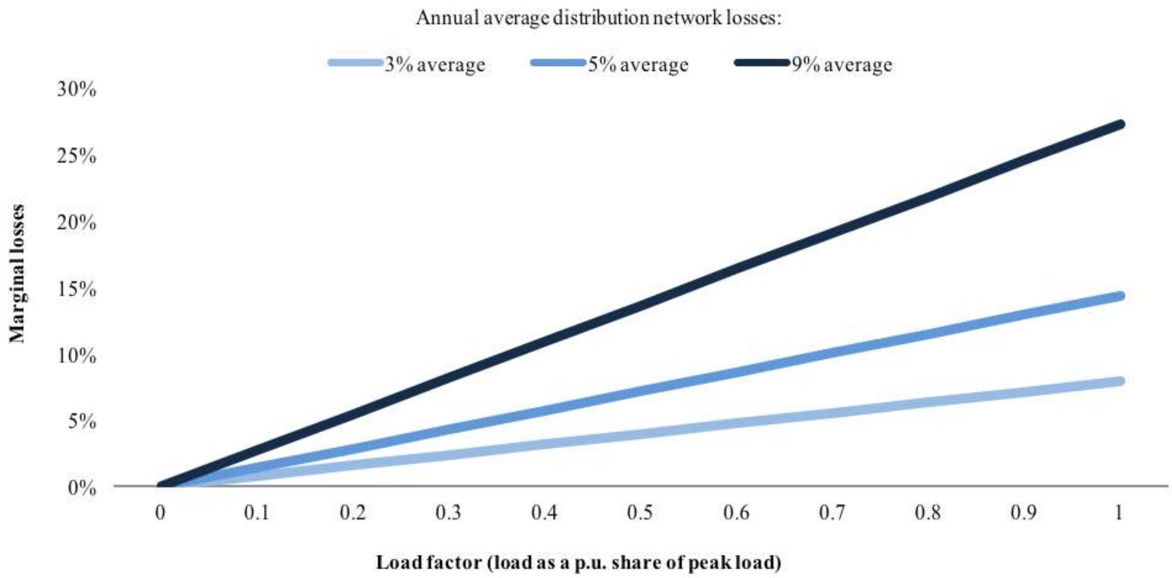
(1) $$P_{loss} = I^2 \mathrm{R} = \frac{\mathrm{R}}{\mathrm{V}^2} P^2$$

Where P_{loss} is total resistive losses, I is electrical current, R and V are resistance and voltage for the network branch, and P is instantaneous power. Thus marginal losses as a function of power are the derivative of **Equation (1)** with respect to power and rise linearly as power increases:

(2) $$\delta/_{\delta P} P_{loss} = 2 \frac{\mathrm{R}}{\mathrm{V}^2} P$$

In times of high demand and therefore high power flow, marginal losses can become significantly greater than in times of low demand and low power flow (**Figure 8.3**). DERs that generate power (or reduce net consumption) at times and locations of heavy network loading and high marginal losses can thus capture higher locational value. Note however that while DERs are typically assumed to reduce losses, they can, in some cases, contribute to losses in the distribution network by creating reverse power flows across lower distribution voltages (Schmalensee et al. 2015; Goop et al. 2016).

Figure 8.3: Marginal Distribution Losses as a Function of Load Factor in Distribution Networks with Differing Annual Average Technical Losses[6]

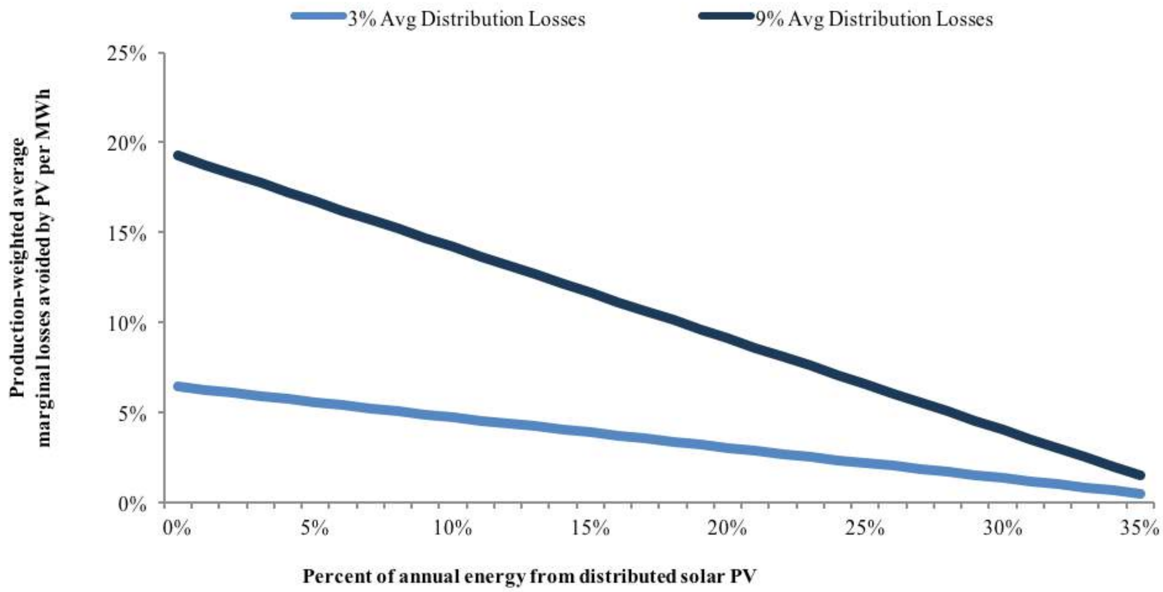


Losses also contribute to the locational value of energy, although they generally have a smaller effect than network constraints. However, DERs that generate power (or reduce net consumption) at times and locations of heavy network loading and high marginal losses can capture higher locational value.

In addition, because marginal losses decrease as the current or power flow through lines declines, the marginal value of loss reduction falls as more DERs produce power or reduce consumption during periods of high network loading. As an example, **Figure 8.4** illustrates the decline in the locational value of distributed solar PV systems due to distribution loss reductions as the share of PV in low-voltage networks increases. This example presents results for a case study that employs load and solar PV profiles consistent with a distribution network in the Electric Reliability Council of Texas (ERCOT) system. Initially, the production-weighted average marginal losses avoided by each MWh of solar PV output ranged from approximately 6 percent to 19 percent in distribution networks with average annual losses ranging from 3 percent to 9 percent. In other words, by reducing marginal distribution losses, each MWh produced by the first few solar PV systems located in the low-voltage portion of a distribution network in ERCOT is worth on average 6–19 percent more than the average MWh of solar PV produced at transmission voltage. However, as solar PV production increases in a given hour, the marginal losses in that hour steadily fall, reducing the marginal value of each additional unit of solar output.

6 As distribution losses are only rarely metered and reported, we estimate distribution losses for **Figure 8.3** by representing the distribution network as a single resistive device. Using aggregate zonal load profiles from the independent system operator for New York (NYISO) and the relationship between power and losses in **Equation (1)**, we establish the equivalent R/V^2 coefficient that corresponds to a given annual average technical distribution losses. In each case, we assume transformer core losses represent a constant 0.5 percent of average load.

Figure 8.4: Marginal Distribution Network Losses Avoided by Distributed Solar PV (Texas ERCOT Example)[7]



Indeed, at high penetration levels, solar PV systems can cause reverse power flows, and losses in distribution networks start to increase again (Schmalensee et al. 2015; Goop et al. 2016). As the share of solar PV increases to 15 percent of annual energy consumed in low-voltage networks, the production-weighted average marginal losses avoided by PV thus declines to roughly 4–12 percent. At 30 percent share of annual energy, this value falls to less than 1.5 percent, with solar causing reverse power flows and thus contributing to losses again in 640 hours of the year. For additional analysis of the locational value of distributed solar PV related to the reduction of transmission and distribution losses, see **Box 8.1**.

7 As with **Figure 8.3**, we estimate distribution losses for **Figure 8.4** by representing the distribution network as a single resistive device. Using aggregate zonal load profiles from ERCOT and the relationship between power and losses in **Equation (1)**, we establish the equivalent R/V^2 coefficient that corresponds to a given annual figure for average technical distribution losses. In each case, we assume transformer core losses represent a constant 0.5 percent of average load. Solar PV profiles are from the National Renewable Energy Laboratory's PVWatts Calculator. We assume any hours with negative net load correspond to reverse power flows and increase resistive losses. For more on the impact of distributed PV systems on distribution network losses, see Schmalensee et al. (2015), Chapter 7.

Box 8.1: The Effect of Energy Losses on the Value of Centralized and Distributed Solar PV

By reducing energy losses in transmission and distribution networks, distributed solar PV may deliver greater locational value than larger-scale centralized solar farms. To illustrate the impact of energy losses on the value of centralized and distributed solar PV, we present two cases: (1) increasing distributed solar PV deployment in low-voltage distribution nodes, and (2) increasing centralized solar PV deployment in high-voltage transmission nodes. In addition, we consider the possible impact of flexible charging and discharging of aggregated electric vehicle (EV) fleets on the value of solar in both cases.[8]

In this case, the power system is segmented into five distinct voltage levels, with a simplified network representation (see **Appendix A** for details). This case study resembles a Spain-like test system in the 2020–2025 timeframe.[9] **Figure 8.5** shows the existing installed generation mix at each voltage level in this system. All conventional generation (including large hydro) is connected at high-voltage transmission and existing renewable sources are connected across a range of voltage levels. Least-cost unit commitment and real-time dispatch of all resources are modeled using the ROM model (see **Appendix A** for details), and LMPs for each voltage level are calculated using a direct current power flow with quadratic losses. Effective resistances are estimated for each voltage level to closely approximate reported annual losses in the real Spanish system, and load profiles are created for each voltage level consistent with aggregate profiles for Spain.[10]

In each case, the penetration of either distributed or centralized solar PV is steadily increased, and the effect on average energy losses and the market revenues of solar PV systems is calculated (assuming all resources are compensated at LMPs, including prices at the distribution level). Smart charging of EVs or energy storage devices may increase demand at low-price hours when the solar generation is available, increasing marginal prices received by solar owners and reducing the prevalence of solar generation curtailment (see Schmalensee et al. 2015). To illustrate this potential impact, we also model cases assuming a fleet of 1.5 million EVs representing around 2 percent of total electricity demand and connected at low voltage levels. In addition, once EVs are charged to fulfill mobility requirements (based on EV parameters and mobility data from Banez-Chicharro et al. [2014]), we model EVs as capable of providing energy back to the system in order to minimize total system costs.[11] The total power capacity of the EV fleet is equivalent to 3 percent of installed electricity-generating capacity in this system, although not all EVs can dispatch simultaneously due to mobility requirement constraints.

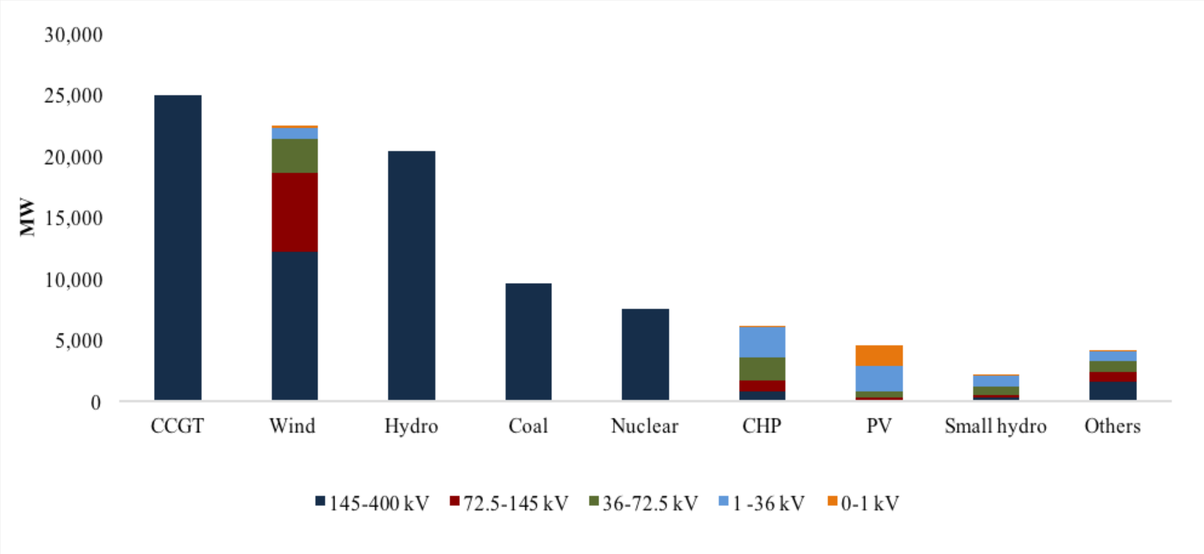
8 Note that as EV manufacturers typically do not support discharging from EVs to the grid at present, these effects could also be generalized to reflect the impact of other forms of energy storage at similar penetration levels.

9 The future generation scenario has been obtained from the Renewable Deployment Plan (Plan de Energía Renovables), approved by the Spanish government at the end of 2011 (IDAE 2011).

10 The Spanish regulator reports aggregated flows, losses, generation, and demand for an electric equivalent network for the whole country. Data are available at CNMC (2014).

11 The EV fleet dispatches to minimize total system costs, subject to constraints on battery state-of-charge and requirements that ensure individual vehicle mobility needs can be met. A round-trip charge–discharge efficiency of 90 percent is considered, but battery degradation is not accounted for in this example. This example may therefore overestimate the degree of discharge flexibility for EV fleets given current battery-life characteristics. If future technological improvement increases battery life, greater EV battery discharge may be cost-effective, consistent with this case study.

Figure 8.5: Assumed Installed Capacity in Spain-like System per Voltage Level



As **Figure 8.6** illustrates, increasing penetration of distributed solar PV at lower voltage levels can help reduce total transmission and distribution losses by reducing power flows across transmission and distribution lines. In contrast, if centralized solar PV is installed at high-voltage transmission nodes, the effect on losses is negligible, as centralized PV displaces other generation at transmission voltage level, with little effect on power flows. At high solar PV penetration levels (higher than 19 percent in our case study), losses start to increase due mainly to reverse flows in distribution networks (in the case of distributed PV) or replacement of distributed generation (in the case of centralized PV). Figure 8.6 also illustrates the same diminishing marginal effect of solar on losses discussed in **Section 8.2.1.1**. Furthermore, EV charging increases demand during periods of high solar output, attenuating the increase in losses due to reverse flows at higher penetration levels of distributed solar PV.

Figure 8.7 depicts average market revenues per MWh produced by both distributed and centralized PV at increasing PV penetration levels, as well as the impact of smart EV charging on solar revenues.[12] At low PV penetration levels, distributed solar earns greater market revenue than centralized solar due to the increase in locational value associated with avoided energy losses (i.e., higher LMPs at low-voltage nodes). However, as PV penetration increases, marginal losses decline during the periods when solar PV generates power, reducing the locational value of distributed PV. At higher PV penetration levels (around 10 percent in Figure 8.7), the locational value of distributed solar starts to become negative (i.e., lower LMPs at low-voltage nodes than at high-voltage nodes) and consequently market revenues of centralized solar become greater than those of distributed solar. Smart EV charging, by shifting stored energy, increases consumption at PV production hours and thus results in higher prices at those hours, thereby increasing PV market revenues for both distributed and centralized solar.

12 Once again, we assume that compensation for all resources reflects LMPs at each voltage level, including distribution. In practice, actual compensation may differ, depending on applicable market rules and tariff designs.

Figure 8.6: Effect of Increased Solar PV Penetration on Total Transmission and Distribution Losses (Spain-like Test Case Example)

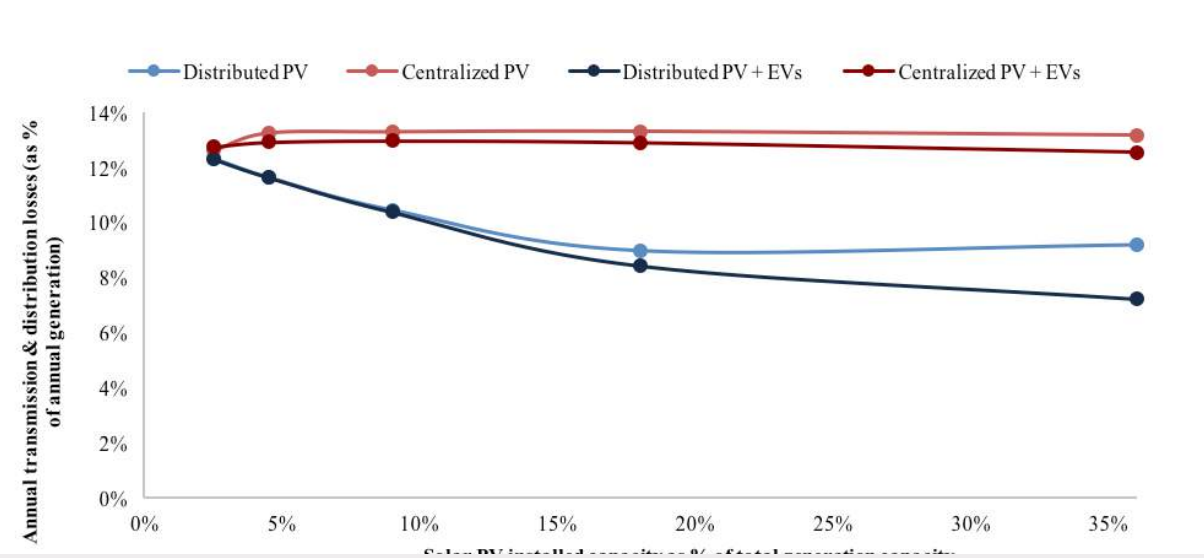
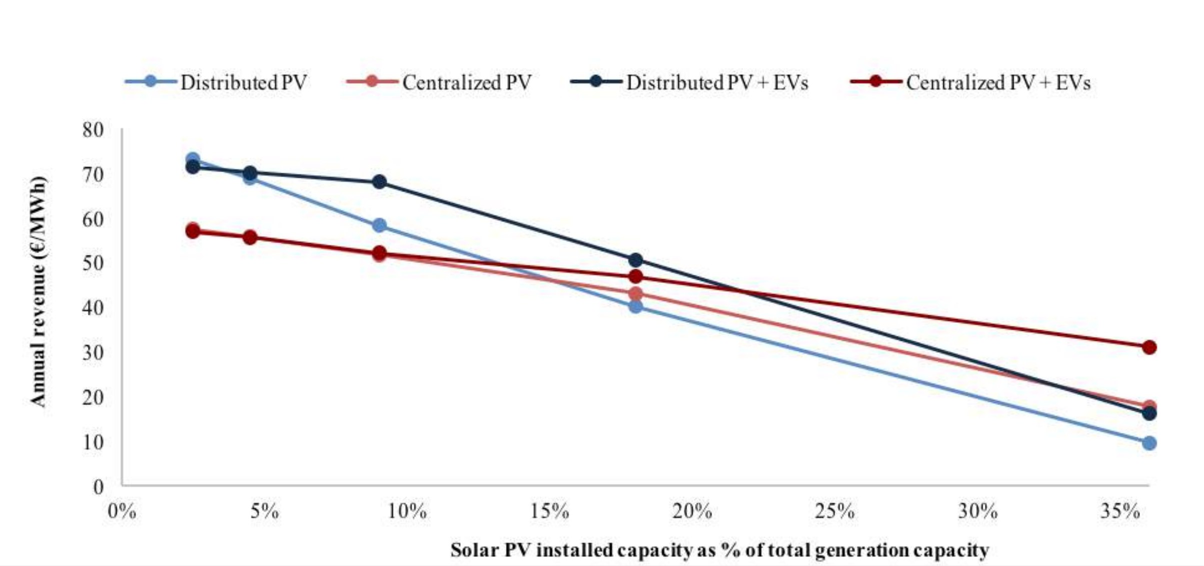


Figure 8.7: Annual Market Revenue per MWh for Solar PV Systems at Increasing Penetration (Spain-like Test Case Example)



8.2.1.2 Network capacity margin and network investment deferral

If sited in ideal locations and operated at ideal times, DERs can potentially deliver significant locational value by avoiding or deferring investments in transmission and/or distribution network capacity. During periods of peak aggregate power withdrawals (i.e., consumption peaks), DERs that can reliably reduce consumption or inject power can help the system avoid reaching network voltage or capacity limits, potentially enabling networks to operate closer to technical limits without requiring physical upgrades.[13] Likewise, if portions of the network experience constraints due to aggregate peaks in power injection (i.e., periods of reverse power flow), DERs can reduce injections or increase withdrawals during these periods to relieve these constraints. If DERs can deliver reductions in aggregate net withdrawals or injections with sufficient reliability,[14] network utilities may not need to expand network margins via traditional investments in network upgrades. The relative costs of both network upgrades and network capacity services offered by DERs must therefore be considered.

If DERs are sited in the ideal locations and operated at the ideal times — and can offer their services with sufficient reliability that distribution network operators can count on their output during periods of network stress—then DERs can substitute partially or fully for conventional investments in transmission and/or distribution infrastructure.

Using optimal power flow models of realistic European distribution networks, we estimate the minimum reduction in aggregate peak demand or withdrawals necessary to accommodate increasing levels of peak demand growth without violating network constraints.[15] This increment of peak demand growth that can be accommodated without investment in network infrastructure can be considered an effective increase in the network's "capacity margin." As **Figures 8.8** and **8.9** illustrate, if targeted to precisely the right locations and times, relatively small reductions in aggregate peak withdrawals in the low-voltage portion of distribution networks can accommodate modest peak demand growth without any additional investments.[16] In other words, if DERs can be sited at the ideal locations and operated at the ideal times—and can offer their services with sufficient reliability that distribution network operators can count on their output during periods of network stress—DERs can substitute in part or in full for conventional network investments. Indeed, pilot projects in various jurisdictions have demonstrated the real-world potential for active network management and dispatch of DERs to reduce or avoid the need for distribution network upgrades and manage both aggregate peaks in network withdrawals (demand peaks) and injections (generation peaks) (**Box 8.2**).

13 In practice, transmission and distribution network operators apply engineering safety margins to all network components to ensure that technical limits are not exceeded. If DERs can be reliably dispatched or counted on to respond to pricing signals, operators may relax these margins and rely on DER behavior to avoid violating network constraints. This would allow network components to be operated more closely to true technical limits and would allow the network to increase its capacity margin without a physical upgrade of network components. See **Section 2.2.1.3** for further discussion.

14 Note that DERs must be capable of providing these services with sufficient reliability that, in aggregate, operation of DERs can substitute for network investments without degrading quality of service. This may require performance contracts or other measures to guarantee sufficient performance and quality.

15 Urban and semi-urban networks from Prettico et al. (2016) are modeled in Matpower (Zimmerman et al. 2011) to calculate optimal power flows subject to alternating current power flow and network constraints. Curtailable loads (or equivalently, distributed generators) are simulated at each load point, and the objective of the model is to minimize the total net load reduction necessary to render power flows feasible under increasing levels of aggregate network load. Increments in aggregate network load are assigned in the following way: First, the network is divided into zones corresponding to each feeder (i.e., any aggregation of consuming nodes for which there is only a single connection point to a MV-LV substation). Then, for each zone/feeder, specific node or bus values are assigned in proportion to the peak load values of those specific buses in the pre-existing network. This is best illustrated with a simple example. Suppose that a specific network zone contains two consumers, A and B. In the pre-existing network, consumers A and B had peak loads of 60 kilowatts (kW) and 40 kW, respectively, and therefore a total peak load of 100 kW. Now suppose that this zone has been assigned a new peak load value that is equal to 1.1x the pre-existing peak load value, or 100 kW x 1.1 = 110 kW. In this case, since customer A's peak load in the pre-existing network accounted for 60 percent of that zone's total peak load, customer A's new peak load is 0.6 x 110 kW = 66 kW, whereas customer B's peak load is 0.4 x 110 kW = 44 kW.

16 If DERs or load curtailment occurs at suboptimal locations, greater total reductions in aggregate peak load are necessary than depicted in **Figure 8.8**. The figure thus presents an optimal case reflecting a lower bound on the required change in aggregate peak net withdrawals to accommodate a given increment of peak load growth.

Figure 8.8: Potential for DERs to Substitute for Network Upgrades in Representative European Distribution Networks (Low-Voltage Distribution Example)

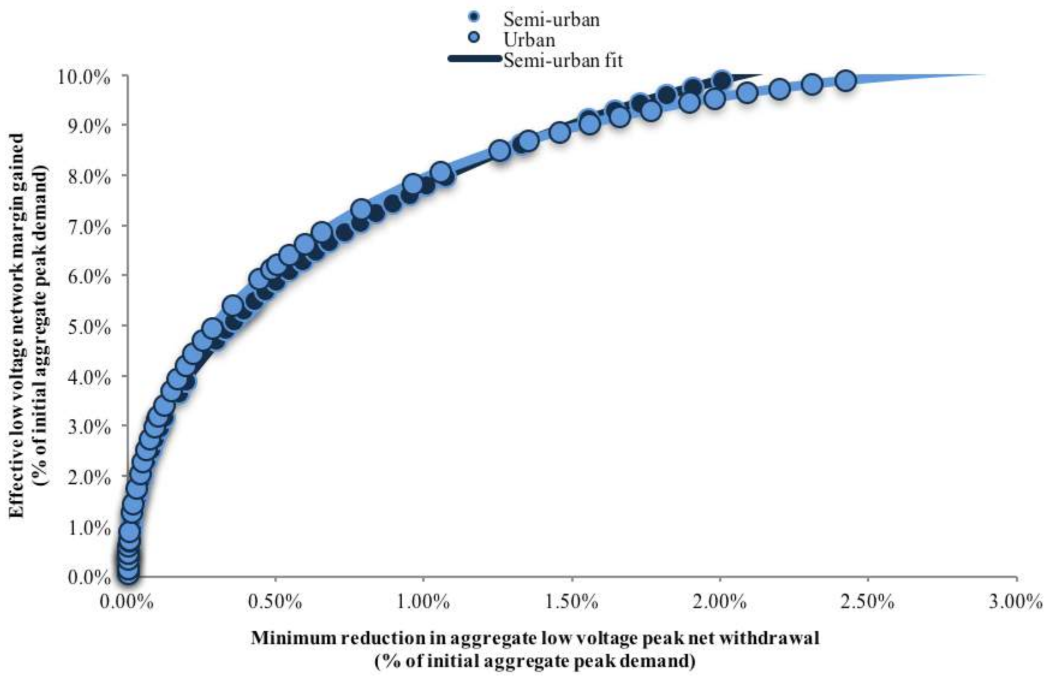
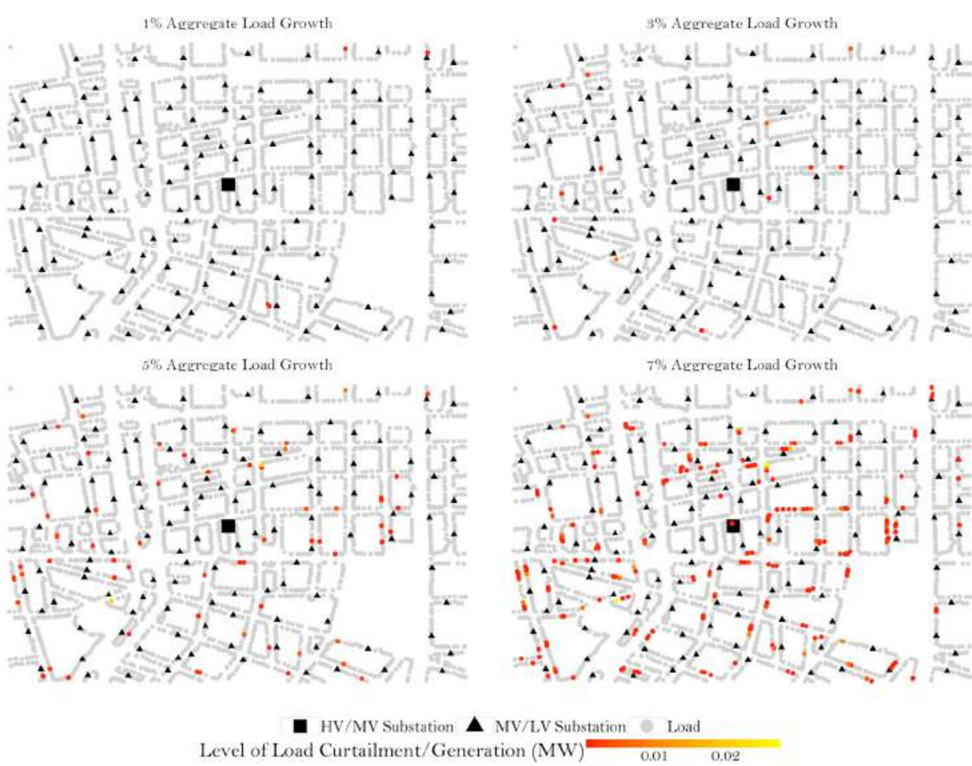


Figure 8.9: Optimal Location and Magnitude of Net Withdrawal Reductions Necessary to Accommodate Aggregate Peak Demand Increases Without Network Upgrades in a Representative European Urban Network



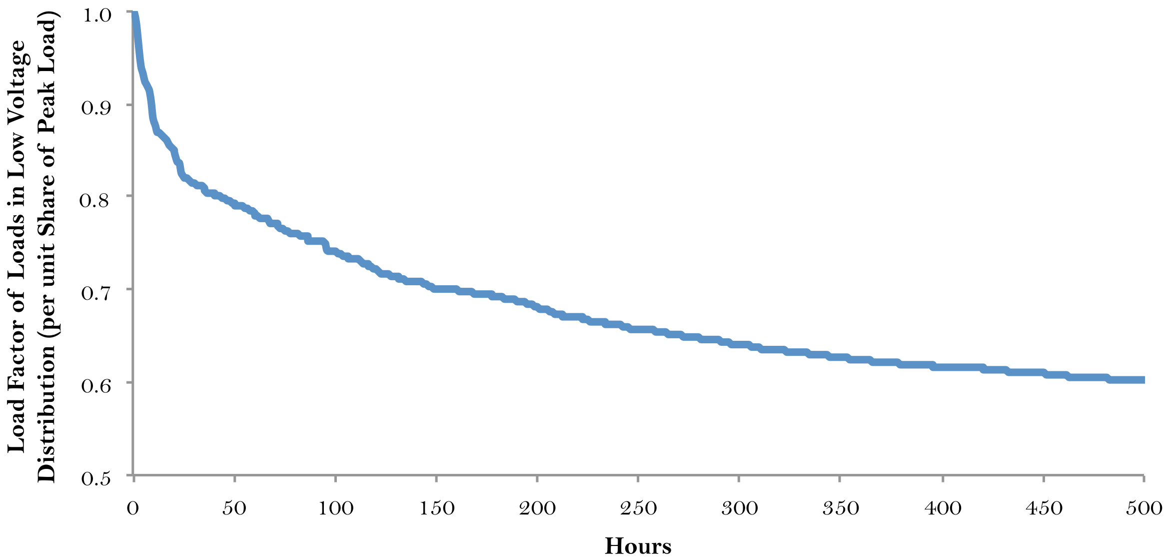
Note, however, that for DERs to provide network capacity value, they must: (1) be located in areas of the network that are experiencing or are projected to experience violations of network constraints that would otherwise necessitate investments in network expansion, such as areas of peak load growth or periods of reverse power flow;[17] (2) be able to reduce net power withdrawals or injections during periods when network congestion is expected; and (3) be able to provide these services reliably, over whatever period of time network investments are to be deferred. Not all locations and not all DERs are likely to meet these conditions. For example, a study of the network deferral value of rooftop solar PV systems connected to distribution feeders in Pacific Gas and Electric's service territory in northern California found no network capacity deferral on 90 percent of distribution feeders modeled (Cohen, Kauzmann, and Callaway 2016). The coincident peak demand in these feeders was either not growing enough to require network investments over the study period or the coincident peak did not coincide with solar PV production periods, preventing solar PV from delivering any network capacity deferral benefit. However, on 10 percent of feeders modeled, solar PV could deliver distribution network capacity deferral values ranging from $10 to $60 per kilowatt-year (kW-yr) of solar capacity deployed, while on a select 1 percent of feeders, this deferral value was $60/kW-yr or greater (Cohen, Kauzmann, and Callaway 2016).

17 Areas of the network with aging assets due for replacement could also afford opportunities for network capacity deferral value.

Not all locations and not all DERs are suited to deliver locational value by deferring network investments. For example, a study of the network deferral value of rooftop solar PV systems connected to distribution feeders in Pacific Gas and Electric's service territory in northern California found no network capacity deferral on 90 percent of distribution feeders modeled.

In addition, the marginal locational value of network deferral also declines, sometimes quite rapidly. As **Figure 8.8** illustrates, a relatively small reduction in peak net withdrawals—if optimally sited—can accommodate the first few percent of peak load growth in representative low-voltage European distribution networks. In this range of load growth, reductions in peak withdrawals have high marginal value because only a few locations in the network potentially violate network constraints. As a result, well-targeted DER investment and operation can mitigate these potential issues without requiring significant reductions in aggregate net withdrawals. However, as peak load grows more significantly, more and more areas of the network are likely to experience network constraint violations if no action is taken. Thus, more and more DER dispatch is required to accommodate increasing levels of peak load growth (see **Figure 8.9**). Eventually, load in all portions of the network is affected by binding network constraints. The marginal reduction in peak withdrawals necessary to accommodate peak load growth eventually converges to 1 megawatt (MW) of reduction per MW of load growth.

Figure 8.10: Load Duration Curve for Top 500 Hours of Demand in a Low-Voltage Distribution Network (Spanish Example)[18]



At the same time, as load grows, network constraints will become binding more frequently, requiring DER dispatch in a greater number of hours of the year as well. This dynamic can be illustrated by examining the load duration curve for a given portion of the network. For example, **Figure 8.10** depicts the 500 hours of highest load for an average low-voltage aggregate load profile in Spain. As this figure illustrates, the number of hours that exceed a given threshold load factor (defined as the hourly load as a per unit share of the annual peak load) steadily increases as the threshold falls. So while only 9 hours exceed a load factor of 0.9 (or 90 percent of highest peak demand), 36 hours exceed a load factor of 0.8 and 157 hours exceed 0.7, and so on. In other words, accommodating each marginal increment of load growth without network upgrades requires both more MW *and* more hours of reductions in peak withdrawals. This steadily reduces the marginal value of employing DERs for network capacity deferral.

Accommodating each marginal increment of load growth without network upgrades requires both more megawatts and *more hours of reductions in peak withdrawals. This steadily reduces the marginal value of employing DERs for network capacity deferral.*

18 Load profile from CNMC (2014).

Box 8.2: Expanding Distributed Generation Hosting Capacity with Active Network Management

The Flexible Plug and Play program in Cambridgeshire, England, provides an example of how actively managing peak injections can dramatically lower the network costs of accommodating distributed generation (DG). Cambridgeshire is a rural area in the east of England that has experienced a significant increase in connection requests from large DG including wind, solar, biomass, and anaerobic digestors. As the area's connected DG capacity has increased, available capacity for new generators to connect to local distribution networks without significant network reinforcement has decreased. Greater levels of DG interconnection pose new challenges and risks for network operation, including the possibility of exceeding line thermal constraints, increased levels of reverse power flows, and voltage increases. UK regulations specify that the distribution utility, in this case UK Power Networks, is responsible for network reinforcements in the high-voltage part of the distribution network that are necessary for the integration of downstream DG projects. In many instances, for such projects to proceed, UK Power Networks must be able to make a compelling business case to the UK regulator that the investment in upstream network reinforcements is warranted. This can be very challenging due to the fact that the new assets often have low utilization rates (in many cases because they will only be utilized during extreme conditions—for example, when distributed solar PV generation is at a maximum). As a result, DG applicants in this part of England have been required to go through costly and time-consuming grid upgrades before they can be connected to the network.

To address this challenge, UK Power Networks partnered with Smarter Grid Solutions to deploy an Active Network Management (ANM) system in an area of the network spanning 70 square kilometers (Smarter Grid Solutions n.d.). Through this program, UK Power Networks offers generators "managed" connections—that is, connections that allow the utility to engage in real-time control of DG in response to prevailing grid conditions. As part of the connection offer, UK Power Networks provides applicants with an estimate of how often their generation will be curtailed to manage network constraints. These estimates help developers with long-term planning and typically show a very low anticipated reduction in annual energy export volumes. In addition to managing DG output, the ANM system coordinates a range of smart devices such as Automatic Voltage Control relays, Dynamic Line Rating relays, and a

Quadrature Booster Control System[19] to manage power flows and voltage and provide real-time ratings of network components. The ANM system has significantly increased network capacity to host DG in the Cambridgeshire area without resorting to traditional investments in network capacity upgrades (Georgiopoulos n.d.). The program allowed two large solar PV arrays (1.2 MW and 6.3 MW) to connect to the distribution network at a cost that was about 90 percent lower than the cost under a business-as-usual interconnection regime. Likewise, these projects have been subject to minimal curtailment (2–2.5 percent). Similarly, six wind projects were able to connect to the distribution network under this program at costs that ranged from 75 percent to 93 percent lower than business as usual, while curtailment associated with these managed connections ranged from 1.7 percent to 5.3 percent. Finally, a combined heat and power project that was installed under this managed connections regime was able to reduce interconnection costs by 95 percent and has only been curtailed about 3.3 percent of the time. As this example illustrates, curtailing DG output in only a few hours of the year can substantially reduce the cost of connecting DG to distribution networks.

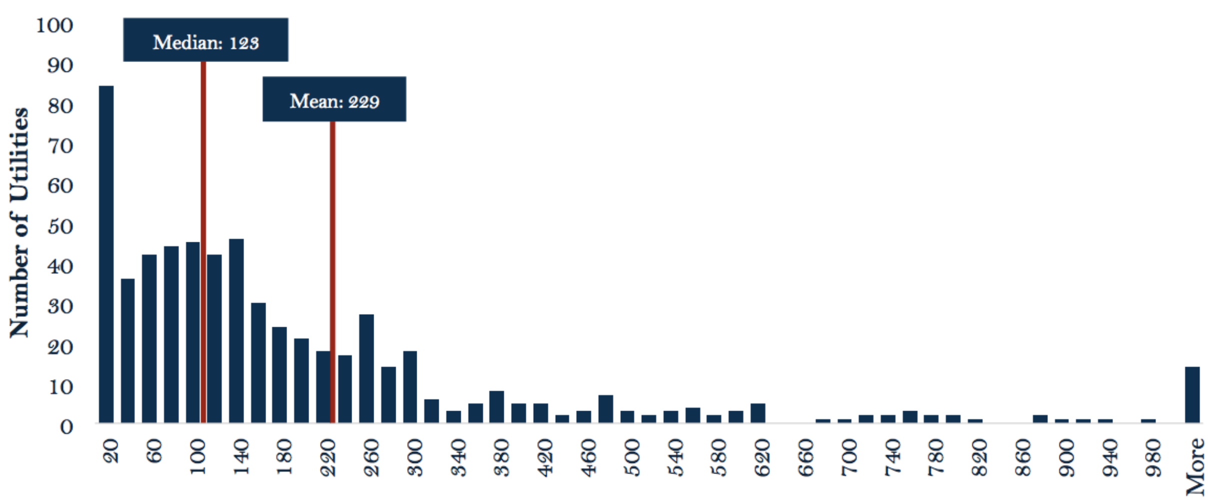
19 Also known as a phase angle regulating transformer or phase-shifting transformer, quadrature boosters are specialized transformers that control power flow along lines downstream of the transformer by shifting the phase angle at the head of one of the lines. Quadrature boosters provide a means to relieve network constraints by redirecting power flows from heavily loaded lines to more favorable paths.

8.2.1.3 Reliability

DERs may be able to increase the reliability of power systems. For example, DERs such as batteries, fuel cells, reciprocating engines, and microturbines can provide power during outages, thus avoiding service interruptions. At present, the benefits of DERs used for reliability purposes (i.e., backup power) flow almost exclusively to the DER owner (i.e., all benefits are private). However, pilot programs (such as the NY Prize or Portland General Electric Dispatchable Standby Generation program) are developing models to enable communities or wider areas to benefit from the reliability improvements provided by DERs (NYSERDA 2015; PGE n.d.). Applying the microgrid[20] concept to enable broader reliability benefits from DERs would require significant changes to system design and operational paradigms (Driesen and Katiraei 2008), such as changes to allow for differentiated quality of service for customers connected to the distribution infrastructure, duplication of circuits, careful coordination of network protection measures, islanding and reconnection protocols, etc. In many cases, these measures may be hard to justify given the low frequency of widespread outages in most locations and the existence of utility-side solutions such as network reconfiguration options or substation-level backup with diesel generators or batteries.

The locational value of increased reliability depends on the frequency and duration of service outages experienced at a particular location and consumers' willingness to pay to avoid such outages (Sullivan et al. 2015). Ratings such as the System Average Interruption Frequency Index (SAIFI) and the System Average Interruption Duration Index (SAIDI) are typically used to measure the frequency and duration of outages. (SAIFI measures the average frequency of outage events for a given network and SAIDI measures the average duration of a loss of service event for a given network.) The frequency and duration of service outages vary dramatically across distribution utilities and networks and from year to year, depending on the incidence of major storms. The median and mean outages in the

Figure 8.11: Histogram of 2014 System Average Interruption Duration Index Numbers for US Utilities



20 According to the US Department of Energy's Microgrid Exchange Group, a microgrid is a group of interconnected loads and distributed energy resources within clearly defined electrical boundaries that acts as a single controllable entity with respect to the grid. Because it can connect and disconnect from the grid, a microgrid can operate in either grid-connected or island mode.

United States lasted roughly two hours and four hours respectively in 2014 (EIA 2016), although a small number of utilities experienced average outages of longer duration (**Figure 8.11**). Meanwhile, the average frequency of outages for US utilities was roughly 1.5 outages per year per customer in 2012 (Campbell 2014). Outage frequency and duration for individual customers or for specific locations within a given distribution network can be higher or lower than these average values.

Lawrence Berkeley National Laboratory has developed estimates of willingness to pay for avoided outages for different types of utility customers (Sullivan et al. 2015). We used these estimates to calculate the cost of outages for several classes of customers and different cumulative annual outage durations (**Table 8.2**). The value of increased reliability is assumed to equal the avoided costs of these outages. To the extent that DERs can prevent service interruptions by providing backup power, they therefore generate additional locational value equivalent to the avoided costs presented below. As **Table 8.2** illustrates, the relative infrequency of average service disruptions in developed networks means that the additional locational value associated with reliability benefits may be modest, unless a given DER can provide power for multiple electricity users during each outage. Exceptions include specific locations with greater-than-average supply disruptions and/or certain high-value or critical loads—facilities such as hospitals, emergency response services, or uninterruptible industrial processes—that exhibit a very high value for avoided supply interruptions (or equivalently, a very high cost of outages). Due to this higher value, many such facilities are already equipped with DERs for uninterruptible power supply.

Table 8.2: Estimated average reliability value for different customer classes in the United States and average cumulative annual outage durations[21]

CUMULATIVE ANNUAL OUTAGE DURATION	RESIDENTIAL (PER CUSTOMER PER YEAR)	SMALL COMMERCIAL (PER CUSTOMER PER YEAR)	LARGE COMMERCIAL & INDUSTRIAL (PER CUSTOMER PER YEAR)
2 hours (median)	$7	$1,100	$30,000
4 hours (mean)	$14	$2,200	$60,000
16 hours (extreme)	$56	$8,800	$240,000

The locational value of increased reliability depends on the frequency and duration of service outages experienced at a particular location and consumers' willingness to pay to avoid electricity outages. In general, this value is modest, unless a given DER can provide power for multiple electricity users during each outage and/or the DER is used to assure uninterrupted supply to very high-value loads, such as hospitals, emergency response services, or other critical loads.

21 Annual average electricity demand is 13,351 kWh for the residential customer class, 19,214 kWh for the small commercial customer class, and 7,140,501 kWh for the large commercial and industrial customer class (Sullivan et al. 2009). The cost of service interruptions is estimated assuming typical outage durations of 2 hours (so the cost of a cumulative annual outage duration of 16 hours equals the cost of a 2-hour interruption multiplied by 8). Note that Sullivan et al. find shorter-duration outages to be more costly on a per-hour basis—that is, a 16-hour outage would be less costly for a customer than eight 2-hour outages.

8.2.1.4 Other sources of locational value

DERs may deliver additional electricity services and public benefits that exhibit locational value, including contributing to conservation voltage reduction programs that reduce final energy consumption, avoiding land-use impacts associated with alternative resources, or even providing local employment benefits.

Modern appliances and machinery are designed to operate within 5–10 percent of a nominal voltage (e.g., 120 volts for household appliances in the United States and 230 volts in Europe). Power electronics, including AC–DC inverters commonly installed along with solar PV systems, electrochemical batteries, and electric vehicle charging stations, can be used to decrease total power consumption by reducing the voltage at which loads are consumed to the low end of the voltage tolerance range (Schnieder et al. 2010; Wang and Wang 2014). This practice, termed conservation voltage reduction (CVR), has been demonstrated to reduce total power consumption by 0.5–4 percent using conventional network infrastructure, such as tap changers at distribution substations (Schnieder et al. 2010). The bulk of this reduction (80–90 percent of energy savings) occurs on the customer side of the meter with additional savings due to reduced line losses (Schnieder et al. 2010). Using DERs to perform CVR close to electricity end uses offers more fine-grained control of consumer voltage and can reduce power consumption by 3.5–5.2 percent (Wang and Wang 2014). In networks with existing CVR programs, DERs may therefore deliver modest incremental savings, while their benefit, in terms of additional annual energy savings, may increase to as much as 5.2 percent in networks without existing CVR programs. In addition, by reducing consumption during periods of peak demand or network stress, CVR can also contribute to meeting network and generation capacity needs. CVR can thus be viewed as a method for consuming both energy and capacity services more efficiently. However, the efficacy of CVR is highly dependent on load type and reaction to voltage drops. In many cases, lowering the voltage at which a device consumes electricity either reduces the quality of service that the device provides or results in no reduction in power as the device draws more current to compensate for the lower voltage. For example, resistive loads such as incandescent light bulbs will dim and resistive heaters will work less effectively at lower voltages. Constant power loads such as computers, LED lights, or other devices with AC–DC conversion devices will increase current to compensate for lower voltage and maintain constant power. As the prevalence of constant power loads increases, the efficacy of CVR may therefore decline. It is also possible that utility adoption of more granular voltage control mechanisms, like switched capacitor banks on the secondary side of distribution transformers, would reduce the additional benefits achieved by utilizing customer-sited inverters for CVR.

Distributed power electronics, including inverters associated with solar PV, electrochemical energy storage, and electric vehicle charging, can contribute to conservation voltage reduction (CVR) programs, reducing energy consumption. However, the efficacy of CVR is highly dependent on load type and reaction to voltage drops. CVR efficacy may decline as devices with constant power loads become more prevalent.

Co-locating generation on customer premises or placing generation on previously impacted sites, such as industrial brownfields or landfills, can mitigate the land-use impacts of power generation. Where DERs serve as alternatives to other generation resources with greater land-use impacts, this can be considered an additional locational value. Furthermore, when DERs are used to defer the expansion of transmission and distribution infrastructure they can generate land-use benefits by reducing the need to acquire additional rights-of-way. If land values are high, which is the case in many urban areas, or if impacts to sensitive environments can be avoided, the resulting locational value can be significant. In addition, as renewable energy penetration increases, locations suitable for large-scale renewable energy development may become scarce or more difficult to access, increasing the potential benefit of distributed resources. Of course, these benefits are not universal, and not all DERs avoid or reduce generation or network capacity and associated land-use impacts.

Finally, deployment of DERs is commonly justified with appeal to local employment or job creation benefits. These employment benefits may be considered locational as well, since the jobs associated with deploying DERs (i.e., construction, installation, etc.) would be created within the local region where the DERs are installed. Analyzing the employment impacts of distributed resources is extremely challenging, however, as it is difficult to incorporate all appropriate short- and long-term equilibrium impacts on overall employment. For example, jobs benefits associated with DERs are often simply a welfare transfer or loss, as the opportunity costs of paying more for DERs translate to employment losses in other sectors of the economy that now see lower expenditures (NYSERDA 2011). A full and accurate accounting would consider the general equilibrium implications of energy investment decisions on overall employment and welfare. In other words, just because a particular technology is more labor-intensive to install doesn't mean it is necessarily a net job creator in the economy as a whole (Tsuchida et al. 2015). Thus, the distributional benefits or costs of DER deployment for a particular locality are difficult to judge.

A CASE STUDY: LOCATIONAL VALUE OF DISTRIBUTED SOLAR PV IN NEW YORK STATE

To provide a concrete example of locational value in different settings, we present a case study of distributed solar PV in the state of New York. We compare the locational value of services provided by solar PV installations within two zones of the independent system operator for New York (NYISO):[22] Long Island and Mohawk Valley in central New York. Long Island experiences the highest average LMPs in the NYISO system due to frequent congestion and relatively high transmission losses, while Mohawk Valley has relatively average LMPs and infrequent congestion. We therefore employ Long Island as a "high-value case," using assumptions grounded in empirical data to construct a case where locational values are on the higher end of the range of possible values exhibited in the NYISO system (**Figure 8.12**). Mohawk Valley, by contrast, represents an "average-value case" with values more typical of most locations in the NYISO system (**Figure 8.13**). Note that in both cases, we assume the solar systems are price takers and do not affect the marginal value of each service. If solar PV penetration increases, the marginal value of solar PV services may decline (as discussed in **Section 8.2.1**). In addition, quantitative results are case-specific and are meant to be illustrative rather than generalizable to other power system contexts. We calculate the locational values for each zone as follows:

- A distributed PV system at a specific location in the power system has an incremental locational energy value equal to the difference between the LMP at the PV system's location and the price at the location of the marginal generator in the NYISO system. We therefore calculate the production-weighted average locational energy value for the two zones using historical 2015 hourly day-ahead NYISO market prices and rooftop solar PV production profiles for Islip (Long Island) and Monticello (Mohawk Valley), New York. This yields an incremental value of $24.0/MWh for Long Island and $2.3/MWh for Mohawk Valley, relative to the energy value at the system-wide marginal generator.
- In addition, distributed PV systems can reduce losses in distribution networks. This component of their locational energy value is not captured in the wholesale energy market price, so we estimate the marginal value of distribution network losses accounting for hourly PV production and aggregate load profiles in each location. We assume the distribution system in the Long Island location has high losses, averaging 9 percent annually, while the Mohawk Valley location exhibits average distribution losses of 5 percent. This yields a production-weighted average locational energy value due to reduced distribution losses of $5.6/MWh in Long Island and $3.1 in Mohawk Valley. Note that these values do not reflect metered losses in the two locations—rather they are chosen simply to create

22 NYISO only publicly reports zonal (as opposed to nodal) electricity prices.

high value and average value cases for illustrative purposes. Actual distribution losses and associated locational energy values in Long Island and Mohawk Valley may differ.

- We assume the inverters associated with the solar PV system can be used to perform CVR, reducing the voltage at which loads co-located with the PV system are supplied. These benefits are calculated by summing the energy reduction benefits associated with operating at lower voltages and the capacity benefits associated with consuming less power during peak periods. To construct average and high value examples, we assume a 1 percent total energy consumption reduction for Mohawk Valley and a 5.2 percent reduction for Long Island, reflecting the typical and extreme values reported in Wang and Wang (2014) for technologies that are capable of directly managing consumption voltage. Accounting for energy and capacity prices in each zone yields locational values for CVR of $11.1/MWh in Long Island and $1.7/MWh in Mohawk Valley.

- Similarly, employing the highest and most typical distribution network investment deferral values from Cohen, Kauzzman, and Callaway (2016), we assume $60/kW-yr of investment deferral value for the Long Island case and $0/kW-yr for the Mohawk Valley case. In other words, we assume that PV output in the Long Island case is coincident with the aggregate peak in local network usage and that the network would otherwise require near-term upgrades, whereas the Mohawk Valley case assumes PV output is noncoincident with local aggregate peak network withdrawals or the network does not require upgrades. Accounting for annual PV production, this is equivalent to a levelized value of $41.2/MWh and $0/MWh for the two cases, respectively. These values are illustrative, and the actual network deferral value of specific PV systems located in Long Island or Mohawk Valley may differ.

- We also account for the fact that the firm generation capacity value of distributed solar PV is enhanced by avoided distribution losses. That is, since generation capacity values are measured at the bulk transmission level, 1 MW of distributed PV avoids 1+*l* MW of capacity installed at bulk transmission voltages, where *l* equals marginal distribution losses during peak-coincident hours. We therefore calculate the additional capacity value of distributed solar by multiplying the solar capacity value[23] by the average distribution loss factor during each of the seasonal peak demand hours used by NYISO to calculate capacity values.[24] Accounting again for annual PV production, this yields a levelized value of $2.9/MWh for Long Island and $0.9/MWh for Mohawk Valley.

- Finally, because engineering safety standards require grid-connected solar PV systems to disconnect and power down in the case of network failures, we assume no reliability value for these distributed PV systems.[25]

As this case study illustrates, the total locational value of the services provided by distributed solar PV—and other DERs—can vary by an order of magnitude within the same power system. The total locational value for the high-value Long Island case is $84.7/MWh (**Figure 8.12**), or roughly three times the average wholesale energy price. This value is an extreme case, representing the highest range of reported values for network capacity deferral and conservation voltage reduction benefits, for example, and the most congested zone in the NYISO system. Meanwhile, the total locational value for the more typical Mohawk Valley case is $7.9/MWh (**Figure 8.13**)—a full order of magnitude less. While this is still a non-negligible incremental value of 29 percent over the average wholesale energy price, it is much more modest.

This dramatic difference in locational value reinforces the importance of a system of prices and charges for electricity services that has sufficient locational granularity. In addition, variation in locational value makes it clear that there is little sense in attempting to define a single "value of solar" or "value of storage" or value of any other resource. The value of each DER depends on the value of the specific services it provides at a specific time and in a specific location.

The total locational value of the services provided by distributed solar PV—and other DERs—can vary by an order of magnitude within the same power system. This dramatic difference in locational value reinforces the importance of a system of prices and charges

23 NYISO calculates the capacity value of solar resources by averaging the production of the resource over the previous peak hours. To account for changes in capacity values over time, we average the capacity values from three previous years (2013, 2014, and 2015).

24 NYISO defines "Summer Peak Hours" as the hours beginning 14, 15, 16, and 17 during the three-month period from June 1 through August 31, inclusive. "Winter Peak Hours" are defined as the hours beginning 16, 17, 18, and 19 during the three-month period from December 1 through the last day of the immediately following February. The capacity value of solar PV is defined as the average capacity factor during these peak hours. See NYISO 2016, "Installed Capacity Manual," Version 6.32, New York Independent System Operator, February 2016, Section 4.5.1.

25 If PV systems are paired with energy storage and configured to safely island during outages, their reliability value could be higher. In that case, however, much of the reliability value could be appropriately attributed to the storage system, which could charge from grid-connected power and provide backup supply without necessarily being co-located with the PV system. The incremental reliability value of the solar system in this case would be equivalent to the extended duration of self-sufficient supply due to the ability to recharge the storage system during islanded operation.

Figure 8.12: Locational Value of Distributed Solar PV — Long Island, New York (High-Value Example)

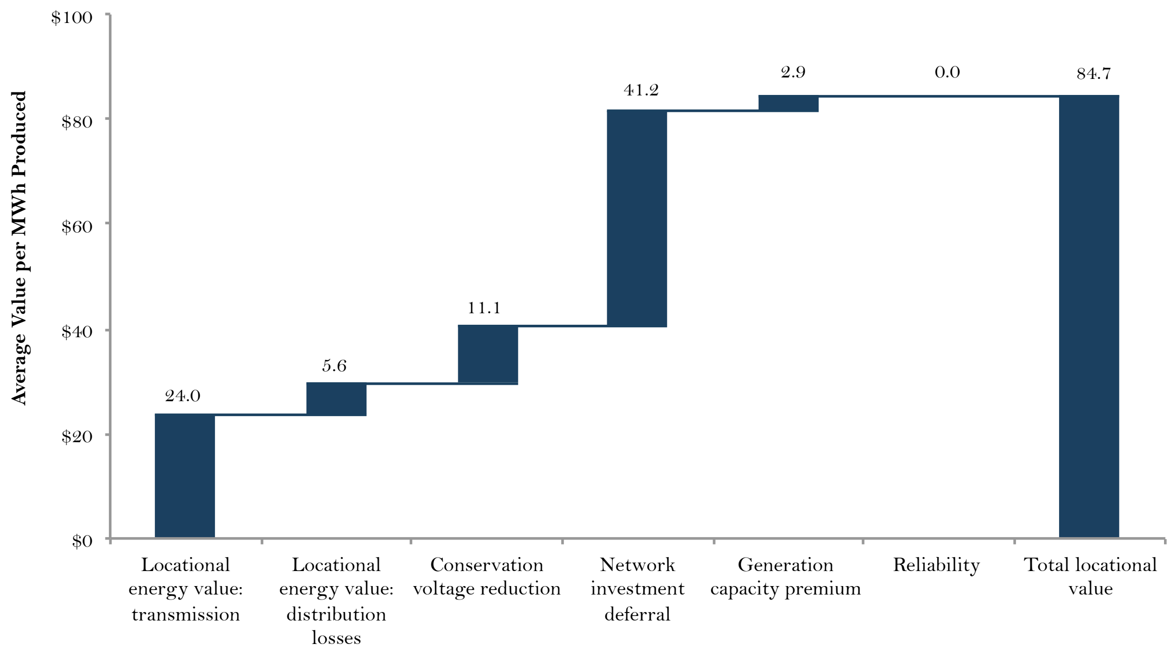
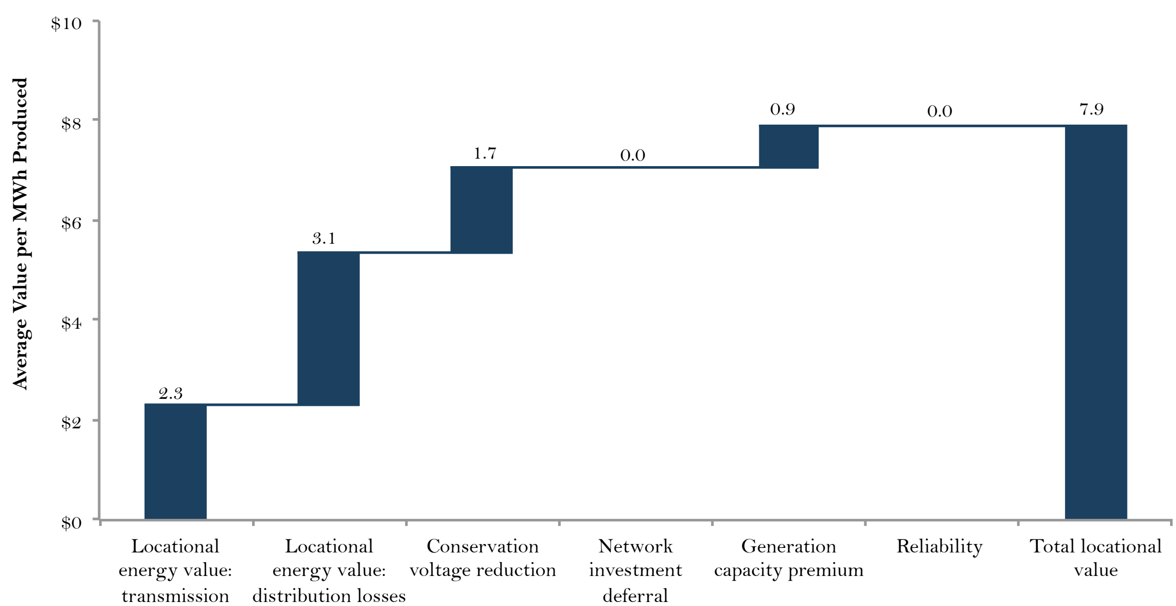


Figure 8.13: Locational Value of Distributed Solar PV — Mohawk Valley, New York (Average-Value Example)



for electricity services that has sufficient locational granularity. It also means that efforts to define a single value for any DER are impractical.

Taking better advantage of existing resources has significant potential to reduce the costs of power systems—if these resources can be unlocked with appropriate prices and incentives for flexible operation.

8.3 Unlocking the Potential Value of Existing Resources

Currently, many resources that already exist in the power system are greatly underutilized. Some of these resources are located on the customer's premises to satisfy specific customer needs: mobility, heating, back-up power, etc. However, the operation of these resources can be optimized to reduce their impact on the power system and even these resources can become cost-effective providers of electricity services.

For instance, transforming greater numbers of electricity users from passive consumers to flexible and price-responsive demand can unlock an inherently distributed resource capable of providing services at the specific times and locations that contribute the greatest value to power systems. In addition, many distributed generation resources, including solar PV systems (and associated inverters) and backup generators, have already been deployed to deliver private benefits to DG owners. Better dispatch and control of these resources and integration into active distribution system operations may yield cost savings and contribute valuable electricity services for which their owners can be compensated. Electric vehicle (EV) penetration is expected to increase, creating another important class of electricity demand that can be charged in a controlled or price-responsive manner to reduce the impact of EV charging on power systems or provide electricity services.

In short, taking better advantage of existing resources has significant potential to reduce the costs of power systems—if these resources can be unlocked with appropriate prices and incentives for flexible operation.

Unlocking the potential of these existing resources can often be more cost-effective than deploying new DERs, conventional generators, or transmission or distribution network assets. Coordinating, dispatching, or incentivizing existing resources to more efficiently consume or supply electricity services may require investment in communications, controls, and metering infrastructure, but in many cases these investments can be quite modest, especially compared to new capital expenditures. In addition to these initial "activation" expenditures, operation of these resources entails opportunity costs of differing magnitudes, depending on the value of the services that the resource owner foregoes to provide system benefits (e.g., reduced occupant comfort, delayed EV charging, etc.). These opportunity costs can be very low, as in the case of flexible heating, ventilation, and cooling (HVAC) systems or reactive power provision with smart inverters, or these opportunity costs can be relatively high, as in the case of demand curtailment. Finally, activating existing resources to provide electricity services entails transaction costs and may incur other regulatory implementation costs that should be fully considered.

Understanding the value of existing DERs requires careful consideration of (1) the temporal and locational value of the services they may deliver, (2) the opportunity costs and transaction costs associated with operating these resources to provide electricity services, and (3) the initial activation costs required to engage these assets in the efficient provision or consumption of electricity services.

Understanding the potential costs and benefits of utilizing existing DERs therefore requires careful consideration of (1) the temporal and locational value of the services the DERs may deliver, (2) the opportunity costs and transaction costs associated with operating these resources to provide electricity services, and (3) the initial activation costs required to engage these assets in the efficient provision or consumption of electricity services. In this section, we present several case studies that illustrate the potential value of existing resources, including flexible demand, controlled EV charging, and smart inverters. Note that these case studies are meant to be illustrative; their quantitative findings should not be generalized to other contexts.

8.3.1 Flexible heating and cooling

Buildings collectively consume nearly three-quarters of total electricity in the United States, with HVAC systems accounting for the single largest category of energy consumption within buildings (EIA 2015; DOE 2012). HVAC loads collectively represent a substantial opportunity for more flexible consumption of electricity and a potential supplier of valuable electricity services, including network capacity deferral or even ancillary services including frequency regulation (Blum and Norford 2014; Kim et al. 2015; Xu 2015). Due to the thermal mass of building materials and the air contained in the building envelope, buildings exhibit thermal inertia—that is, changes in electricity consumption for heating and cooling have a lagged effect on the internal temperature of building spaces and thus on occupants' comfort. This thermal inertia makes HVAC loads well suited for flexible operation in response to energy prices or control signals. Building energy management systems or "smart thermostats" that employ methods such as model predictive control (MPC) can enable HVAC systems to follow an optimal heating or cooling strategy that exploits thermal inertia to minimize energy costs and/or provide electricity services while maintaining a comfortable environment for building occupants.

As an example of the potential benefits of flexible HVAC loads, we simulate a residential building in Westchester County, New York,[26] with electric cooling using the Demand Response and Distributed Resources Economics Model (see **Appendix A**). The building's air conditioning (AC) demand is controlled by a home energy management system or "smart thermostat" that uses optimization to schedule the AC unit's operation, based on an expectation of future hourly energy prices and ambient temperatures, to minimize the customer's electricity bill subject to occupants' comfort preferences. We model thermal flexibility using a simplified building energy thermal model derived from Mathieu et al. (2012), with temperatures varying around a target set point (e.g., 70°F / 20.5°C). Penalty factors within the model's objective function account for the comfort impact of deviations away from the target temperature set point. These factors represent the opportunity cost of shifting cooling demand.[27] When confronted with hourly prices that reflect wholesale LMPs for energy supply[28] and marginal distribution losses,[29] and peak-coincident charges that reflect the household's contribution to generation capacity and distribution network capacity requirements, flexible control of the building's AC system shifts the household's electricity demand away from the highest priced hours. We model two cases, one reflecting a high load-growth scenario with higher peak-coincident generation capacity and network capacity charges and a low load-growth scenario where these charges are more

26 The household consumes 6,600 kWh annually and is equipped with a 4 kW central air conditioning system.

27 We assume a temperature deviation of 0.5°C with no penalty to reflect the fact that HVAC equipment naturally fluctuates around a target temperature. Deviations from 0.5°C to 1.5°C are penalized at $0.10 per hour, while penalties increase to $0.50 per hour for deviations of 1.5°C–2.5°C and rise steeply for each additional degree of deviation beyond 2.5°C. These penalties are designed to reflect the fact that residential consumers have a demonstrated price elasticity for air conditioning and have shown a willingness to tolerate both pre-cooling and higher temperatures during event hours (EIA 2014; Herter and Okuneva 2013).

28 We use historical 2015 hourly wholesale energy prices for NYISO Zone I, which contains Westchester County, New York.

29 We assume average losses in a distribution network are 5 percent and use a quadratic function to estimate marginal losses in each hour as a function of historical 2015 NYISO Zone I aggregate load (i.e., Westchester County, New York).

modest.[30] Note that these prices differ significantly from current tariff practices in New York and are meant to reflect the response of users to cost-reflective tariffs (see **Chapter 4** for additional discussion).

Figure 8.14 illustrates an example of hourly prices, household electricity demand, and indoor and outdoor temperatures with and without flexible HVAC during four days in July for the low load-growth case. Flexible AC operation results in significant reductions in annual energy costs. Employing flexible AC control reduces the household's annual electricity bill by $108 or 6 percent under the low load-growth case and $407 or 22 percent under the high load-growth case. In addition, by reducing total electricity consumption during the few hours of the year that account for the greatest share of electricity costs, flexible HVAC loads and other flexible demand can substantially reduce total power system costs.[31] A simple calculation provides a rough sense of the potential aggregate impact of these measures: Assuming the tariffs in our simulation accurately reflect the marginal costs of serving this class of customer and assuming flexible HVAC devices are adopted by even one-quarter of New York State's 8.2 million households, total electricity system costs in the state could be reduced by hundreds of millions of dollars annually.[32]

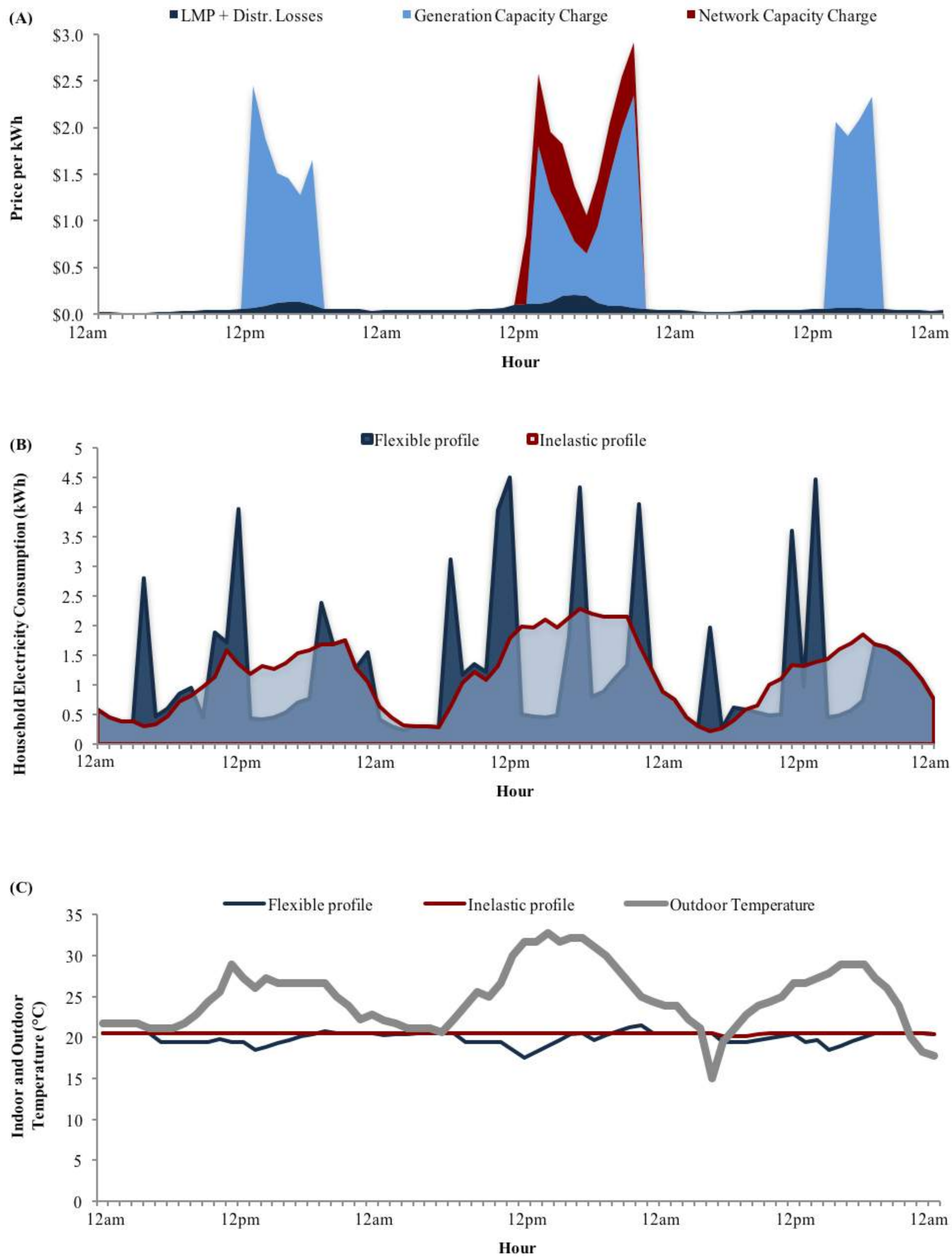
Commercial customers may have even greater potential to harness thermal inertia and become more flexible consumers of energy for heating and cooling. Commercial buildings account for nearly half of total building electricity consumption in the United States (EIA 2015), and with higher total energy consumption per building, unlocking flexible HVAC control in commercial buildings may entail lower transaction and activation costs than residential homes. In addition, with the right controls, commercial buildings with variable speed heat pumps are able to provide ancillary services, such as providing operating reserves, to power system operators. Indeed, Kim et al. (2016) estimate that as much as 59 gigawatts of reserve capacity could be provided by the 6.6 billion square-meters of floor area in the current commercial building stock in the United States. For more on the potential for flexible HVAC in commercial buildings to provide electricity services, see Xu (2015).

30 Generation capacity costs are assumed to be $70/kW-yr in the low growth case, reflecting 2013–2015 average NYISO capacity clearing prices, and $115/kW-yr in the high growth case, reflecting NYISO load and capacity forecasts for 2016–2026 (NYISO 2016). Distribution network capacity costs range from $25/kW-yr in the low growth case to $120/kW-yr in the high growth case, reflecting two values within the range of transmission and distribution deferral benefits estimated in several regulatory avoided cost filings (see, e.g., Neme and Sedano 2012). Costs are converted to peak-coincident charges by allocating annual costs proportionately to the top 100 hours of aggregate demand at the NYISO zonal level for generation capacity charges and top 100 hours of average residential load profile for the distribution network charge.

31 Note that generation and network coincident peaks occur during summer hours in this New York case example. In systems with winter-peaking demand at either the system-wide or distribution feeder level, households or businesses with electric heating systems could employ similar flexible control strategies to reduce energy bills and total power system costs.

32 Note that this case assumes the HVAC controller has perfect information about future prices. In practice, HVAC systems will be controlled with imperfect foresight, which may result in lower savings. Algorithms for optimal control of HVAC systems with imperfect information can perform quite well, but never as effectively as with perfect information.

Figure 8.14: Results for Westchester County, New York, Example (Low Load-Growth Case) with Respect to Hourly Energy Prices (A), Household Electricity Consumption (B), and Indoor and Outdoor Temperature (C) over Four Days in July

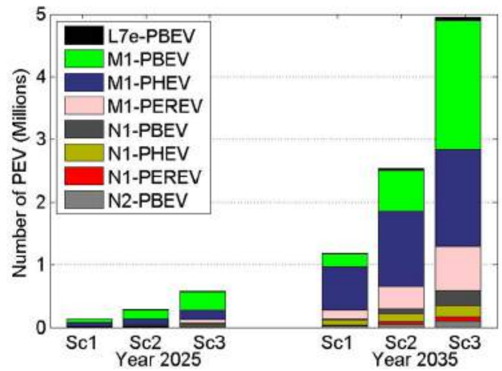


8.3.2 Smart charging of electric vehicle fleets

Electric vehicle (EV) adoption is expected to steadily increase over the next decade. Projections vary, but one source estimates that EV market share may reach approximately 20 percent of new vehicle sales worldwide by 2030 (BNEF 2016). Such estimates suggest that EV charging could account for a significant increase in electricity demand. Enabling smart charging of EVs is therefore key to reducing the impact of this new class of demand on power system costs.

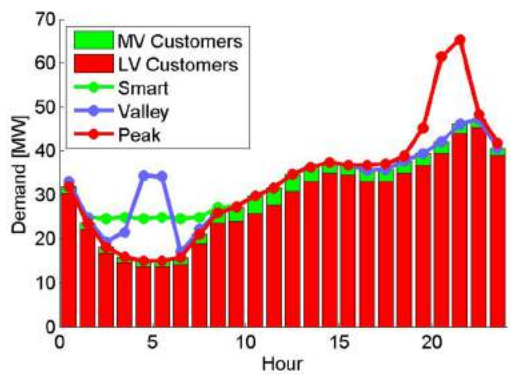
In particular, different EV charging strategies can have a substantial effect on the cost of distribution networks as well as on the marginal cost of energy and generating capacity. Mateo, Frías, and Miralles (2016) use the Reference Network Model (RNM) to analyze the effect of charging strategies on five large-scale distribution networks (see **Appendix A** for a description of the RNM). They consider three scenarios for increasing EV penetration in Spain based on policy scenarios developed by Hassett and Bower (2011) (**Figure 8.15**).[33] The authors consider four categories of vehicle size: quatricycle (L7e), passenger vehicle (M1), goods-carrying vehicles with a maximum laden mass of 3,500 kilograms (N1), and goods-carrying vehicles with a maximum laden mass of 12,000 kilograms (N2). The authors also distinguish between three specific types of EV technology: plug-in battery electric vehicles (PBEVs), which have no energy source besides the battery; plug-in hybrid electric vehicles (PHEVs), which use a combustion engine after their batteries are depleted; and plug-in extended-range electric vehicles (PEREVs), which use a combustion engine to provide electrical power and overcome range limitations. Additional assumptions about specific connection times, energy consumption for each type of electric vehicle, and distance traveled by each type of vehicle are based on Hassett and Bower (2011) and Ball, Keers, and Alexander (2010).[34]

Figure 8.15: EV Penetration Scenarios for Spain



Source: Mateo, Frías, and Miralles (2016)

Figure 8.16: Illustration of EV Charging Strategies on Top of Energy Consumed by Medium- (MV) and Low-Voltage (LV) Consumers



Source: Mateo, Frías, and Miralles (2016)

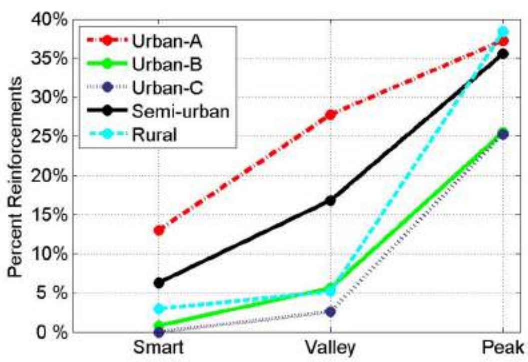
EVs are modeled under three different charging strategies (smart, valley, and peak). **Figure 8.16** illustrates EV charging demand in each hour as the difference between the dots and bars, where the bars reflect other non-EV electricity demand in each hour. The "smart charging" strategy is responsive to price signals or dispatch signals that reflect local distribution network constraints and is designed to minimize impact on distribution network costs. The "valley charging" strategy charges vehicles on

33 EVs are assigned to each of the five representative test networks in proportion to the aggregate load in each network.

34 See Mateo, Frías, and Miralles (2016) for a detailed description of technical parameters for different types and sizes of EVs.

a schedule that considers only wholesale energy prices, without accounting for distribution network constraints. Finally, in the “peak strategy,” EV owners start charging their vehicles whenever they arrive home without regard to marginal energy prices or distribution network constraints.

Figure 8.17: Distribution Network Reinforcement Costs (as Percent of Initial Network Cost) in Five Representative European Distribution Networks—2035 EV Penetration, Scenario 3



Source: Mateo, Frías, and Miralles (2016)

At low EV penetration levels (e.g., those predicted for 2025), EVs can be charged with very low (if any) network reinforcement costs. However, under high penetration levels, the importance of unlocking the flexibility of EV loads becomes apparent. **Figure 8.17** depicts the increase in distribution network costs associated with reinforcements to accommodate EV charging for the year 2035 under Scenario 3 (reinforcement costs are presented as a percent of initial network costs). A peak charging strategy wherein EV owners plug in their vehicles whenever they arrive home would entail roughly 25–40 percent increases in distribution network costs (as well as additional impacts on generation and transmission network costs) if EVs reach 5 million vehicles or roughly 20 percent of Spain’s current light-duty vehicle fleet. A valley charging strategy that is responsive to wholesale energy prices reduces the impact on distribution networks somewhat. Yet EV charging may still drive network reinforcements under this strategy on circuits that experience local coincident demand peaks that do not align with system-wide demand peaks. By contrast, a smart charging strategy that reflects distribution network constraints achieves the greatest reduction in distribution network reinforcement costs, containing reinforcement costs to less than 14 percent of existing network costs. As this example illustrates, prices or charges for electricity services that reflect the cost of distribution network constraints or reinforcements can be important to avoid unnecessary network reinforcements. These signals are especially relevant in scenarios that assume high penetration of EVs, as EV owners or aggregators can respond to prices and often exhibit considerable flexibility in both the time and location of EV charging. However, smart charging strategies may require higher investments in control and communications technologies. These activation costs must be compared to potential benefits in terms of avoided network or generation costs.

8.3.3 Reactive smart inverters with smart inverters

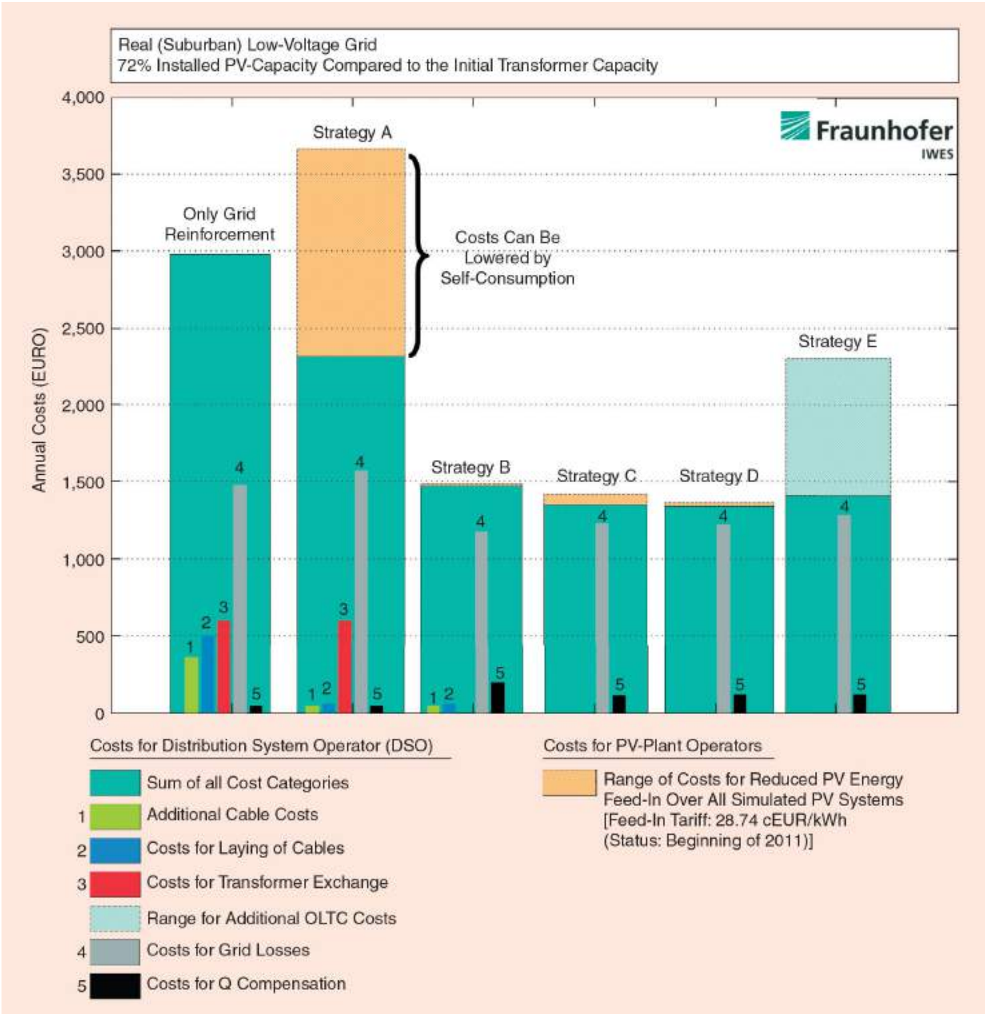
Inverters associated with DERs that generate direct current power, such as solar PV, battery storage, and fuel cells, can be employed to provide voltage control and reactive power compensation in distribution networks. Inverters can control real and reactive power output locally, in response to measured physical conditions; via communication with an upstream substation or some other aggregation of inverters in a region of the distribution network; or via communication with the distribution system operator or utility (von Appen et al. 2013). The value of utilizing smart inverters depends on the benefits offered by offsetting network reinforcements (such as investments in transformers or lines) relative to the costs of integrating smart inverters that are capable of responding to centralized or distributed controls. Enabling access to the capabilities of smart inverters entails both operational and investment costs.

Typically, the apparent power rating of inverters is sized to accommodate at least the peak production of a PV unit. PV production at this full rated capacity typically spans only a small number of hours each year. During these hours of peak PV output, the inverter may be utilizing all of its rated capacity for real power provision. In order to absorb reactive power during these hours, the inverter would have to curtail real power output. This would yield an opportunity cost for real power output equal to the cost of having to purchase power from the grid (instead of self-generating at virtually zero marginal cost). Alternatively, if real power production exceeds load, the opportunity cost is the lost remuneration for curtailed real power output that would otherwise have been injected into the network. During most hours of the year, however, inverter units have excess capacity to inject or absorb reactive power without curtailing real power, and thus, at little-to-no opportunity cost.

The primary "activation" cost for smart inverters is the cost of enabling existing inverters to respond to some local, distributed, or central control signals or price signals. According to some industry estimates, the cost premium for a smart PV inverter relative to a conventional inverter amounts to 1 percent of the total cost of a distributed PV installation (Niggli 2013). Efficient use of smart inverters implies additional costs and considerations, however, including the need to ensure adequate availability of sensing and communication capabilities. For example, in cases where the distribution system operator (DSO) exercises centralized control over inverters, the primary information and communication technology (ICT) requirements involve sensors to measure system state and communication capabilities to exchange information between DERs or local sensors and the distribution utility. In cases of localized control over inverters (meaning that inverters respond to measurements of local conditions without control or dispatch instructions from the DSO), communication requirements may be less, but limited coordination of resources across the distribution system may inhibit use of the lowest-cost combination of DERs.

Figure 8.18 depicts a cost comparison between multiple voltage control strategies from von Appen et al. (2013). It illustrates some of the benefits of utilizing inverters for voltage control as alternatives to conventional distribution system investments. Strategy A restricts the output of PV units to 70 percent of installed PV capacity and shows, in orange, the opportunity cost of curtailed PV output. This leads to higher total costs than costs incurred for grid reinforcement. Strategies B, C, and D rely on automatic, local control of active and reactive power by inverters. Utilizing inverters for voltage control under these strategies reduces total costs (including operational and investment costs) by roughly 50 percent relative to the costs of network reinforcements. Finally, Strategy E relies on an on-load tap-changing (OLTC) transformer, which can be used in a decentralized control scheme that requires coordination between inverters in the low-voltage distribution network and an upstream MV/LV transformer.

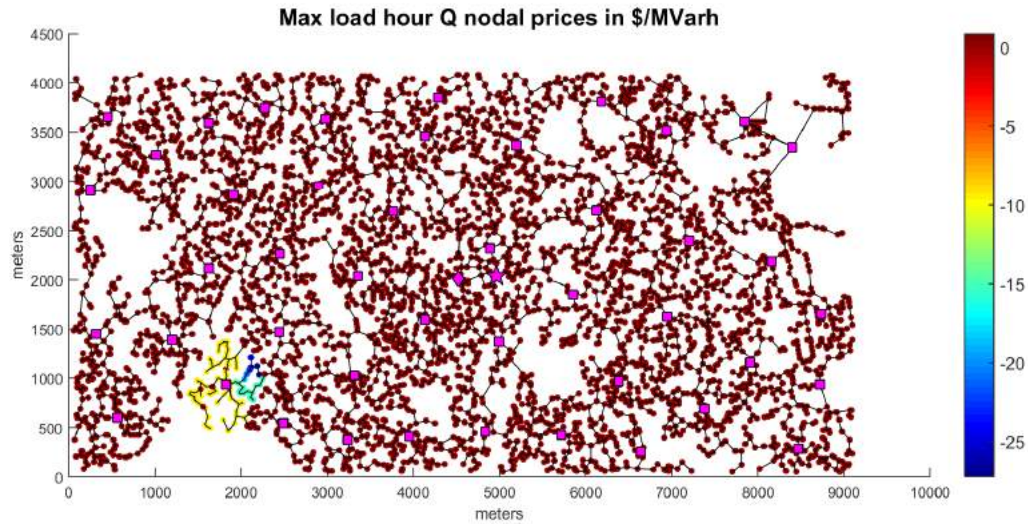
Figure 8.18: Cost Comparison for Multiple Voltage Control Strategies



Source: von Appen et al. (2013)

The value of utilizing smart inverters is, again, dependent on where in the network distributed resources are located. For example, **Figure 8.19** illustrates the locational marginal price of reactive power for one hour in an urban distribution network with PV capacity equal to 10 percent of peak load. Reactive power absorption is particularly valuable in one region of the network with greater PV output than load. In contrast, because PV output is sufficiently matched to load elsewhere in the network, the presence of smart inverters yields no value for reactive power absorption at these locations.

Figure 8.19: Localized Value of Reactive Power Absorption in a Distribution Network with 10 Percent PV Penetration



8.4 Economies of Scale and Distributed Opportunity Costs

Many technologies suitable for distributed deployment, including solar PV, electrochemical energy storage, and fuel cells, can be deployed across a range of scales. Each of these resources also exhibits varying degrees of economies of unit scale, or declining costs per unit of capacity as the size of the system increases.

For example, solar PV systems can be deployed at the kilowatt (kW) scale on residential rooftops, at a scale of several hundred kW to several MW on commercial or industrial rooftops or in ground-mounted arrays, or at a scale of tens to hundreds of MW in so-called "utility-scale" solar farms. **Figure 8.20** depicts the estimated annual cost of ownership for PV systems of varying scales in the United States, including annuitized capital costs[35] and fixed operations and maintenance (O&M) costs. Estimates are given for costs in 2015[36] and for a range of possible cost declines through 2025.[37] As **Figure 8.20** illustrates, smaller-scale solar systems have higher annual costs than larger systems due to economies of unit scale in installation, system component costs, and maintenance.[38] Failure to exhaust economies of unit scale for solar PV systems therefore results in incremental unit costs relative to larger-scale systems.

Resources that can be deployed at multiple scales, such as solar PV and energy storage, exhibit economies of unit scale—that is, declining costs per unit of capacity as the size of the system increases. By failing to exhaust economies of unit scale, smaller-scale deployment of these resources therefore results in incremental unit costs relative to larger-scale systems.

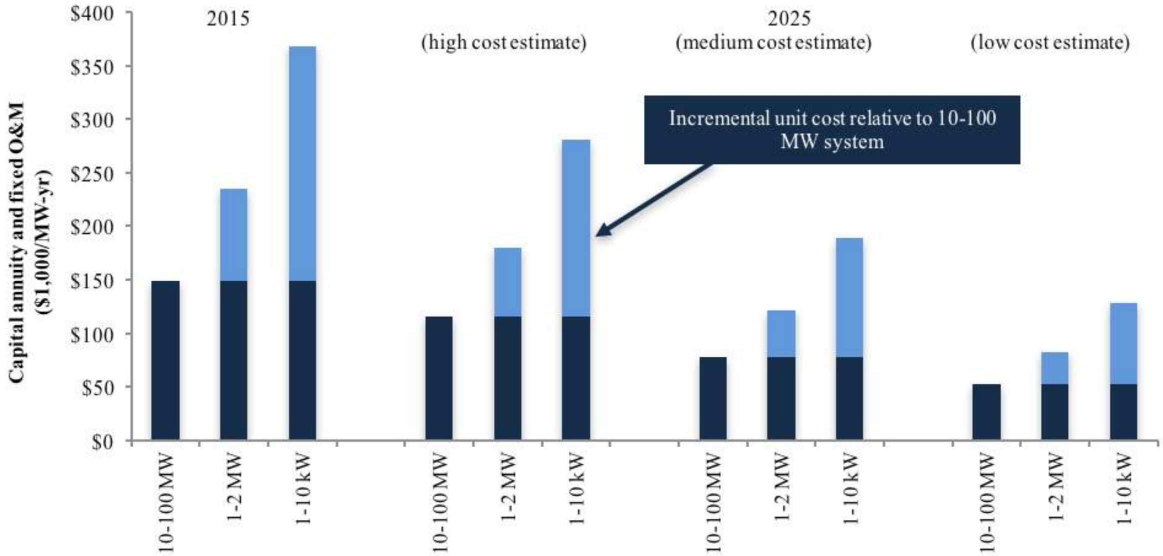
35 Per Lazard (2015a), the capital annuity assumes a 7.68 percent after-tax weighted average cost of capital (8 percent cost of debt, 12 percent cost of equity, 40 percent total business tax rate, 40 percent equity share) and 25-year asset life.

36 The cost shown for 10–100 MW systems in 2015 is from Lazard (2015a). The costs shown for 1–2 MW and 1–10 kW systems are based on average economies of unit scale across estimates from Lazard (2015a) and NREL (2016).

37 Projected cost reductions from 2015 to 2025 assume a 25, 50, and 66 percent decline in installed cost and a 0, 25, and 50 percent decline in annual O&M costs across our high, medium, and low cost estimates, respectively. For comparison, NREL (2016) projects cost reductions through 2025 ranging from 4 to 65 percent for installed costs and 0 to 53 percent across solar systems of various sizes. Note that economies of unit scale depend significantly on installation and balance-of-system costs; in some countries (e.g., Germany, Australia), installed costs for smaller-scale systems can be lower than in the United States.

38 Note that interconnection costs or transmission extensions needed to access locations suitable for utility-scale solar projects can result in greater variation in the cost of larger-scale solar projects. The figures shown here assume typical interconnection costs and assume no transmission expansion.

Figure 8.20: Estimated Economies of Unit Scale for Fixed-Tilt Solar PV Systems—Annual Costs of Ownership in 2015 and Projected for 2025



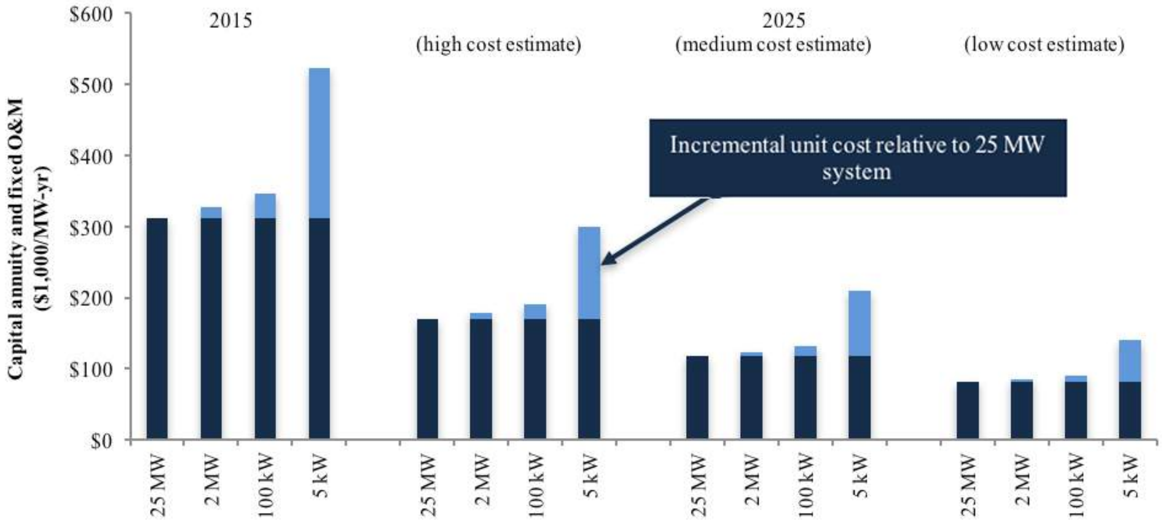
The rate at which economies of scale are exhausted varies significantly from technology to technology. Traditional nuclear, hydroelectric, natural gas, and coal-fired power plants are typically installed at scales of several hundred MW to more than a thousand MW, thereby maximizing cost reductions due to economies of unit scale. In contrast, economies of scale for solar PV units appear to be exhausted at the scale of tens to hundreds of MW. Electrochemical energy storage systems tend to exhibit even more rapidly diminishing economies of unit scale, as depicted in **Figure 8.21**.[39]

In any case, economies of scale matter, even for distributed resources. For resources that can be deployed at multiple scales, incremental unit costs must be considered alongside possible increases in locational value associated with smaller-scale, distributed installation of these resources. Trade-offs between locational value on the one hand and incremental unit costs due to economies of unit scale on the other hand can help identify the ideal locations and applications for these resources. In many cases, after factoring in both costs and benefits, smaller is not better. Indeed, when incremental unit costs exceed locational value, society incurs "distributed opportunity costs" when small-scale DERs are deployed in lieu of more cost-effective resources.

To illustrate this trade-off, consider our case study of locational value for distributed solar PV systems in New York (presented in **Section 8.2)**. The high-value case in Long Island, New York, produced nearly $85 per MWh in locational value, while the average-value case in the Mohawk Valley produced roughly $8 per MWh in locational value. However, solar PV systems deployed in a distributed fashion to capture this additional locational value also incur incremental unit costs. **Figure 8.22** compares the locational value for both the Mohawk Valley and Long Island cases to incremental unit costs for distributed PV systems at the kW and MW scales. The figure includes both estimated costs for 2015 and a range of possible cost declines through 2025. To facilitate this comparison and account for annual energy production, we present incremental unit costs as the difference

39 Per Lazard (2015a), the capital annuity assumes 7.68 percent after-tax weighted average cost of capital (8 percent cost of debt, 12 percent cost of equity, 40 percent total business tax rate, 40 percent equity share) and 10-year asset life. Estimates for 2015 cost, economies of unit scale, and 2025 cost declines are based on a range of estimates in Lazard (2015b) and on personal correspondence with GTM Research. Projected cost reductions for 2025 assume 50, 66, and 75 percent declines in installed cost and 0, 25, and 50 percent declines in annual O&M costs across high, medium, and low cost estimates, respectively.

Figure 8.21: Estimated Economies of Unit Scale for Lithium Ion Energy Storage Systems (with a 1:2 Power to Energy Ratio)—Annual Costs of Ownership for 2015 and Projected for 2025



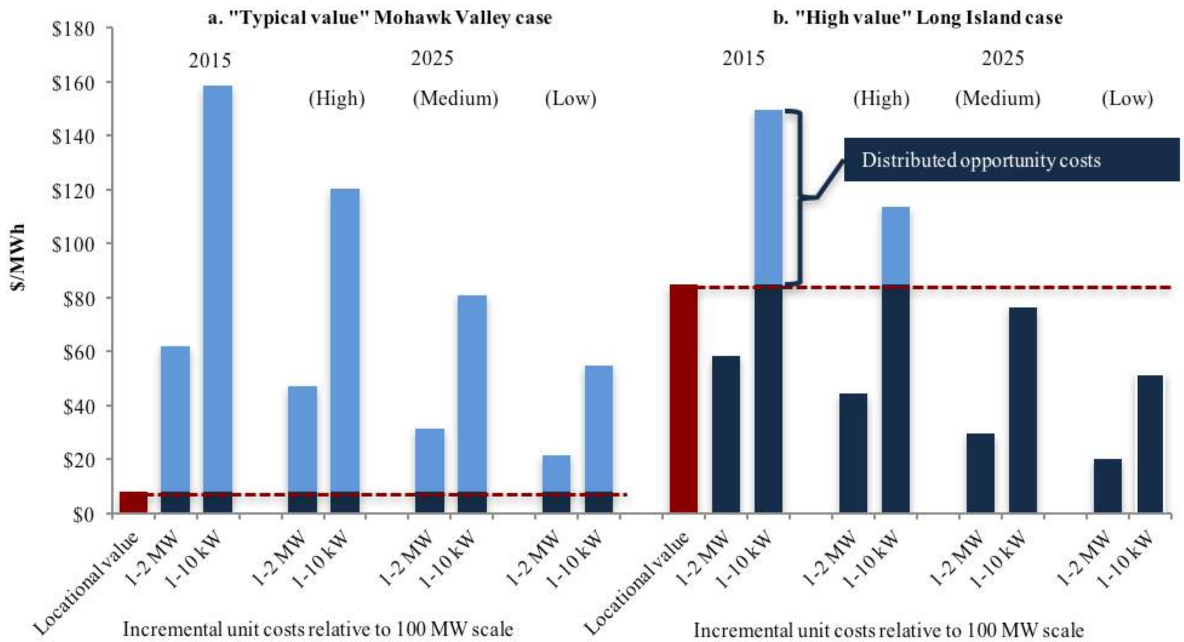
in levelized cost per MWh[40] between distributed PV systems in Long Island and Mohawk Valley, New York, and a 100 MW, utility-scale, fixed-tilt solar PV system with an equivalent solar insolation quality.[41] As **Figure 8.21** illustrates, incremental costs for both kW- and MW-scale distributed solar PV systems exceed the relatively modest locational benefits in the typical value case we constructed for Mohawk Valley. This holds for both 2015 estimated costs and all 2025 cost projections. In short, locational benefits are not sufficient in the illustrative Mohawk Valley case to justify deploying solar PV at a distributed scale. Larger, utility-scale solar would deliver greater net benefits in this case, while small-scale PV deployment incurs distributed opportunity costs.

In contrast, the much greater locational value in the high-value case we constructed for Long Island outweighs the incremental unit costs for MW-scale distributed solar systems at estimated 2015 and projected 2025 costs. At estimated 2015 costs and at the high end of projected 2025 costs, smaller, kW-scale rooftop systems would still entail a distributed opportunity cost in the Long Island case, as incremental unit costs still exceed locational value. However, incremental unit cost for rooftop-scale PV systems falls below the locational value for the medium- and low-range 2025 cost projections. In the high-value Long Island case, distributed solar PV is thus economically justified, although the greatest net benefit would accrue to the largest-scale PV system capable of capturing the locational value in that part of the network. In other words, even at locations where distributed deployment is economically justified, exhausting economies of unit scale to the greatest extent possible in that setting yields the greatest economic benefits.

40 Note that levelized costs are inadequate to compare technologies with very different production patterns or locations (e.g., comparing solar PV systems to combined cycle gas plants or nuclear generators) because the value of these resources can differ substantially. In this case, where we are comparing solar PV systems with very similar production profiles and location, and explicitly factoring in differences in value, levelized cost is an appropriate metric.

41 Solar PV production profiles from NREL PVWatts Calculator assuming fixed-tilt systems at all scales. Annual production for Long Island location is 1,458 MWh per MW installed and 1,376 MWh for Mohawk Valley. See note 40 above for cost and discount rate assumptions. Levelized cost of electricity is calculated using a formula from NREL (2016). Note that utility-scale solar systems with one-axis tracking can exhibit even lower levelized costs than the fixed-tilt system used for comparison here.

Figure 8.22: Comparison of Locational Value and Incremental Cost for Solar PV Systems in New York State Example



Economies of scale matter, even for distributed resources. For resources that can be deployed at multiple scales, trade-offs between locational value and incremental unit costs due to economies of unit scale can help identify the ideal locations and applications for these resources. In many cases, after factoring in both costs and benefits, small-scale deployment of DERs can incur "distributed opportunity costs." Smaller is not always better.

This example illustrates important trade-offs between locational value and economies of unit scale. While the values and distributed opportunity costs presented here are case-specific, the basic heuristic can be broadly applied to develop insight about both the value and cost of distributed deployment of solar PV, energy storage, and other resources that can be deployed at multiple scales.

In particular, for resources that exhibit significant incremental unit costs—as is the case today for solar PV, at both kW and MW scales, and for electrochemical energy storage at the kW scale—distributed deployment is likely to be inefficient in many locations. At specific sites, such as in portions of the network that are experiencing frequent congestions or in areas that are confronting rapid growth in electricity demand, locational value can be significant and can give distributed solar or storage technologies an economic advantage over larger-scale deployment of these resources. On the other hand, in well-developed networks that experience congestions and service interruptions only infrequently, locational value can be modest, as the Mohawk Valley case illustrates. In such locations, economies of unit scale generally outweigh locational value given current and projected costs for solar PV and electrochemical energy storage systems. In these settings, innovation in the underlying technologies and/or learning-by-doing to reduce system installation costs may be needed to flatten economies of unit scale before widespread distributed deployment of these resources becomes economically efficient.

For resources that exhibit significant distributed opportunity costs, distributed deployment is likely to be inefficient in many locations. Exceptions may include heavily congested networks or areas experiencing rapid growth in electricity demand.

Once again, it should be noted that a cost-reflective system of prices and charges with sufficient spatial granularity plays an essential role in incentivizing patterns of DER investment that balance trade-offs between economies of unit scale and locational value. Without a system of prices and charges that conveys sufficient locational granularity, investors will be unable to identify sites where DERs are likely to deliver the greatest locational value. Likewise, inaccurate price signals that overcompensate distributed resources relative to their true locational value may encourage patterns of DER deployment that entail significant distributed opportunity costs. Electricity prices, regulated charges, and other incentives therefore need to become much more sophisticated in terms of capturing the marginal locational value of DERs with sufficient temporal and locational granularity (see **Chapter 4**). Accurate price signals or other incentives are essential to unlock the full value of DERs, put all resources on a level competitive playing field, and facilitate decision-making that internalizes trade-offs between locational value and distributed opportunity costs.

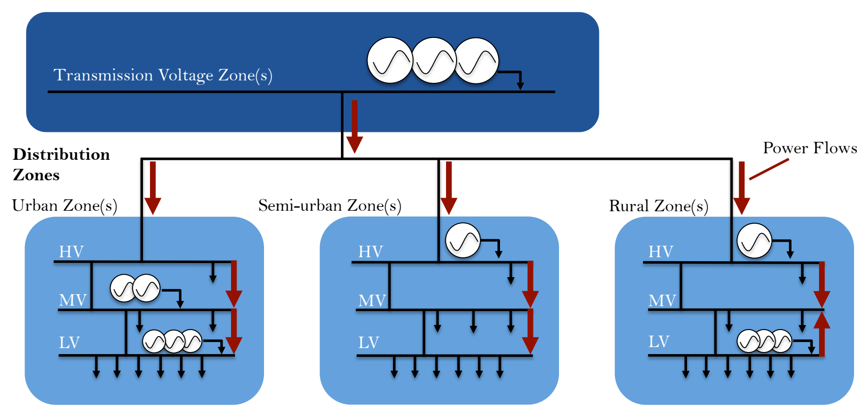
Accurate price signals or other incentives are essential to unlock the full value of DERs, put all resources on a level competitive playing field, and facilitate decision-making that internalizes trade-offs between locational value and distributed opportunity costs.

Box 8.3: Comprehensive Cost-Benefit Assessment for Distributed Energy Resources

As this chapter illustrates, DERs offer both new options for the delivery of electricity services and new trade-offs for power systems. DERs compete with, complement, and coexist alongside one another and alongside a diverse portfolio of conventional generation resources and network infrastructure. At the same time, DERs have very different impacts on electricity transmission and distribution networks compared to conventional resources and can often be deployed at multiple voltage levels and scales, exhibiting trade-offs between locational value (**Section 8.2**) and economies of unit scale (**Section 8.4**).

To fully evaluate these complex trade-offs and interactions in a long-term context, new models must be employed that capture the salient dynamics. GenX, a comprehensive electricity resource capacity expansion model, is described more extensively in **Appendix A**. It incorporates state-of-the-art formulations that capture and optimize trade-offs between economies of unit scale and locational value for a variety of DERs and conventional resources. GenX employs a simplified zonal representation of electricity networks with electricity resources available for deployment at one or more transmission and distribution voltage levels, each exhibiting different unit costs, with power flows and constraints between zones (**Figure 8.23**). Using mathematical formulas that capture in sufficient detail the results of extensive offline modeling of distribution network power flows, the GenX model accounts for the impact of DERs and electricity demand on distribution network losses and reinforcement costs as well as transmission-level losses and constraints or reinforcements. The model also captures operational constraints on generators, operating reserve requirements, and opportunities for flexible demand or price-responsive demand curtailment.

Figure 8.23: Power System Representation in the GenX Model



As an example of the opportunity this kind of modeling presents for evaluating DERs in a complex power system, we present the results of a case study of energy storage in a Spain-like test system.[42] Battery energy storage can be deployed almost anywhere and at multiple scales, ranging from a few kW to the multi-MW range, and can provide multiple services to the power system, such as reserves, capacity, energy arbitrage, and firm capacity. In this case, we explore opportunities for using energy storage to avoid or defer transmission network reinforcement costs in a constrained link between a bulk power system generation zone and a distribution network zone where energy consumers are located. Storage can be employed at the 5 kW scale in low-voltage distribution or the 100 kW scale in medium-voltage distribution, in both cases "downstream" from the transmission constraint. Storage can also be employed at the 25 MW scale in the bulk transmission voltage zone "upstream" of the constraint. Battery options with both two hours and four hours of storage duration are considered (a 2:1 and 4:1 energy-to-power ratio, respectively).

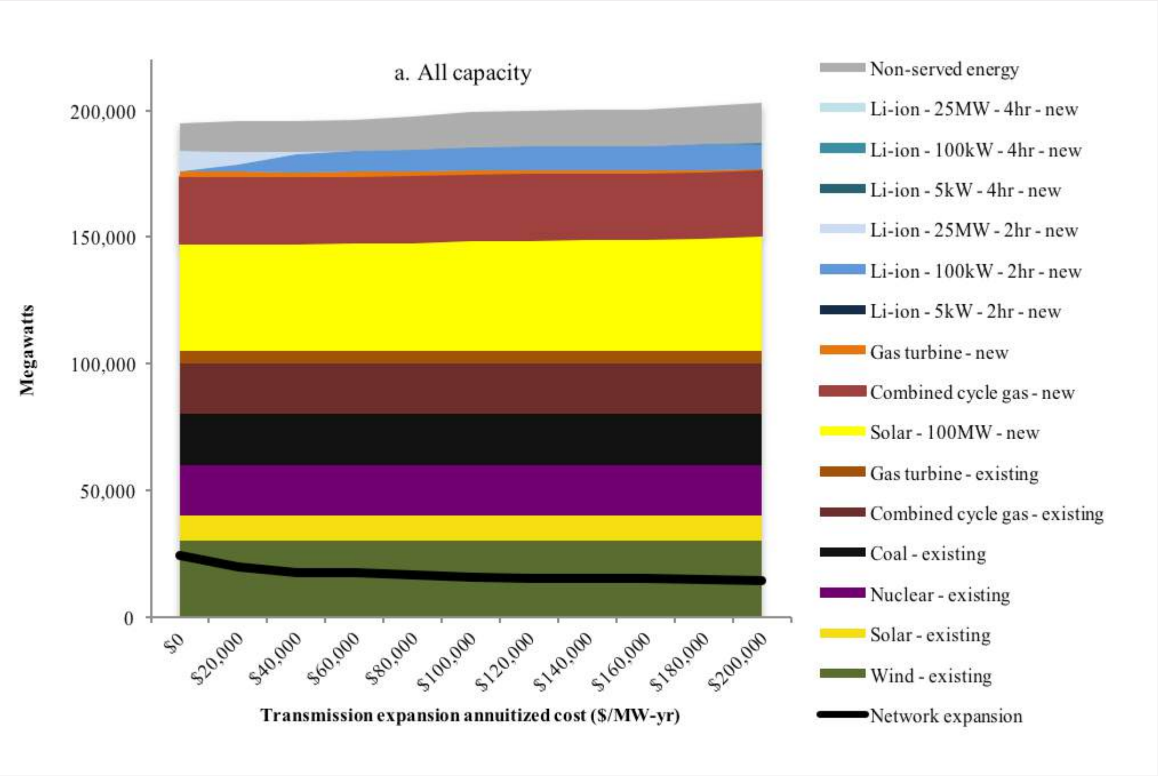
We steadily increase the cost of transmission network reinforcements to explore trade-offs between distributed storage with greater locational value and MW-scale storage with lower unit costs. (**Figure 8.24**). As transmission network costs increase, the optimal location for energy storage shifts from transmission voltage at the 25 MW scale to medium-voltage distribution at the 100 kW scale, while expansion of transmission capacity also declines. The 100 kW system exhibits the lowest unit cost for a system that can be sited where it relieves the transmission constraint and defers network investments. As such, the model never selects the more costly 5 kW storage option for inclusion in the least-cost resource portfolio. Distributed solar PV (at 5–10 kW and 100 kW–1 MW scales) is also an option, but it is not selected by the model in this case, because PV production is non-coincident with peak demand in this test system and thus cannot contribute to transmission network deferral. Furthermore, the optimal duration of energy storage in this case is two hours, with a small amount of four-hour storage appearing only when annuitized transmission reinforcement costs exceed $180,000 per MW-year.[43] Shorter duration storage thus appears sufficient to reduce peak demand during periods of binding transmission constraints, thereby avoiding the need for reinforcements—at least given this specific load profile. In addition, as total storage capacity increases modestly across the cases, solar PV capacity increases, reflecting the complementary nature of these resources. Likewise, the capacity of natural gas open-cycle and combined-cycle units declines slightly as storage and PV capacity increases, indicating that solar and storage are at least partial substitutes for gas-fired resources in this context.

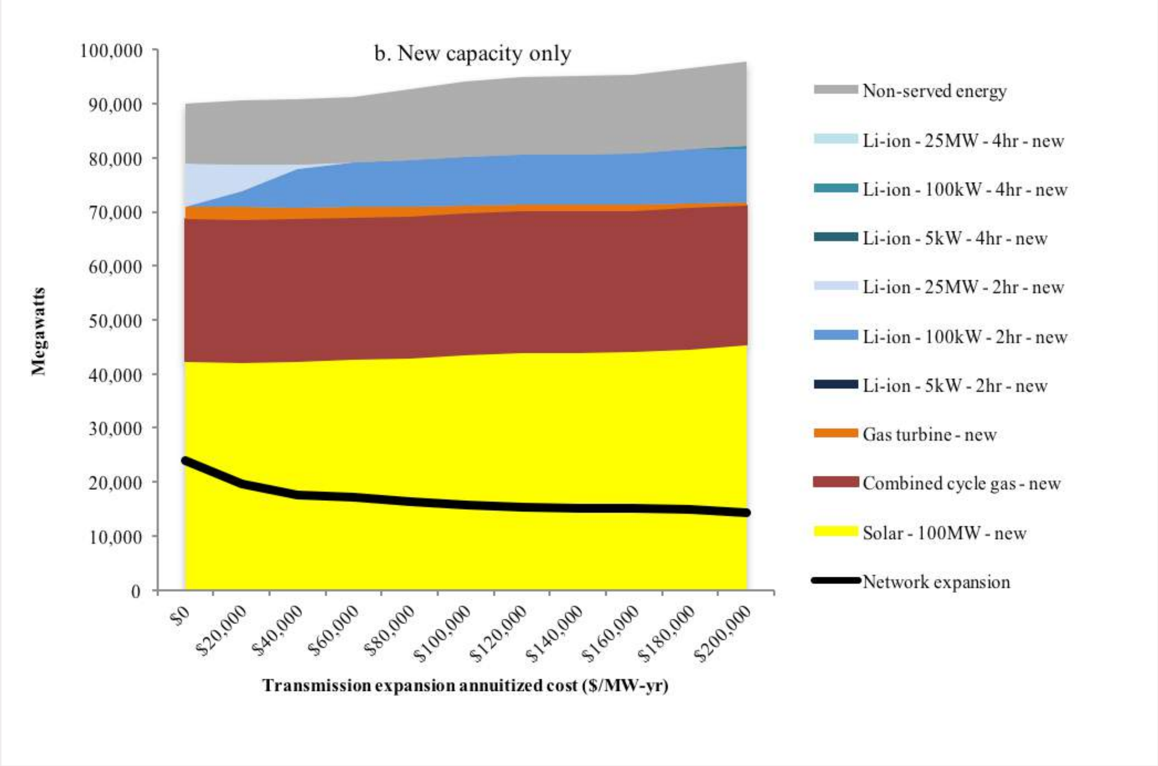
These results are meant to be purely illustrative and to demonstrate the kind of system-wide analysis that is important to develop insights about the role of DERs in the power systems of the future and to explore when and how DERs can add value to power systems.

42 Hourly demand profiles are based on data from the Spanish regulatory commission (CNMC 2014) and scaled to a peak demand of 100 GW. For this test system, we assume the following existing capacity mix: 5,000 MW of open cycle gas turbines, 20,000 MW each of combined cycle gas turbines, pulverized coal, and nuclear reactors; 30,000 MW of onshore wind; and 10,000 MW of solar PV. Energy storage and solar PV costs are the mid-range 2025 costs reported in **Figure 8.20** and **Figure 8.21**. Fuel price assumptions are mid-range assumptions for Europe in 2025 from IEA (2015), and capital and operating cost estimates for other resources are from NREL (2016).

43 In this case, 100 kW-scale lithium-ion storage systems with a 2:1 energy-to-power ratio have installed system costs of $724/kW ($362/kWh), while 100 kW-scale systems with a 4:1 energy-to-power ratio have installed costs of $1,305/kW ($326/kWh). With a lower installed cost per kW of rated power capacity, the model thus favors the shorter-duration storage except in cases when larger energy capacity is significantly more valuable.

Figure 8.24: Effect of Transmission Expansion Cost on the Optimal Generation and Energy Storage Capacity and Transmission Expansion Decisions (Spain-like Test System)





References

ACCC. 2012a. "Benchmarking Opex and Capex in Energy Networks." Australian Competition and Consumer Commission and Australian Energy Regulator Working Paper 6. May 2012.

Ball, E.B.R., N. Keers, and M. Alexander. 2010. *Mobile Energy Resources in Grids of Electricity*. MERGE Project Deliverable D2.1. Janaury 4, 2010. www.transport-research.info/sites/default/files/project/documents/20140203_154622_76425_Deliverable_2.1_Modelling_Electric_Storage_devices_for_Electric_Vehicles.pdf.

Banez-Chicharro, F., J.M. Latorre, and A. Ramos. 2014. "Smart Charging Profiles for Electric Vehicles." *Computational Management Science* 11 (1–2): 87–110. doi:10.1007/s10287-013-0180-8.

Blum, D. and L.K. Norford. 2014. "Dynamic Simulation and Analysis of Ancillary Service Demand Response Strategies for Variable Air Volume HVAC Systems." *HVAC&R Research* 20(8): 908-921. November 2014.

Campbell, R.J. 2012. *Weather-Related Power Outages and Electric System Resiliency*. Washington, D.C.: Congressional Research Service. doi:R42696.

CNMC. 2014. "Methodology for Computing of Transmission and Distribution Charges" (in Spanish). Comisión Nacional de los Mercados y la Competencia. www.cnmc.es/Portals/0/Ficheros/Energia/Circulares/Circular_3_2014/03_2014_Circular_memoria.pdf.

Cohen, M.A., P.A. Kauzmann, and D.S. Callaway. 2016. "Effects of Distributed PV Generation on California's Distribution System, Part 1: Engineering Simulations." *Solar Energy* 128: 139-52. doi:10.1016/j.solener.2016.01.002.

Con Edison. 2016. Brooklyn Queens Demand Management Demand Response Program Guidelines. Consolidated Edison. New York, NY. conedbqdmauctiondotcom.files.wordpress.com/2016/03/bqdm-dr-program-overview-6-28-161.pdf.

DOE. 2012. *2011 Buildings Energy Data Book*. Washington, DC: US Department of Energy. March 2012.

Driesen, J. and F. Katiraei. 2008. "Design for Distributed Energy Resources." *IEEE Power Energy Magazine* 6: 30-40. doi:10.1109/MPE.2008.918703.

EIA. 2016. Electric power sales, revenue, and energy efficiency Form EIA-861 detailed data files. Washington, DC: US Energy Information Administration. www.eia.gov/electricity/data/eia861/. Accessed January 3, 2016.

EIA. 2014. *Price Elasticities for Energy Use in Buildings of the United States*. Washington, DC: US Energy Information Administration. October 22, 2014.

EIA. 2015. *Annual Energy Outlook 2015*. Washington, DC: US Energy Information Administration. April 2015.

EPA. 2016. *The Social Cost of Carbon*. Washington, DC: US Environmental Protection Agency. September 29, 2016. www.epa.gov/climatechange/social-cost-carbon. Accessed October 28, 2016.

EnerData. 2015. Energy Efficiency Indicators: Rate of Electricity T&D Losses, 2014. www.wec-indicators.enerdata.eu/world-rate-of-electricity-T-D-losses.html. Accessed October 17, 2016.

EPRI. 2015. Alternatives to the 15% Rule: Modeling and Hosting Capacity Analysis of 16 Feeders. Palo Alto, CA: Electric Power Research Institute.

Georgiopoulos, S. n.d. Interruptible Connections: Lessons from Flexible Plug and Play. UK Power Networks. innovation.ukpowernetworks.co.uk/innovation/en/Projects/tier-2-projects/Flexible-Plug-and-Play-(FPP)/Project-Documents/UK-Power-Networks.pdf.

Goop, J., M. Odenberger, and F. Johnsson. 2016. "Distributed Solar and Wind Power—Impact on Distribution Losses." Energy 112: 273-284. August 2016.

Hassett, M.A. and E. Bower. 2011. Evaluation of the Impact That a Progressive Deployment of EV Will Provoke on Electricity Demand, Steady State, Operation, Market Issues, Generation Schedules and on the Volume of Carbon Emissions—Electric Vehicle Penetration Scenarios in Germany, UK, SP. Project MERGE. textlab.io/doc/877048/deliverable-3.2-ev-penetration-scenarios.

Herter, K. and Y. Okuneva. 2013. SMUD's 2012 Residential Precooling Study—Load Impact Evaluation. Prepared for Sacramento Municipal Utility District. Sacramento, CA: Herter Energy Research Solutions. June 11, 2013.

IEA. 2015. *World Energy Outlook*, 2015. Paris, France: International Energy Agency. November 2015.

Jamasb, T., P. Nillesen, and M. Pollitt. 2004. "Strategic Behaviour Under Regulatory Benchmarking." *Energy Economics* 26(5): 825–843. dx.doi.org/10.1016/j.eneco.2004.04.029.

Kim, Y-J., L.K. Norford, and J.L. Kirtley. 2015. "Modeling and Analysis of a Variable Speed Heat Pump for Frequency Regulation Through Direct Load Control." *IEEE Transactions on Power Systems* 30(1): 397-408. January 2015.

Kim, Y-J, D.H. Blum, N. Xu, L. Su, and L.K. Norford. 2016. "Technologies and Magnitude of Ancillary Services Provided by Commercial Buildings." *Proceedings of the IEEE* 104(4). April 2016.

Lazard. 2015a. *Lazard's Levelized Cost of Energy Analysis—Version 9.0*. New York, NY: Lazard. November 2015.

Lazard. 2015b. *Lazard's Levelized Cost of Storage—Version 1.0*. New York, NY: Lazard. November 2015.

Mateo, C., P. Frías, and A. Sánchez Miralles. 2016. "Distribution Planning with Hourly Profiles for Analysing Electric Vehicle Charging Strategies." *International Journal of Electric and Hybrid Vehicles* 8(1). doi:10.1504/IJEHV.2016.076959.

Neme, C. and R. Sedano. 2012. *US Experience with Efficiency as a Transmission and Distribution System Resource*. Montpelier, VT: Regulatory Assistance Project. February 2012.

NREL. 2016. *2016 Annual Technology Baseline, Discussion* Draft. Golden, CO: National Renewable Energy Laboratory. April 2016.

NYISO. 2016. *2016 Load and Capacity Data "Gold Book."* Rensselaer, NY: New York Independent System Operator. April 2016.

NYSERDA. 2011. *New York Solar Study*. Albany, NY: New York State Energy Research and Development Agency.

NYSERDA, 2015. RFP 3044: NY Prize Community Grid Competition. Albany, NY: New York State Energy Research and Development Agency.

PGE. n.d. "Dispatchable Standby Generation." Portland Generation Electric. www.portlandgeneral.com/business/get-paid-to-help-meet-demand/dispatchable-standby-generation. Accessed October 17, 2016.

Prettico, G., F. Gangale, A. Mengolini, A. Lucas, and G. Fulli. 2016. *Distribution System Operators Observatory: From European Electricity Distribution Systems to Representative Distribution Networks*. European Commission, Joint Research Centre. doi:10.2790/701791.

Schmalensee, R., V. Bulovic, R. Armstrong, C. Batlle, P. Brown, J. Deutch, et al. 2015. *The Future of Solar Energy*. Cambridge, MA: Massachusetts Institute of Technology.

Schneider, K.P., J.C. Fuller, F.K. Tuffner, and R. Singh. 2010. PNNL-19596: Evaluation of Conservation Voltage Reduction (CVR) on a National Level. Richland, WA: Pacific Northwest National Laboratory. doi:10.2172/990131.

Schweppe, F.C., M.C. Caramanis, R.D. Tabors, and R.E. Bohn. 1988. *Spot Pricing of Electricity*. Kluwer Academic Publishers.

Smarter Grid Solutions. n.d. "Managed Grid Connections for Generators." Glasgow, Scotland: Smarter Grid Solutions. www.smartergridsolutions.com/media/112363/ukpn_case-study_us_final.pdf.

Sullivan, M.J., M. Mercurio, and J. Schellenberg. 2009. *Estimated Value of Service Reliability for Electric Utility Customers in the United States*. Berkeley, CA: Lawrence Berkeley National Laboratory. doi:LBNL-2132E.

Sullivan, M.J., J. Schellenberg, and M. Blundell. 2015. *Updated Value of Service Reliability Estimates for Electric Utility Customers in the United States*. Berkeley, CA: Lawrence Berkeley National Laboratory. doi: LBNL-6941E.

Tsuchida, B., S. Sergici, B. Mudge, W. Gorman, P. Foxpenner, and J. Schoene. 2015. Comparative Generation Costs of Utility-Scale and Residential-Scale PV in Xcel Energy Colorado's Service Area. Cambridge, MA: The Brattle Group.

Wang, Z. and J. Wang. 2014. "Review on Implementation and Assessment of Conservation Voltage Reduction." *IEEE Transactions on Power Systems* 29: 1306-15. doi:10.1109/TPWRS.2013.2288518.

Xu, N. 2015. "Active Participation of Buildings in the Power Sector: The Case of Small Office Buildings." Master's thesis. Cambridge, MA: Massachusetts Institute of Technology. September 2015. hdl.handle.net/1721.1/103567

Zimmerman, R.D., C.E. Murillo-Sánchez, and R.J. Thomas. 2011. "MATPOWER: Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education." *IEEE Transactions on Power Systems* 26(1): 12-19. February 2011. For latest version of model see www.pserc.cornell.edu/matpower/.

PART 4: A POLICY AND REGULATORY TOOLKIT FOR THE FUTURE POWER SYSTEM

09

A Toolkit for Regulators and Policy Makers

Throughout this study we offer recommendations for proactive regulatory, policy, and market reform designed to enable the efficient evolution of the power system over the next decade and beyond. The purpose of our recommendations is to enable all energy resources, whether distributed or centralized, to participate in the efficient provision of electricity services while achieving other public policy objectives. The overarching framework for these recommendations has two interrelated pillars: Establish a comprehensive system of prices and regulated charges that applies to all network users, and remove inefficient barriers to the integration and competition of distributed resources and centralized resources alike. Those barriers include inadequate remuneration schemes for distribution utilities, power sector structures that may impede fair competition, and wholesale electricity market design flaws.

This chapter collects our core policy and regulatory recommendations for establishing efficient prices and charges and for removing inefficient barriers. It is intended to be used as a quick reference for policy makers, market designers, regulators, and other practitioners. **Chapters 4** through **8** contain detailed analyses, modeling results, and theoretical underpinnings for each of the recommendations made herein; where more detail is desired on the recommendations or on questions of practical implementation, readers should refer to previous chapters in this report.

The task facing those responsible for the reliable and cost-effective planning, regulation, and operation of future power systems is daunting; the sheer length and weight of this study seem to confirm that. However, many of the recommendations outlined in this chapter and throughout the report are implementable and can deliver significant benefits today. Throughout the report we pay careful attention to the practical implementation challenges of our recommendations and, where possible, we highlight examples of their real-world implementation and outline steps that bridge theory and practice. These discussions are intended to enable proactive policy and regulatory reform of the power sector; proactive reform is the only defense against being caught flat-footed by challenges that seem minor today, but could seem insurmountable tomorrow.

9.1 Recommendations for Establishing an Efficient System of Prices and Charges

Centralized and distributed resources are installed at different locations and different scales in power systems. In many cases they are also operated with different temporal patterns of production and consumption. Nonetheless, centralized and distributed resources compete to provide a limited set of electricity services. The services that DERs are uniquely positioned to provide — those services with high locational value — often compete with network infrastructure or other means of providing electricity services. Explicit coordination of the multiplicity of agents that can provide electricity services is necessary. We recommend using a comprehensive system of prices and charges as the method for coordination. **Chapter 4** discusses the recommendations made in this subsection in greater detail.

Recommendation 1: Create a comprehensive and cost-reflective system of prices and charges.

The only way to enable centralized and distributed resources to jointly and efficiently operate, expand, compete, and collaborate, is to establish a comprehensive and cost-reflective system of economic signals — prices and regulated charges — with adequate granularity in service type, time, and location. Prices and regulated charges collectively determine, at each connection point and time, the value of the services provided or consumed by any particular agent. Incorrect economic signals can drive inefficient investment and operational decisions, enable costlier resources to displace more efficient ones, enable inefficient business models to crowd out efficient ones, and result in more expensive electricity services and a loss of societal welfare.

Establish a comprehensive system of cost-reflective prices (for those services provided in markets) and regulated charges (for remuneration of network activities and any policy costs included in electricity rates) that apply equally to all network users — centralized or distributed — and that adequately reflect time- and location-specific system conditions.

Recommendation 2: Ensure that all prices and charges are non-discriminatory and technology neutral.

As network users become more responsive to prices and as new loads, such as electric vehicles (EVs), or new devices, such as distributed generation and storage, emerge, it will no longer be meaningful to continue to use existing customer classifications. Network utilization patterns are becoming increasingly diverse, forcing us to abandon the assumption that underlies most current tariff designs: that consumers can be lumped into homogenous customer classes. Furthermore, impacts on the power system depend on the specific pattern of injection or withdrawal of power at a given location. Prices and charges should thus be technology agnostic and based only on the injections and withdrawals of electric power at a given time and place, rather than on the specific devices behind the meter. In addition, cost-reflective prices and regulated charges should be symmetrical, with injection at a given time and place compensated at the same rate that is charged for withdrawal at the same time and place.

Cost-reflective electricity prices and regulated charges should be based only on what is metered at the point of connection to the power system — that is, the injections and withdrawals of electric power at a given time and place, rather than the specific devices behind the meter. These cost-reflective prices and regulated charges should be symmetrical, with injection at a given time and place compensated at the same rate that is charged for withdrawal at the same time and place.

Recommendation 3: Ubiquitously deploy modern metering infrastructure.

Deploying modern metering infrastructure[1] — whether in the form of today's advanced meters or something altogether different — is imperative. Without this infrastructure, it is impossible to meaningfully develop a comprehensive system of prices and charges and accurately meter, compensate, and charge a diversity of electricity resources. As a result, the value of the services that distributed resources and flexible demand provide cannot be measured and tapped.

> Modern information and communications technologies must be deployed to all points of injection and withdrawal within the network

Recommendation 4: Find the efficient level of granularity of prices and charges for each power system given its characteristics and regulatory context.

The prices and charges for electricity services may vary significantly with space and time when computed according to the most precise cost-reflective methods. However, this does not necessarily imply that regulators must adopt prices and charges with the maximum level of granularity in order to achieve a satisfactory level of efficiency. Implementing more granular prices and charges involves costs. Thus, temporal and locational granularity should be increased only as long as the incremental efficiency gains exceed the incremental implementation costs. Complexity and equity concerns related to implementation (to be discussed later) are relevant factors that should also be considered.

> Progressively improving the temporal and locational granularity of prices and charges can deliver increased social welfare; however, these benefits must be balanced against the costs, complexity, and potential equity concerns of implementation.

Some simple-to-adopt measures may allow regulators to capture a significant fraction of potential efficiency gains. The following recommendations highlight potential intermediate improvements toward "getting the prices and charges right" that can capture the low-hanging fruit at reasonable implementation costs in most power systems.

Recommendation 5: Collect policy costs and residual network costs in a minimally distortive fashion.

The first measure to be adopted is not about enhancing granularity, but about reducing the distortion caused by charging policy costs (e.g., for efficiency programs, subsidies for renewable energy, etc.), taxes, and residual network costs (i.e., network costs that are not recovered via cost-reflective charges) within the volumetric, dollar-per-kilowatt-hour ($/kWh) component of the tariff, as is today the practice in most power systems. Any policy costs, taxes, and residual network costs that are not directly affected by changes in electricity consumption or injection should be removed from the volumetric ($/kWh) component of the tariff and charged in a manner that minimizes distortions of cost-reflective prices and charges for electricity services.

1 Modern metering infrastructure encompasses a broad swath of technologies that could be capable of providing temporally and spatially granular data on real and reactive power consumption and/or production. The suite of technologies capable of providing these functions is rapidly evolving, so we are not prescriptive as to the exact form that these technologies should take.

Regulators and policymakers must also carefully monitor conditions that could lead to a serious threat of inefficient grid defection. If inefficient grid defection is a serious possibility, regulators should reconsider the costs that are included in the tariff, or the adoption of other measures (e.g., an exit charge), to prevent substantial cross-subsidization among consumers and a potential massive defection with unforeseen consequences.

The costs of public policies, residual network costs, and taxes that are not directly affected by changes in electricity consumption should not be recovered with volumetric charges ($/kWh) in electricity tariffs. We recommend a fixed charge (an annual lump sum conveniently distributed in monthly installments). The magnitude of each customer's charge should be dependent upon some proxy metric of lack of price elasticity or some measure of wealth, such as the property tax or the size of a system user's dwelling. Depending on the seriousness of the threat of grid defection, which costs are included in the electricity tariff must be carefully considered.

Recommendation 6: Ensure that time differentiation is reflected in energy prices.

In most power systems, energy prices vary significantly over time throughout each day (except in systems with large hydro storage). Large penetrations of intermittent renewables with zero variable cost, in typical power systems with significant proportions of thermal generation, increase this time variability. Wholesale energy prices (or short-term marginal costs, in the absence of markets), reflect this time variation in the price of energy and can be communicated to all network agents.

Communicate and apply the time variation of wholesale energy prices (hourly or sub-hourly, typically) to end customers via advanced meters. Where wholesale markets do not exist, communicate and apply short-term marginal costs.

It is easy to add a first "locational component" to the wholesale energy prices (see **Recommendation 8** for a complete discussion on locational granularity) to reflect the impact of the distribution network on the prices seen by end customers connected at different voltage levels, to a first approximation. A simple approach, which does not raise equity issues, consists of using just one loss factor per voltage level at a given time. Loss factors may change significantly with time, due to varying operating conditions in power systems.

Recommendation 7: Apply forward-looking peak-coincident capacity network charges to network users.

Depending on the characteristics of the power system, peak-coincident signals may become critically important to reduce the need for future investments in generation and networks. Once there is an advanced meter available that can record the hourly or sub-hourly injection or withdrawal of power at every connection point, there is no conceptual difficulty in applying peak coincidental capacity charges for firm capacity and for responsibility in network investment. Peak-coincident charges for network investments generally should have locational differentiation, as networks do not get congested uniformly. If locational marginal prices (LMPs) capturing congestion in distribution networks are employed, peak-coincident network charges must be carefully coordinated with the LMP to send signals that lead to efficient short- and long-run outcomes. Of course, simplifications can be made if regulators or policy makers are concerned about equity issues or other complications of establishing peak prices that are highly granular with respect to location. Sensible peak pricing notably enhances the level of comprehensiveness of a system of prices and charges without causing significant implementation efforts or costs (although the difficulty of regulatory design and public acceptance should not be underestimated).

Peak-coincident capacity charges that reflect a user's contribution to incremental network costs incurred to meet peak withdrawals and injections, as well as scarcity-coincident generating capacity charges, can unlock flexible demand and distributed resources and enable significant cost savings.

Recommendation 8: Progressively increase the locational granularity of economic signals.

The locational component of prices for energy and other services can be introduced or augmented to reflect differences in prices and regulated charges at the transmission level. We recommend a shift towards providing more spatial granularity at the wholesale level, with nodal prices representing the ideal target, if workable. The most important hurdles to implementing nodal prices involve institutional and governance issues. This is particularly true when the aim is to create regional markets — that is, integrating markets across many different countries, as in the European Internal Electricity Market, or across power systems in the Eastern or Western Interconnections in the United States.

Significant and persistent differences in the values of transmission LMPs among nodes — either actually existing or estimated — indicate that locational price signals should be employed at the wholesale level. Generators connected to the transmission system should be the first to be exposed to locational signals in order to avoid distortions in the economic dispatch. Second, agents connected at the distribution level should also be exposed to locational prices that capture differences in marginal prices at the transmission level (as well as any distribution locational components, to be discussed later), thereby allowing them to react in an efficient manner. Many US wholesale markets calculate locationally granular energy prices as frequently as every five minutes — however, in very few cases are these prices communicated to agents in the distribution system or used to settle consumption in the distribution system. Clearing a market where participants face different prices creates arbitrage opportunities and is computationally complex.

In principle, locational price signals can be further improved to reflect the effects of losses and congestion in the distribution system. To date, only a handful of markets have somehow included the impact of distribution network losses on energy prices, usually in a very simplified manner;[2] no markets today account for the impact of distribution network constraints on energy prices. There may be good reasons for the application of detailed locational signals at the distribution level, even if LMPs are not used in the transmission network.

Locational granularity matters. Including a locational component in electricity prices and charges at transmission and also distribution levels will deliver efficiency gains, but also adds complexity and may raise equity concerns.

The application of locational prices to end customers creates geographical differences in tariffs, which could be contrary to well-established regulatory practices and may raise equity issues.[3] All these factors, plus the complexity and cost associated with implementing locational prices and the expected efficiency gain have to be considered when deciding the level of locational granularity to be adopted in economic signals.

2 See the comment on loss factors in **Recommendation 6**.

3 Being connected to one distribution network feeder or another is the result of a decision taken solely by the distribution company. That decision can have a significant effect on applicable network charges for small and medium customers, and may result in very different tariffs for network users that are not very far apart. This issue is addressed in **Recommendation 10**.

Recommendation 9: Explore the potential to use the reactive component of the electricity price to detect voltage problems and enable DERs to respond to these problems.

Network constraints at the distribution level are typically related to voltage problems[4] — usually of a local nature. These constraints have to be modeled with a full AC[5] representation of the power flows, with active and reactive components. These computer models can determine the reactive component of the energy price, which is directly associated with the existence of binding active constraints, as well as marginal network losses.

The network investments necessary to mitigate these network problems might be avoided or deferred by the actions of demand response and DERs. The distribution system operator may wish to establish specific contracts for DERs or aggregations of DERs, either directly or through competitive auctions (if there are enough participants to guarantee minimal competitive conditions), to enable cost-effective alternatives to traditional network investments.

The reactive component of the energy price is a good indicator of the existence and economic importance of voltage problems in electricity networks. Because of its local characteristics, diverse ad hoc solutions can be employed at the distribution level to bring about the cost-effective participation of DERs in managing network constraints.

Approaches that make use of reactive LMPs at the distribution level are an area of active research, since the interaction between DERs and reactive prices is not well understood yet. In the future, reliable and scalable methods may be implemented for the purpose of improving the locational signal in energy prices to capture the marginal value of reactive power and the cost of voltage constraints. In the meantime, simpler solutions can be employed at the distribution level to bring about the cost-effective participation of DERs in managing network constraints.

Recommendation 10: Explore methods to address distributional concerns associated with the use of cost-reflective prices and charges.

Improvements in electricity tariff design may raise distributional concerns, including concerns that tariff modifications could significantly change end customers' total expenditures for electricity, produce prices and charges that are very volatile, or lead to prices and charges that are strongly spatially granular and thus impose very different costs on otherwise similar consumers.

Rebates can roughly equalize the average charges seen by consumers (similar to the current situation, where most consumers connected at the same voltage level are subject to roughly equal tariffs within a distribution company or even a country) while retaining the efficient, cost-reflective price signals needed to reward consumers for reducing energy use during costly periods and thereby unlock cost savings for the power system and all consumers. In addition, hedging arrangements are available to address month-to-month bill volatility when such volatility is a concern.

4 By "voltage problems" we mean deviations beyond acceptable limits in the magnitude or shape of the voltage waveform, or extreme volatility in this magnitude, which disrupt the operation of connected devices. The acceptable magnitude of such deviations and associated limits vary between power systems and within power systems.

5 An "alternating current (AC) model" of the power flows, explicitly including the real and reactive components.

Distributional concerns can be overcome and low-income consumers can be protected without giving up the implementation of a more efficient and comprehensive system of prices and charges for broad swaths of network users. Rebates to equalize average charges and mechanisms to hedge month-to-month bill volatility are some of the instruments that can be used to properly address these concerns.

9.2 Recommendations for Removing Barriers

Efficient pricing alone is not sufficient to ensure the cost-effective development of the power sector over the coming decade and beyond. Establishing sound economic pricing is a critical step in enabling all resources to compete efficiently to provide the electricity services required in the power system. However, many inefficient barriers exist today that prevent DERs from participating on an equal footing with centralized resources and vice-versa. These barriers can be grouped into three categories: (1) inadequate regulation of distribution networks for the presence of DERs; (2) power sector structures that lead to conflicts of interest between the network and the commercial components of distribution companies; and (3) biased or inadequate wholesale market rules.[6]

9.2.1 Distribution network regulation

Today's distribution system regulations are ill-adapted for a world with DERs. New cost drivers and new network uses are creating greater uncertainty about the trajectory of distribution system costs, as well as new opportunities to make trade-offs between capital expenditures (CAPEX) and operational expenditures (OPEX), including contracts with or payments to DERs that are capable of contributing to more efficient network operation. Dealing with these new uncertainties and challenges in an efficient manner will require utilities to become more innovative; however, today's regulations lack incentives for innovative behavior. Finally, as the power system becomes more digitalized and as DERs proliferate, distribution systems will face new cyber threats that current regulations are not adequately prepared to manage.

Powerful regulatory tools exist to manage these changes. These regulatory tools can be divided into two categories: (1) improved approaches to distribution remuneration that can account for new cost drivers, new demands and uses of the distribution network, and increased uncertainty; and (2) additional incentive mechanisms for achieving specific outcomes that are not captured by improved distribution remuneration, including performance and quality of service improvements and long-term innovation. Our key regulatory recommendations are summarized here; a detailed discussion can be found in **Chapter 5**.

Recommendation 11: Implement cost efficiency incentives for distribution utilities.

Our first recommendation is to more closely align the business incentives of the distribution company with the remuneration structure, starting with a forward-looking revenue trajectory. Regulators can better align utility incentives for cost-saving investments and operations and ensure that the benefits of improved utility performance are shared between utility shareholders and ratepayers by adopting state-of-the-art regulatory mechanisms. The

6 Distribution network regulation is discussed in detail in **Chapter 5**, power sector structure is discussed in **Chapter 6**, and wholesale market design is the subject of **Chapter 7**.

principal mechanism used to reward efficiency efforts is the multi-year revenue trajectory with profit sharing. This mechanism is a key component of regulatory regimes that go by many names, including "revenue caps," "RPI-X," "multi-year rate plans," and others — such regimes are being successfully used in many power systems today. In each case, the basic objective is the same: Utilities should be assured that, over some defined period of time, their revenues will be to some degree decoupled from their costs, so that utilities will be able to partly (if there is "profit sharing" with their customers) retain any cost savings they may implement during that time period. An essential characteristic is that such multi-year trajectories are set in advance and are forward-looking, as the future will not look like the past.

Forward-looking, multi-year revenue trajectories can reward distribution utilities for cost-saving investments and operations, aligning utilities' business incentives with the continual pursuit of more efficient solutions.

Recommendation 12: Introduce profit-sharing mechanisms and offer a menu of regulatory contracts.

Profit-sharing mechanisms are successfully used today in some power systems in conjunction with multi-year revenue trajectories to share potential profits from efficiency gains and distribute risks between utilities and ratepayers, thereby balancing incentives for productive efficiency with concerns about allocative efficiency. Under a profit-sharing mechanism, utilities retain only a portion of any reductions in cost below the revenue trajectory, with the remaining share accruing to ratepayers in the form of lower rates. Likewise, if actual expenditures exceed the revenue trajectory, utilities bear only a portion of the excess cost, with rates increasing to share the remainder of the burden with ratepayers. In this manner, the profit-sharing mechanism retains utility incentives for cost reduction and improved performance (productive efficiency) but does not fully decouple allowed revenues from realized utility costs, thus improving rent extraction and allocative efficiency (if costs fall below the revenue cap) and mitigating utility exposure to uncertainty (if costs rise above the revenue cap).

Furthermore, the regulator can improve on a single profit-sharing factor by offering regulated utilities a menu of regulatory contracts[7] with a continuum of different sharing factors. A menu of contracts allows the firm to play a role in selecting the strength of cost-saving incentives. If constructed correctly, this menu will establish "incentive compatibility" — that is, the menu is designed to ensure that, *ex ante*, a profit-maximizing utility will always be better off (i.e., earn the greatest profit and return on equity) when the firm accurately reveals its expectations of future costs. Incentive compatibility thus eliminates incentives for firms to inflate their cost estimates while rewarding firms for revealing their true expected costs to the regulator. This helps minimize strategic behavior and information asymmetries.

Introduce profit-sharing mechanisms in multi-year, forward-looking remuneration trajectories. Profit-sharing mechanisms can be enhanced with a menu of contracts that enables distribution utilities to play a role in selecting the strength of the cost-saving incentives to which they will be exposed.

7 This is an elaborate procedure that is explained in **Chapter 5**.

Recommendation 13: Equalize incentives for efficiency in capital and operational expenditures.

Utilities are facing increased trade-offs between traditional capital investments in network assets and novel operational and network management strategies that harness DERs. Equalizing incentives for efficiency across capital expenditures (CAPEX) and operational expenditures (OPEX) is a key step toward giving utilities the flexibility to incorporate novel means of providing network services.

Incentives are typically skewed by conventional regulatory approaches, which add approved capital expenditures directly to a utility's regulated asset base or rate base, while operational expenditures are expensed annually. Even if utilities are properly incentivized to pursue cost savings via a profit-sharing incentive, under traditional remuneration frameworks, saving one dollar of CAPEX will reduce the utility's regulated asset base, thereby reducing the allowed return on equity and net profit for the utility's shareholders. Distribution companies will therefore be fundamentally disincentivized to trading CAPEX for OPEX, including contracting with DERs to defer network investments.

The UK Office of Gas and Electricity Markets has developed a mechanism for equalizing these incentives, known as the total expenditure or "TOTEX-based" approach. Alternative measures have been proposed by the New York Department of Public Service. Whatever mechanism is pursued, the policy objective is to ensure that utilities are free to find the most cost-effective combination of conventional investments and novel operational expenditures (including payments to DERs) to meet demand for network services at desired quality levels.

Recommendation 14: Implement measures to manage inherent uncertainty in utility remuneration and to reduce information asymmetry.

Regulators should start to test and use newly available tools to confront current lack of experience in several areas, including estimating distribution costs under the strong presence of DERs, managing errors in the estimation of relevant distribution network cost drivers when developing multi-year trajectories for remuneration, and addressing heightened information asymmetry between regulators and utilities.[8]

State-of-the-art regulatory tools, including incentive-compatible menus of contracts, tested and reliable engineering-based reference network models, and automatic adjustment factors, should be used to ensure continued cost-efficiency under uncertain future conditions and increased information asymmetry.

Financial incentives related to capital expenditures (CAPEX) and operational expenditures (OPEX) need to be equalized to encourage utilities to find the most efficient combination of expenditures that meet demand for network services at desired quality levels.

8 **Chapter 5** discusses examples of these tools in great detail.

Recommendation 15: Create output-based incentives for advancing explicit regulatory goals.

Other strategies are required to incentivize utilities to move toward critical objectives or outcomes that are unrelated to short-term economic efficiency but are nonetheless important. These include objectives related to commercial quality of service, continuity of electrical supply, voltage quality (which together comprise quality of service), and energy loss reduction. These outcomes are frequently not incentivized by core remuneration frameworks, since achieving improved performance may impose increased investment and operating costs on distribution companies, and most core remuneration frameworks only reward cost reductions. **Chapter 5** discusses best practices in Europe and the United States for each of these topics. In both Europe and the United States, the implementation of outcome-based performance incentives has helped steer network utilities in directions that are beneficial for network customers and that have led to safer, more reliable, and more efficient networks.

Regulators should leverage outcome-based performance incentives to reward utilities for improvements in quality of service or other objectives, such as enhanced resiliency, reduced distribution losses, and improved interconnection times.

Recommendation 16: Create explicit incentives for long-term innovation.

Increased uncertainty about the evolution of network needs, cost drivers, and opportunities will intensify the need for long-term innovation, including expanded investment in demonstration projects, as well as the need for technological learning from such projects and dissemination of that knowledge between network utilities. The objective is for utilities to transform into consistent adopters and integrators of novel solutions.

Chapter 5 has presented case studies of some of the novel approaches used by regulatory authorities in Europe and the United States to create input-based incentives and competitive rewards for the promotion of long-term innovation in electricity distribution networks. These examples could provide guidance to regulators as they consider adopting incentives for innovation.

Incentives for longer-term innovation are needed to accelerate investment in applied R&D and demonstration projects and learning about the capabilities of novel technologies that may have higher risk or longer-term payback periods.

Recommendation 17: Proactively address current and potential future cybersecurity issues.

Most of the recommendations in this report rely on or imply a greater degree of digitalization and interconnectedness in the power system. This sets the stage for our final recommendation concerning distribution system regulation.

Widespread connection of DERs will increase digital complexity and attack surfaces, and therefore requires more intensive cybersecurity protection, with a multi-pronged approach to cybersecurity preparedness. Providing cybersecurity for the electric grid requires developing a risk management culture; understanding the characteristics of baseline or "within-band" operations; rapid sharing of information about cyber threats; and active, skilled, and coordinated teams to detect and respond to anomalous cyber activity, defend against cyber attacks, reduce the "dwell time" of cyber attackers, and implement layered cyber defenses. System operators must have the capacity

Proactive regulation and best practices to address cybersecurity threats are critical as the power system becomes increasingly digitalized.

to operate, maintain, and recover a system that will never be fully protected from cyber attacks. Relevant issues that need to be addressed include advanced cybersecurity technologies, machine-to-machine information sharing, cloud security, regulation and implementation of best practices to avoid prolonged outages and increase system resilience, and international approaches to cybersecurity.

9.2.2 Establishing an effective structure for the electricity industry

The growth and future potential of a variety of distributed energy resources and more price-responsive and flexible electricity demand has sparked a new wave of debate over the structure of the electric power sector, this time focused on the role of distribution network owners and operators as well as end consumers, aggregators, retailers, and new DER business models. This new and complex situation has strong parallels to the introduction of competition and new actors at the bulk power system level during the last three decades in many countries. Useful insights can be drawn from this experience regarding the best structure — for each specific context —for integrating DERs in the provision of electricity services at the distribution level. The goal of these recommendations is to minimize potential conflicts of interest and ensure an efficient, well-functioning electricity sector.

Recommendation 18: Establish independence between the distribution system operator and any agents that perform activities in competitive markets.

Market platforms, network providers, and system operators perform the critical functions that sit at the center of all transactions in electricity markets. Properly assigning responsibilities for these core functions is thus critical to promoting an efficient, well-functioning electricity sector and establishing a level playing field for the competitive provision of electricity services by traditional generators, network providers, and DERs.

The restructuring of wholesale electricity markets demonstrated that establishing a market platform alone is insufficient to ensure competitive electricity generation and supply. In practice, both the system operator and network provider functions can significantly affect the ability of market agents to buy or sell electricity services. Previous attempts at functional and legal unbundling in the bulk power system (i.e., among wholesale generators and transmission system operators) have generally proven insufficient at enabling effective competition. Of course there are differences between the wholesale and transmission context and the retail and distribution context that must be accounted for when making recommendations at the level of distribution networks and retail markets.

We find that the best solution, from a market efficiency perspective, is structural reform that establishes financial independence between the distribution system operator (DSO) and any agents performing activities in competitive markets, including adjacent wholesale generation and ancillary services markets and competitive retail supply and DER markets within the DSO's service territory.

As experience with restructuring in the bulk power system has demonstrated, structural reform that establishes financial independence between distribution system operation and planning functions and competitive market activities would be preferable from the perspective of economic efficiency. It would also facilitate more light-handed regulation.

Recommendation 19: Legal or functional independence requires significant regulatory oversight and transparent mechanisms for provision of services.

Successful experiences in the implementation of competitive wholesale markets indicate that there are two possible structural options — with some variants — for ensuring proper independence and efficiency: (1) a combined distribution network owner/system operator with financial independence from agents that perform competitive activities, and (2) an independent distribution system operator. Only the former model has so far proven practically viable at the distribution level.[9]

9 Their pros and cons are discussed in detail in **Chapter 6**.

Effective regulatory oversight for network operation and planning and transparent mechanisms for the provision of distribution system services are critical to prevent conflicts of interest.

Regulators must contend with the implications of the existing industry structure and the costs of any proposed transition in each context. Diverse conditions exist in various jurisdictions, and the objective of making the DSO independent must be considered alongside the industry restructuring implications of this strategy in every regulatory jurisdiction and power sector setting. As a second-best alternative, various forms of legal and functional independence can be established. These structures will need to be complemented by transparent mechanisms (e.g., auctions or markets) for selecting services where DERs and centralized network services might compete to ensure that no conflicts of interest are being exercised.

Facilitating a level playing field between DERs and conventional approaches to generation and network services is likely to remain challenging if distribution network utilities are vertically integrated into competitive market segments — in that case, significant regulatory oversight will be required.

Recommendation 20: Carefully review and implement best practices in the management of power sector data.

To facilitate level-playing-field competition between aggregators (including retailers), DER providers, and a diversity of competitive agents active within distribution systems, a fourth core function is becoming increasingly important: that of data platform or data hub.

Experience in retail markets in Europe and elsewhere has demonstrated that all market participants need equal and non-discriminatory access to a degree of customer information sufficient to facilitate a level playing field for competition. Likewise, timely and non-discriminatory access to data on network conditions and operation and planning decisions, as well as information on network customers, could be important for facilitating competition among DER service providers and aggregators.

As the increasing connectivity of electric and telecommunications devices generates ever larger quantities of data, more attention needs to be given to who has what rights to use the data, and for what purposes. Utilities, DSOs, data hubs, and data managers have responsibility to act as data stewards to protect customer privacy and the security of the information to which they have unique access, and to abide by government regulations and good practices. Examples of such practices include the European General Data Protection Regulation and the US Federal Trade Commission's Fair Information Practices. Privacy and security safeguards also need to be implemented by third parties that provide analytics and have access to raw power sector data.

Regulators should review and carefully assign responsibilities for data management, while considering multiple goals, including non-discrimination, efficiency, and simplicity. Data on customer usage, telemetry data on network operation and constraints, and other relevant information must be securely stored and made available in a non-discriminatory, timely manner to registered market participants. Consumers must be provided with timely and useful access to data on their own use of electricity services. Data privacy, data protection, and data rights will require increasing attention.

Economies of scope between metering, system operation, and data access argue for combining responsibility for the data hub or data management function with the distribution system operator function. This once again places the focus on ensuring the effective independence of the DSO. If this independence cannot be guaranteed, an independent data hub agency is a second-best option. If the DSO is independent of other competitive agents, the DSO can act as a neutral manager of the data hub. To the degree DSOs are integrated with competitive market segments, the importance of the data hub responsibility argues for further enhancement of functional independence or the establishment of an independent data hub manager. Decisions about the governance of the data hub are thus strongly related to the other structural choices discussed above. Furthermore, the possibility cannot be ruled out that future ICT developments might render a centralized data hub unecessary, while making compatible the selective preservation of privacy, non-discriminatory access to data of commercial value, and DSO access to data needed to perform security functions.

9.2.3 Improving wholesale market design

Wholesale market design should be improved to remove unnecessary barriers that impede the full participation of all resources — including DERs — or distort their competitive performance. Market design improvements are also needed to acknowledge the value of flexibility and to create a level playing field for all technologies.

Recommendation 21: Expand price signals to the intraday timeframe.

The first key step toward improving wholesale market design is to create a more flexible short-term market sequence, adapted to the different planning horizons of all available resources and capable of producing intraday price signals that reflect the operation of the power system between day-ahead markets and the time of electricity delivery.

The sequence of short-term markets present in today's power systems was strongly influenced by the operational procedures of centralized generating plants prior to market liberalization. Today, the day-ahead time horizon is no longer adequate for key market functions given the presence of price-responsive demand, intermittent energy resources, and more constrained operation of thermal generators, among other trends. Therefore, price signals are also needed during the intraday timeframe. In the current US context, with independent system operators (ISOs), we argue for an alternative settlement system that produces intraday price signals (without necessarily implementing intraday markets, as in the European Union). These price signals are necessary to create incentives for efficiently forecasting and managing increasingly variable generation and demand patterns, and should reach all agents in the market.

Power markets should include intraday pricing mechanisms to reward flexible resources and improve forecasting and control of variable renewable resources and electricity demand.

Recommendation 22: Enhance market liquidity and transparency by using more centralized and discrete (rather than continuous) market mechanisms.

Removing information asymmetries and opportunities for market power abuse is necessary to allow small new entrants to compete on a level playing field and thereby obtain the most efficient use of all resources. Current power markets present numerous instances where large market players have a significant advantage over smaller ones. We recommend concentrating market transactions in centralized market sessions (which favors liquidity and reveals market participants' information) instead of relying on bilateral arrangements. Reducing or eliminating the use of bilateral mechanisms is more critical as markets approach real time and the number of market participants shrinks. Where intraday markets exist, we find that a sequence of discrete intraday auctions is superior to continuous trading for the same reasons mentioned above.

Closer to real time, all resources should participate in balancing and real-time markets for the management of their energy imbalances. Most system operators allow large market players to self-manage their imbalances (netting positive and negative deviations of several power plants within the same portfolio) and settle only their net deviations in the balancing market. Liquidity and transparency in balancing markets would improve if this kind of aggregation were avoided.

Recommendation 23: Adapt auction rules to incorporate new operational constraints and new resources.

In addition to the constraints that are imposed by the network, electricity generators are subject to internal physical constraints (e.g., start-up, minimum power output, ramping constraints) that have a significant impact on their operation and cost. New resources (including demand-side resources) in electricity markets will introduce their own operational constraints, which will also need to be addressed. The design of new bidding formats will be a key element in allowing market agents to include information about their costs and operating constraints. Present electricity market rules have been designed for traditional resources and large generators that can manage large portfolios, and current bidding formats are not suitable for some new technologies, such as small-scale storage and demand response. One approach to deal with this problem in the European Union is to continue increasing the complexity of existing bidding formats, but the scalability of this solution is questionable. A better alternative is the US ISO approach in which customized bidding formats are created for each type of resource that capture well their various constraints. In ISO markets, where multi-part bidding formats are already in place, it will be necessary to design new bidding formats (or significantly enhance existing ones) to accommodate new resources.

Market clearing and pricing mechanisms will become more complex as refinements in bidding formats are introduced. An outstanding debate is whether discriminatory side payments (uplifts) should be used in addition to uniform market prices, and if so, how to allocate their cost. Different approaches create significantly different price signals, which affect both the short- and long-term decisions of increasingly responsive market agents. To avoid discrimination against demand resources, the allocation of side payments (when they cannot be avoided) should be incorporated in the market-clearing procedure in a way that ensures revenue sufficiency for all market agents (whether on the generation side or the demand side).

> *Market operators should adapt the definition of services and the bidding formats, with an adequate spatial and temporal granularity, to incorporate new operational constraints and new resources.*

Recommendation 24: Align reserve markets with energy markets and new flexibility requirements and capabilities.

An efficient definition and computation of prices for operating reserves is key to encourage efficient operational and investment decisions. Reserve products are currently tailored to the characteristics of conventional generating technologies. It is important to remove inefficient barriers in current product definitions (e.g., reduce minimum bid sizes and allow aggregation) and to create innovative reserve products that value different capabilities. The objective of creating new reserve products should be to provide value to the power system as a whole. New reserve products should accommodate the diversity of new technologies, rather than creating a new product for each individual new technology.

The connection between energy and reserves to reflect scarcity situations is an essential component of market design to promote efficient operation and investment decisions. Properly defined reserve requirements can help reveal the value of flexibility. In most markets, the quantity of required reserves is fixed, regardless of the cost of those reserves or the value of additional reserves to the system. Instead, these reserve requirements should reflect the changing value of reserves to the system (i.e., the reserve market price) and should be periodically reexamined. The connection between reserves and energy prices should either be enforced explicitly in energy auctions or at least be facilitated by closely aligning the timelines of reserves and energy markets.

Remuneration of reserves must be integrated with or connected to energy markets to reflect the value of reserves and account for the performance of different resources in providing reserves. This integration is particularly important during scarcity events.

Recommendation 25: Enable all resources to participate in long-term capacity markets.

Efficient pricing of energy and reserves may still not encourage adequate levels of investment in firm capacity resources, in the eyes of policy makers or regulators. These regulators or policy makers may institute long-term markets to procure firm capacity resources. If this is the case, non-conventional technologies should be allowed to participate on the same terms as traditional resources.

Capacity mechanisms have a two-fold objective: to provide a long-term hedge for investors and to ensure security of supply for consumers. Demand response, storage, and intermittent renewable generation can contribute in various degrees to solving the security of supply problem and should therefore be eligible to compete to provide capacity services. The key challenge is to ensure that different resources compete on equal footing. To the best of DER capabilities, these resources should be subject to the same conditions as any other technology participating in the capacity mechanism. Although technology neutrality is often argued as a necessity, long-term capacity mechanisms cannot be "technology neutral" if only one product is auctioned or directly remunerated. The solution lies in the definition of the products procured through capacity mechanisms, along with potential penalties for underperformance, where these definitions should adequately value and compare the contribution of each resource to supply electricity in periods of scarcity. If coherently designed, the implementation of capacity mechanisms is fully compatible with other support schemes.

Non-conventional technologies (renewables, storage, demand response, etc.), whether distributed or not, should be integrated in capacity mechanisms on an equal footing with other technologies.

Recommendation 26: Support schemes for clean technologies should be designed to minimize the distortion of efficient prices and signals.

Regulators and policy makers may want to explicitly incentivize the installation of certain resources — for example, renewable electricity generation. This should be done in a manner that minimally distorts the efficient price signals received by market participants. Support schemes that distort the marginal price of energy can create many challenges, for instance, by exposing renewable producers to distorted short-term prices and incentivizing inefficient operating decisions that have a market-wide effect, such as artificially induced negative prices.

One effective method for minimizing distortions is to introduce a capacity-based support mechanism that reflects an estimate of the support needed for reference facilities. The regulator would assign any new plant coming into operation to a particular reference facility that corresponds to the same technology type (wind, solar, etc.), similar commissioning date, and similar project size. The investment cost of the reference facility is known and its market revenues and operating costs (only applicable to fuel-based renewables like biomass) can also be determined. The support payments would then be calculated as the difference between the reference facility's market revenues and the annuity of its investment cost, including a reasonable rate of return. This approach can be combined with auctions for competitive price discovery.[10]

Support mechanisms for DERs or other resources (e.g., centralized solar PV) should be implemented in a fashion that minimally distorts short-term pricing in electricity markets. Auctioned capacity-based support payments that make use of reference facilities to estimate market revenues are an excellent example of such a mechanism.

10 See **Chapter 7** for a detailed description.

9.3 Understanding the Value of DERs: Implications for Policy-making and Regulatory Decisions

Our remaining recommendations focus on understanding how DERs create value, and provide insights about the factors that are most likely to determine the portfolio of cost-effective resources, both centralized and distributed, deployed in different power systems. These insights hold important implications for policy-making and regulatory decisions, as highlighted in the following recommendations.

Recommendation 27: Focus on the locational value of services provided by DERs, rather than on the value of specific resources.

DERs can provide a range of services, which can have locational and non-locational value. Due to their distributed nature, DERs have the potential to be sited and operated to provide services in areas of the power system where these services are most valuable. Understanding the specific services that have locational value — and the range of different resources that can compete to provide these services — should thus guide and inform policy and regulatory decisions.

Due to their distributed nature, DERs have the potential to be sited and operated to provide services in areas of the power system where these services are most valuable. Understanding the specific services that have locational value is thus critical to identifying ways that DERs can create value and should guide policy and regulatory decisions.

Recommendation 28: There is no single "value of solar" or "value of storage" or "value of DERs" to distribution networks. Policy incentives and tariff designs should accurately reflect variation in the marginal value of services provided by DERs.

The value of certain electricity services can differ by orders of magnitude within the same wholesale market or transmission system, and even within a given distribution network. Just as there is no single value for the services that DERs provide, there is no single value of DERs. Indeed, the value that DERs can provide is extremely context-dependent. Additionally, the value of many services can decline quickly as more resources are deployed or activated to supply that service in a particular location and time. This dramatic difference in locational value reinforces the importance of a system of prices and charges for electricity services with sufficient locational granularity. In addition, the variation in locational value makes it clear that it makes little sense to define a single "value of solar" or "value of storage" or value of any other resource. The value of each DER depends on the value of the specific services it provides at a specific time and a specific location. To accurately value the services provided by DERs, prices, regulated charges, and other incentives must therefore reflect the marginal value of these services to the greatest extent practical.

The total locational value of the services provided by distributed solar PV and other DERs can vary by an order of magnitude within the same power system. This dramatic difference in locational value reinforces the importance of a system of prices and charges for electricity services with sufficient locational granularity and renders efforts to define a single value for any DER impractical.

Recommendation 29: Don't neglect the potential to unlock greater value from existing resources.

In many cases, unlocking the potential of existing resources — such as flexible demand, smarter electric vehicle charging, and power electronics or distributed generation that is already installed — can often be more cost-effective than deploying new DERs, conventional generators, or transmission or distribution network assets. Policy and regulatory decisions should therefore avoid skewing incentives toward the deployment of new resources and should aim to create a level playing field and appropriate incentives for existing resources and flexible load to compete in the provision of electricity services. Understanding the potential costs and benefits of existing resources requires careful consideration of (1) the temporal and locational value of the services they may deliver, (2) the opportunity costs and transaction costs associated with operating these resources to provide electricity services, and (3) the initial activation costs required to engage these assets in the efficient provision or consumption of electricity services.

Unlocking the contribution of existing resources — such as flexible demand, electric vehicles, power electronics, or distributed generation that is already deployed — can be an efficient alternative to investing in electricity generation and network capacity.

Recommendation 30: Economies of scale still matter, and "distributed opportunity costs" should be avoided.

Many technologies suitable for distributed deployment, including solar PV, electrochemical energy storage, and fuel cells, can be deployed across a range of scales. Each of these resources also exhibits varying degrees of economies of unit scale, or declining costs per unit of capacity as the size of the installation increases.

For resources that can be deployed at multiple scales, incremental costs associated with failing to exhaust economies of unit scale can outweigh locational value. This can result in a "distributed opportunity cost," making distributed deployment of these resources inefficient. Trade-offs between the incremental cost and additional locational value associated with deploying distributed resources on a smaller scale must be considered in each context.

Economies of scale matter, even for distributed resources. For resources that can be deployed at multiple scales, trade-offs between locational value and incremental unit costs due to economies of unit scale can help identify the ideal locations and applications for these resources. In many cases, after factoring in both costs and benefits, small-scale deployment of DERs can incur "distributed opportunity costs." Smaller is not always better.

Trade-offs between locational value and incremental unit costs further reinforce the importance of technology-neutral price signals with sufficient locational granularity to incentivize and reward the siting and operation of DERs in locations that deliver the greatest net value to the power system.

Accurate price signals or other incentives are essential to unlock the full value of DERs, put all resources on a level playing field, and facilitate distributed decision-making that internalizes trade-offs between locational value and distributed opportunity costs.

Appendix A

A Description of the Computational Models Used in the Utility of the Future Study

A.1 Introduction

Modeling electricity systems is challenging because of their enormous size and complexity. Electricity systems contain millions of elements with diverse functionalities and capabilities, and distributed energy resources (DERs) significantly add to this complexity. Representing such systems with computer-based models requires striking a balance between computational tractability and preserving key system characteristics. The computational tools used in this study consider the technical and economic characteristics of electricity systems, as well as the interactions among various parts of the system. These tools aim to quantify the economic trade-offs between providing electricity services from centralized versus distributed energy resources and to evaluate the impact of sectoral trends and changes (technology cost declines, changing regulatory environments, new energy and environmental policies, etc.) on the efficient mix of these resources.

These computational tools vary in their modeling capabilities, in the degree to which they use simplifying assumptions to represent networks, and in the geographical scope of the power systems that they analyze. They also have different objectives: For example, some are focused on the effects of resource integration on system operations while others center on network expansion and planning.

Some of the models examine the operation and scheduling of resources from an end-user or system perspective. In these types of models, investment decisions are fixed, and the objective is to determine how resources would be utilized (e.g., when generation units would be scheduled and dispatched, when storage facilities would be charged or discharged, etc.) to minimize end-user costs and/or to minimize total system costs. Other models focus on the investment and planning decisions that would minimize costs in the medium to long term. These types of models tend to use more simplified representations of the operation of the system. In general, a single model cannot represent in full detail the complexity of both operation and investment decisions for both networks and energy resources (this limitation of electricity system models is sometimes referred to as the "short blanket" problem[1]).

Modeling the technical capabilities of DERs in detail is challenging due to the complexity of the underlying physical processes as well as to limitations in computational power. Examples of the former include the thermodynamic performance of thermal storage systems and the degradation characteristics of electrochemical storage technologies. These processes are complex but can have significant impacts on DER costs, on the services that DERs are capable of providing, and on the value that these resources can capture in markets. Furthermore, single DER technologies are often capable of providing multiple electricity services, either one at a time or simultaneously with certain restrictions. Therefore, detailed and sufficiently accurate representations of DER technology characteristics are essential for determining what services might be provided by these technologies and at what costs. The most relevant parameters that affect technology performance can be obtained from detailed, technology-specific models and integrated into less detailed models that represent a variety of DER technologies. These less-detailed models can then be used to estimate how different DERs might compete with one another for the provision of electricity services, a task that would be computationally difficult with more detailed representations. The Utility of the Future study utilizes the Distributed Energy Resources Consumer Adoption Model (DER-CAM)[2] and Demand Response and Distributed Resources Economics (DR DRE) model to analyze the DER technical capabilities and end-user responses to price signals.

1 If you pull a short blanket up to cover your head, you cannot also cover your feet—and vice versa.

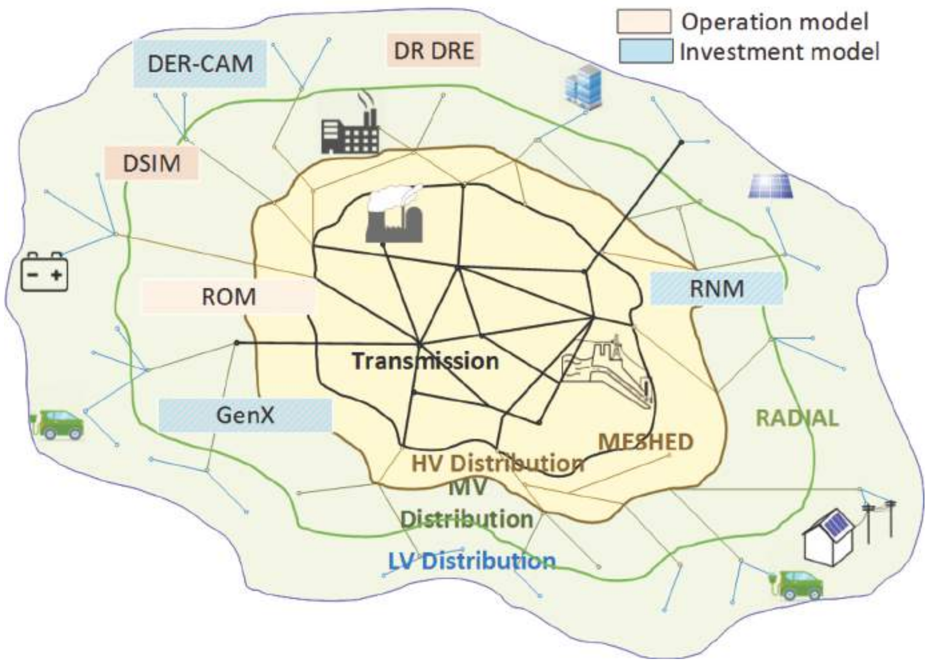
2 DER-CAM was developed by the Lawrence Berkeley National Laboratory. This model has been used in the Utility of the Future study through a collaboration agreement. Further description of this model can be found in building-microgrid.lbl.gov/projects/der-cam.

Electricity network modeling is also a core aspect of the Utility of the Future study as it allows us to quantify the mutual impacts of different elements connected at various parts of the network, either at the level of transmission or of distribution voltage. The network characteristics condition the methodologies used to model network effects. For instance, transmission networks usually have a tractable number of nodes that can be individually included in a unit commitment or a security-constrained economic dispatch model. Since the focus of this study is on distribution, we only use simplified representations of the transmission network in using the Reliability and Operational Model (ROM) and the Optimal Electricity Generation Expansion (GenX) model, which facilitates inclusion of distribution network features at an abstracted but salient level of detail.

Distribution networks have a complex structure with millions of connection points and assets. They have a meshed topology in high voltage, and a radial topology in medium and low voltages. The Reference Network Model (RNM) allowed us to design large-scale networks that resembled what an efficient distribution company would build. We were then able to use those networks to run power flow and optimal power flow simulations. RNM also allowed us to do distribution network expansion planning studies under different load growth and DER penetration scenarios.

By combining the capabilities of RNM, DR DRE, and Matpower, the Distributed Energy Systems Simulator (D-Sim) model is able to calculate the economically rational DER investment and operation decisions made by several thousand network users coordinated by price signals. D-Sim can also calculate the impact of these decisions on the operation of the distribution network.

Figure A.1: Computational Models Represented by Scope and Network Considerations



Given the complexity of incorporating a detailed description of distribution networks into an operational model that also provides economic dispatch and locational marginal prices at the wholesale level, we derive simplified representations for the two models with wider coverage: ROM, which is a detailed operational model that combines day-ahead scheduling and real-time economic dispatch, and GenX, which solves the energy resources expansion problem for centralized and distributed energy sources jointly, considering (in a simplified way) both transmission and distribution network effects. **Figure A.1** shows the different models used in the Utility of the Future study, including what part of the system each model represents and whether the model mostly addresses operational or investment decisions.

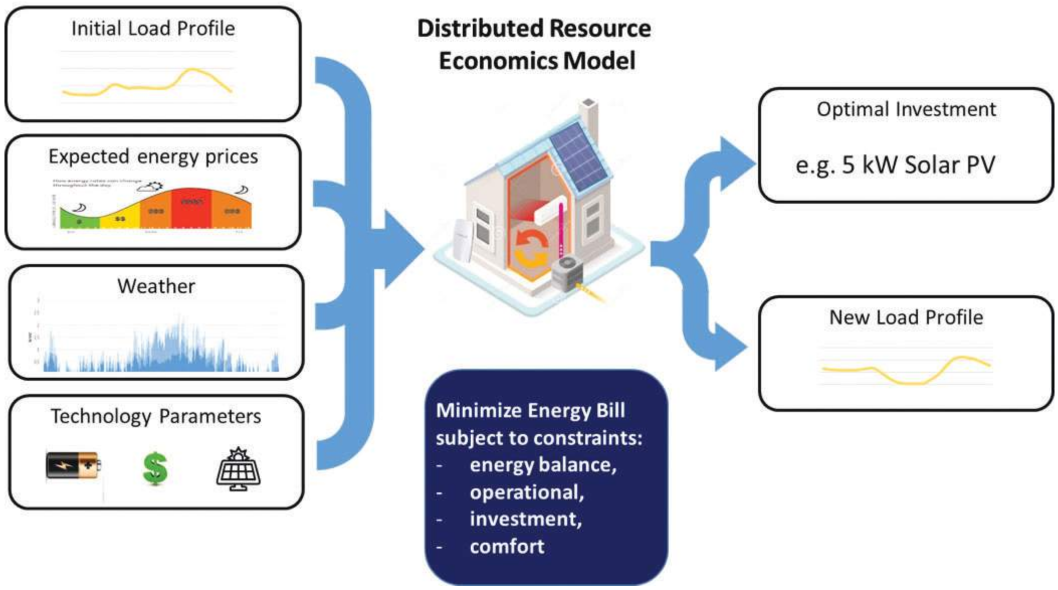
The main features of each model, including the model description, methodology, main inputs and outputs, and the experiments performed in the Utility of the Future study, are provided below.

A.2 Demand Response and Distributed Resources Economics Model

A.2.1 Description

The Demand Response and Distributed Resources Economic (DR DRE) model is designed to optimize the operation of a customer's on-site energy resources and energy purchases in response to economic signals (as a price-taker). It considers as investment options solar photovoltaics (PV) and batteries, as well as flexibility from loads. Economic signals can include hourly energy prices (in dollars per kilowatt hour [$/kWh]), capacity charges (dollars per kilowatt [$/kW]), and fixed (dollar per year or month) components.

Figure A.2: Schematic Representation of Demand Response and Distributed Resources Economics Model



A.2.2 Methodology

The model is a mixed integer linear program coded in the Julia/JuMP programming language (currently using the Gurobi solver). A strength of this language is its open-source nature, its speed, and its flexibility, making the model suitable for many iterations and for computationally intensive problems. DR DRE is designed to optimize a single end-user's consumption and/or production, considering both an operations and an investment point of view, with the purpose of minimizing the overall net cost for the user. The main inputs and outputs of DR DRE are represented in **Section A.2.3**.

A.2.3 Main DR DRE inputs

- **DER technologies and characteristics:** PV capacity and production data (location-specific PVWatts[3] production data are used to scale PV capacity), battery characteristics (energy capacity, power capacity, charge and discharge efficiency, degradation), gas-based combined heat and power (CHP), fuel cell or microturbine characteristics (gas price, heat-to-power ratio, power capacity)
- **Building characteristics:** non-controllable load (e.g., lighting, bathing or showering routines), controllable load (e.g., dishwasher, dryer), thermal loads (e.g., power rating of heat pump, water heater)
- **Rates:** hourly energy rates ($/kWh), use of network rate ($/kW-coincident or $/kWh), firm capacity rate ($/kW), and reserves
- **Weather data:** hourly ambient temperature, insolation data

3 NREL's PVWatts Calculator estimates the energy production and cost of energy of grid-connected photovoltaic (PV) energy systems. It is available at pvwatts.nrel.gov.

A.2.4 Main DR DRE Outputs

- **Customer purchases:** purchases from existing utility/retailer and associated savings
- **Customer investments in DERs**
- **Operating schedule of controllable DERs:** batteries, schedulable loads, and thermal storage

A.2.5 Experiments run in DR DRE for the Utility of the Future study

Experimental designs and case studies were conducted in DR DRE to answer the following questions:

- Under what sets of service pricing structures or technology parameters do DERs complement and compete with each other? (**Chapter 3.2.1**)
- What is the optimal DER deployment at the end-user level, and how do DER economics change under different service pricing structures? (**Chapter 4.5**)
- What are the most sensitive regulatory parameters affecting DER economics? (**Chapter 4.5**)

A.3 Distributed Energy Systems Simulator

A.3.1 Description

The Distributed Energy Systems Simulator (D-Sim) examines the impact of a large number of price-responsive network users on distribution networks under different price signals, using the DR DRE model and iterative optimal power flow calculations.

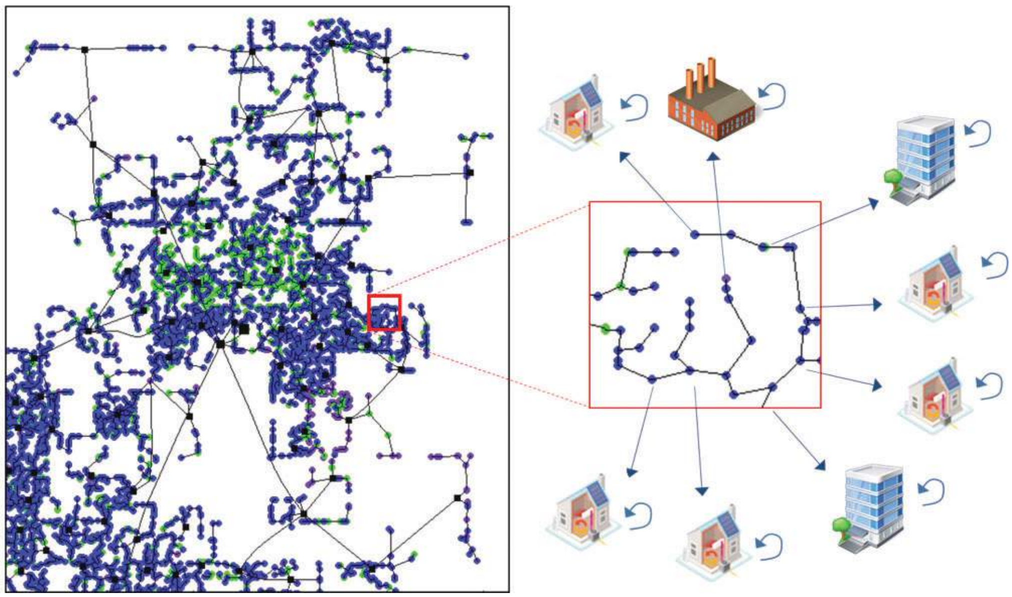
A.3.2 Methodology

The model focuses on the lower voltage levels of distribution networks. Based on distribution network characteristics, the model populates representative buildings with realistic geographic zoning. The simulations use sequential hourly power flow and optimal power flow calculations to compute distributed locational marginal prices for fractions of a year. DR DRE then responds to these prices and computes the optimal response from network users based on real user constraints (e.g., room temperature tolerances and cost of temperature deviations).

The model has three simulation layers:

- **A distribution network** connected to the transmission network through one grid substation. As the three-phase power system is assumed to be balanced, the three-phase circuit is represented by a single line equivalent
- **Distribution network users** connected to the network through a single meter. Each user is represented in detail, and internal dynamics and constraints are both included
- **An external electric power system.** The entire infrastructure upstream of the grid substation is summarized in a single node and modeled as a generator with very large active and reactive power capacity. The cost of energy purchased from the system is calculated from a simplified production cost function

Figure A.3: Distribution System Used in the Distributed Energy Systems Simulator



This example of a distribution system used in the Distributed Energy Systems Simulator includes the distribution network itself and models of price-responsive network users connected to each node. The network users can react to expected future price signals by changing their short-term behavior and by adopting DER technologies.

A.3.3 Main D-Sim inputs

- **A map image and aggregate statistics for the distribution area:** Load density, distribution of network user types, and electrical equipment characteristics are included
- **Definition of network user types:** Responsive users are represented through a combination of manageable load and fixed, active, power-consumption profiles

A.3.4 Main D-Sim outputs

- **Hourly data:** for nodal prices (locational marginal pricing at distribution level), nodal voltages, congestions, and losses per line
- **Yearly data:** for energy payments by network users, network charges by user, total remuneration of the distribution network owner, as well as identification of efficient network reinforcements
- **Optimal DER investments:** from a private point of view (either centrally coordinated or not), assuming that the users are price-takers

A.3.5 Experiments run in D-Sim for the Utility of the Future study

- Competitiveness of DER investments under different price signals (**Chapter 4.5**)
- Sensitivity of the economic performance of a given DER investment given other DER investments and network reinforcements (**Chapter 4.4**)
- Calculation of locational marginal pricing at distribution level and network cost recovery (**Chapter 4.4**)

A.4 Reference Network Model for Large-Scale Network Distribution Planning

A.4.1 Description

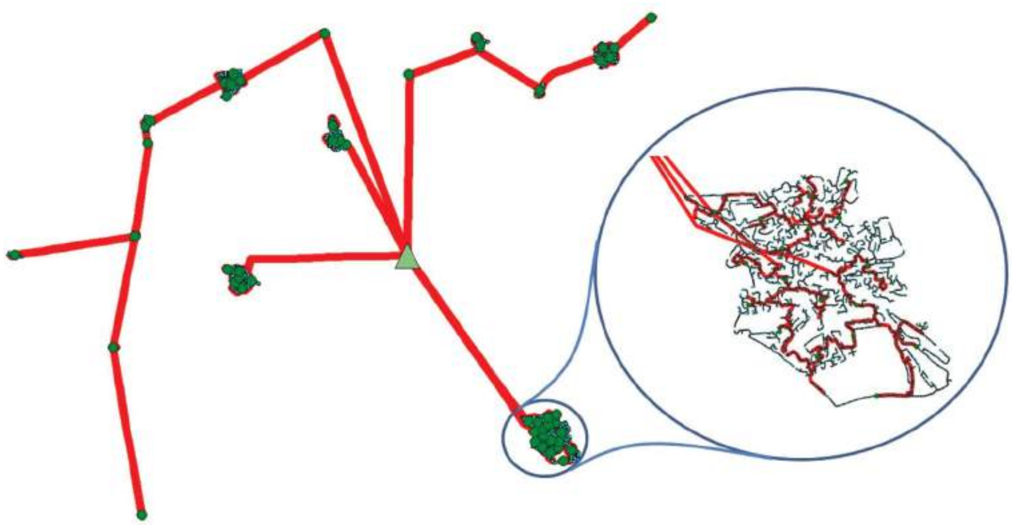
The Reference Network Model (RNM) is a very large-scale planning tool developed by Instituto de Investigación Tecnológica—Comillas (Mateo et al. 2011; Gómez et al. 2013) that plans the electrical distribution network starting from the geographic coordinates and profiles of network users. It designs networks comprising several voltage levels from low to high and includes plans for both substations and feeders. The planning algorithms consider network technical constraints, such as voltage limits, thermal constraints, and continuity of supply targets. RNM also considers geographical constraints such as the street map, topography, and forbidden paths (e.g., through nature reserves or lakes). RNM has been used in different projects to test the effect of demand changes and different penetrations of DERs on distribution networks (MIT 2015; MERGE Project 2011; Vallés et al. 2014).

A.4.2 Methodology

There are two versions of the RNM: Greenfield and Brownfield. The Greenfield RNM builds the network from scratch, starting with network users and moving up to transmission substations. The Brownfield RNM expands and reinforces a pre-existing network to accommodate additional demands or new DERs. Both models have been programmed in C++ and are able to deal with very large-scale networks comprising millions of network users.

Figure A.4 shows a rural network example built in RNM. Thick red lines are medium-voltage (MV) feeders, thin black lines are low-voltage (LV) feeders, the green triangle is the high-voltage (HV)/MV substation, green circles are MV/LV transformers, and small points are LV customers.

Figure A.4: Example of Rural Network Built by the Reference Network Model



A.4.3 Main RNM inputs

- **Customers, DERs, and supply HV substations:** locational data (rectangular coordinates), power demand, installed capacity, and daily load or generation profiles
- **Library of standard installations:** substations, transformers, electrical lines, switching and protection equipment
- **Technical and economic planning parameters:** continuity of supply targets, demand increase rate, loss factors, simultaneity coefficients
- G**eographic data and constraints:** orography, lakes, nature reserves

A.4.4 Main RNM outputs

- **Graphical results:** feeder layout, substation locations, customers, and DER connections
- **Description of the planned network:**
  - » customers and DERs per voltage level
  - » length of LV, MV, and HV networks, including an indication of aerial and underground ratios
  - » total cost of substations and feeders per voltage level broken down into investment cost, preventive and corrective maintenance cost, cost of energy losses, and cost of switching and protection equipment
  - » specifications for infrastructure required in urban areas (costs of ditches, aerial wires, and posts)
  - » continuity of supply indexes (SAIDI, SAIFI)[4] for specific regions
  - » substations, feeders, and protection equipment installed at each voltage level, including both the number of installations and their main technical and economic parameters

4 System Average Interruption Duration Index (SAIDI) and System Average Interruption Frequency Index (SAIFI) are used as reliability indicators by electric power utilities.

A.4.5 Experiments run in RNM for the Utility of the Future study

- Impact of specific DERs (e.g., PV, storage, demand response, and electric vehicles) and a combination of DERs on distribution network costs and investment (**Chapter 8.3.2**)
- Design of network charges considering different networks and categories of network users (**Chapter 4.4.4**)
- Implications of DER penetration on the remuneration of distribution system operators (**Chapter 5.2.2**)

A.5 Reliability and Operational Model with Distribution Network Representation

A.5.1 Description

The Reliability and Operational Model (ROM) simulates the operation of an electric power system, including day-ahead scheduling and real-time economic dispatch. ROM uses an Electric Equivalent Network (EEN) to represent distribution and transmission networks in a simplified manner that accounts for features that affect the operation of the larger power system, such as resistive losses. In this manner, ROM computes the dispatch of energy resources located both at transmission and at distribution networks.

A.5.2 Methodology

An EEN example is represented in **Figure A.5**, where the transmission network consists of two voltage levels and the distribution network has four different voltage levels (other topologies can also be modeled).

Figure A.5: Electric Equivalent Network

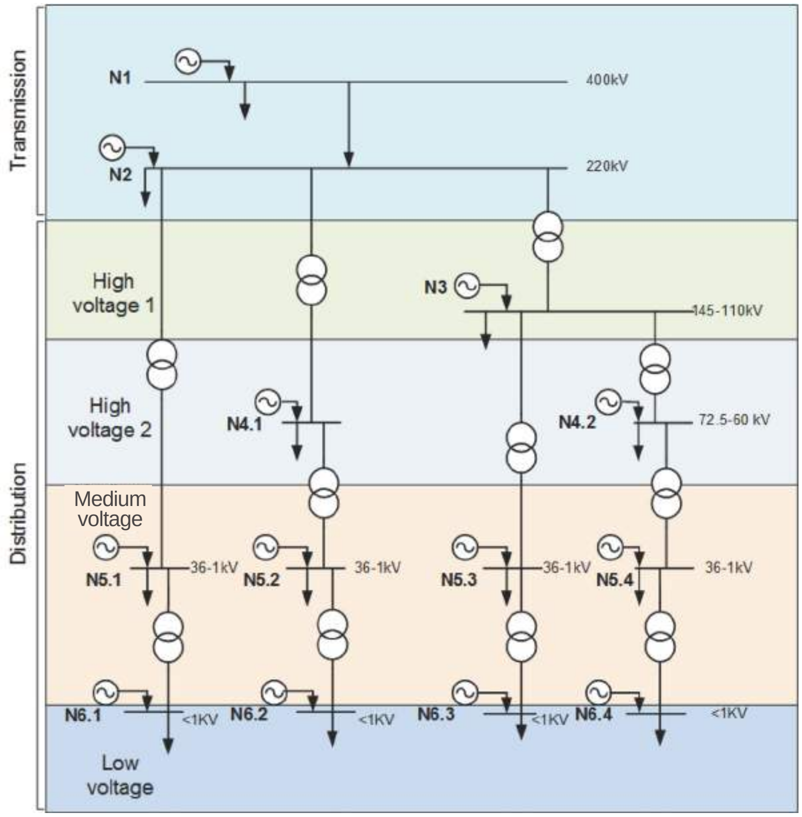


Figure A.5 shows different possible ways to transform voltage levels, reflecting the complexity of real distribution networks. For instance, urban networks usually have direct transformation from high-voltage transmission levels to medium-voltage distribution levels (as shown at the left-hand side of Figure A.5), whereas in rural areas transformation typically happens at several distribution voltage levels (as represented in the right-hand side of Figure A.5). Each branch of the network represented in Figure A.5 can aggregate many thousands of networks with similar topology.

A.5.3 Main EEN inputs

An EEN can be computed based on available measured data from distribution companies. The following data are required at each voltage level: generation, aggregated demand, energy flows, and network energy losses. These data may only be available for specific time periods. For example, measures are usually taken during tariff periods, which occur monthly or even less frequently. Therefore, different samples are needed to compute the EEN parameters.

A.5.4 Computation of equivalent resistances

The EEN represented in **Figure A.5** has a radial configuration. Since parallel flows are nonexistent, and a DC power flow approach is being used, it is not necessary to calculate an equivalent network reactance. In the DC power flow formulation, reactive power flows and voltage drops are also neglected. A quadratic approximation for energy losses is considered for the calculation of the equivalent network. Some of the ROM model results are operation costs, locational marginal prices, and reliability indexes. This model has been employed in research projects funded by the European Commission, such as MERGE (2011) and TWENTIES (2013).

Some of the main characteristics of ROM are:

- a day-ahead unit scheduling optimization followed by a sequential hourly simulation in a one-year period;
- Monte Carlo simulation of many yearly scenarios to deal with the stochasticity of demand and intermittent renewable generation; and
- detailed operation constraints regarding unit commitment and dispatch.

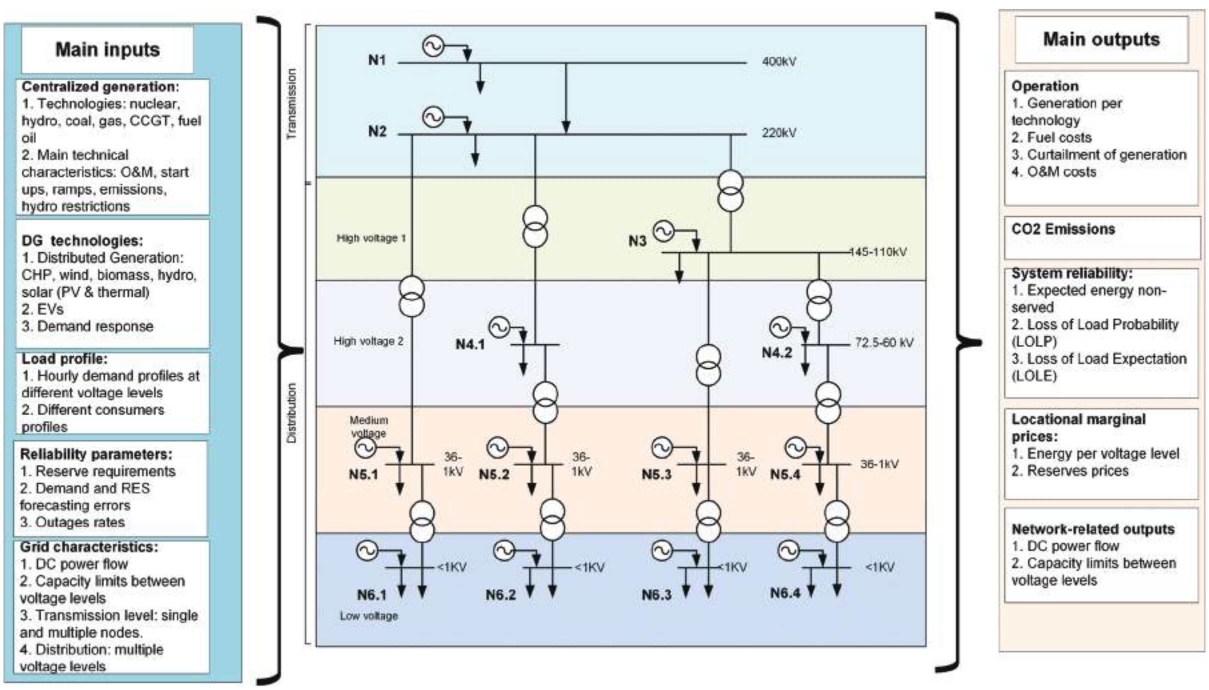
In addition, when management of hydro resources and seasonal pumped storage exceeds the analysis time frame, the hydro operation can be computed by another higher–level model and taken as an input in ROM.

Figure A.6 shows some of the model's inputs and outputs, including the EEN representation.

A.5.5 Experiments run in ROM for the Utility of the Future study

- Effect of energy losses on marginal prices for each voltage level (**Chapter 3.2.3** and **Chapter 4.4.1.3**)
- Effect on locational marginal pricing of a mix of DERs connected at LV—e.g., PV plus storage and backup generation (**Chapter 8.2.1.1**)
- Effect of prices and charges on the dispatch of DERs (e.g., electric vehicles, solar PV, demand response, and backup generation) connected at low voltages (**Chapter 4.5**)

Figure A.6: Summary of the ROM Model



A.6 The GenX Model: Optimal Electricity Generation Expansion Model with Distributed Energy Resources

A.6.1 Description

The Optimal Electricity Generation Expansion (GenX) model represents the expansion of generation capacity while distinguishing between distributed and centralized resources. The formulation includes requirements for operating reserves and expected impacts on transmission and distribution network losses and reinforcement costs. By changing certain parameters, GenX can also simulate the effect of different network tariffs, subsidies, and other policy or regulatory decisions on the equilibrium capacity mix.

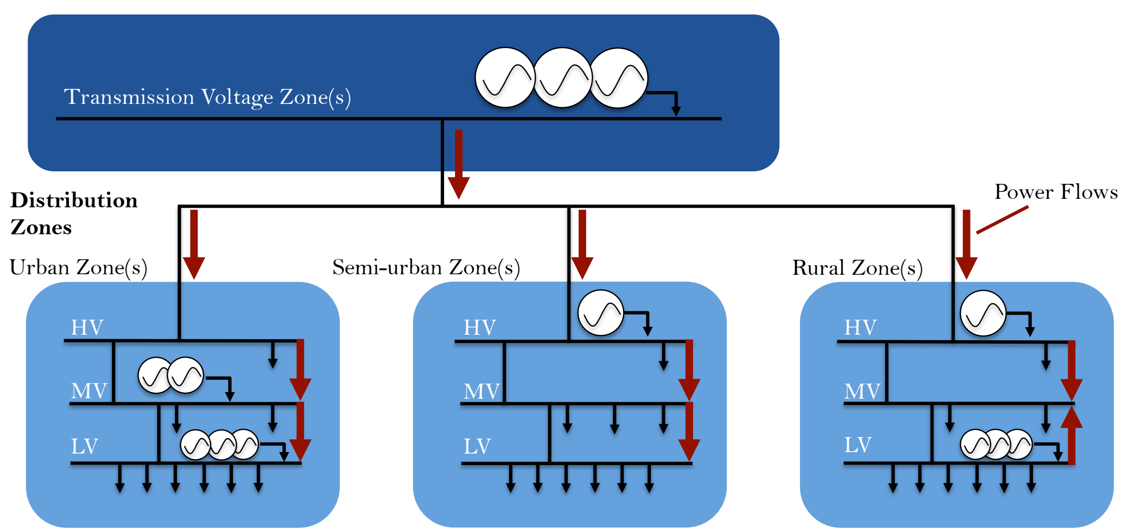
A.6.2 Methodology

At the core of GenX is a state-of-the-art, compact, capacity-expansion formulation that includes a detailed representation of operational constraints. A distinctive feature of the model is its ability to capture the locational value of different investment options through a simplified representation of the transmission and distribution network, which comprises a number of zones.

The capacity of all generating units and DERs is represented as a continuous decision variable, except for large thermal units, which may be represented as integer plant clusters by zone (as per Palmintier [2013]) if desired. Units of incremental capacity of all DERs, large-scale wind and solar, and open-cycle gas turbines (OCGTs) are all small enough that this abstraction is minor, while larger thermal units can be represented as integer clusters if the discrete nature (or lumpiness) of these investment decisions is considered important.

Operational decisions for generating units and DERs are continuous decisions, with the exception of cycling decisions for large thermal units, which can be represented as either continuous decisions or integer decisions (e.g., how many units within each cluster of similar plants to turn on or off) as desired. Integer clustering of similar plants entails the simplifying assumption that all plants within a cluster are identical and that all committed units within a cluster are operating at the same power output level. Treating commitment decisions as continuous variables

Figure A.7: Network Representation by Zones in the GenX Model



further relaxes the problem and allows commitment of fractions of a plant. Both options introduce modest abstraction errors but significantly improve computational performance, enabling greater detail in other features, such as network complexity. Since on/off decisions for individual DERs and even OCGTs are fast and occur in small increments, representing them as continuous decisions is also a minor abstraction.

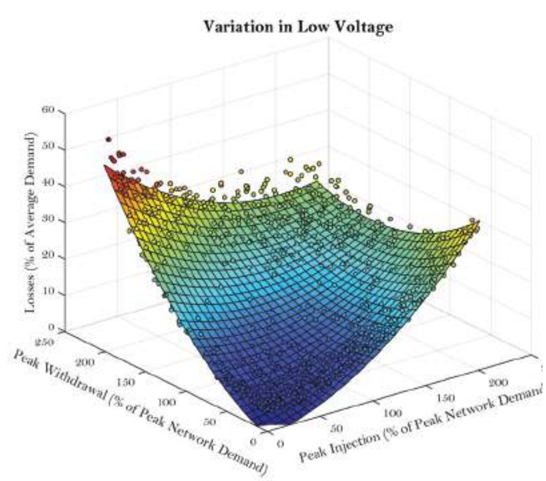
Capacity investment and operational decisions are indexed across each node or zone in the system, enabling the model to select the optimal location of capacity investments and operations in each location. Thus the model balances the different economies of scale at different voltage levels on the one hand, with the differential impacts or benefits of location at different zones or voltage levels on the other hand—a key advantage over other models.

Eligible central station generation resources include: combined cycle gas turbines, open cycle gas combustion turbines, pulverized coal, nuclear, wind, and solar PV. Other possible central station resources include geothermal, biomass, solar thermal, pumped hydro storage, and possible thermal storage for solar thermal or nuclear units. Eligible DERs include: solar PV, electrochemical storage, thermal storage, flexible demands, and batteries.

Power flows between zones and voltage levels are modeled as simple transport flows. Maximum power flows across these interfaces capture key network constraints. Future versions could include nodal network representation in the meshed portion of the system with DC power flow constraints, although this dramatically increases the dimensionality of the problem.

Losses are a function of power flows between voltage levels or zones, implemented as a piecewise linear approximation of quadratic resistive losses. Losses due to power flows within zones are related to injections and withdrawals by means of several hundred power flow simulations in realistic networks, as we show in **Figure A.8** for a semi-urban network.

Figure A.8: Example Results from Multiple Power Flow Simulations



These results taken from multiple power flow simulations show the relationship among losses, injections, and withdrawals within the low- and medium-voltage portions of a semi-urban European distribution network.

Distribution network reinforcement costs associated with changes in peak power injections or withdrawals at each node are represented as linear or piecewise linear functions parameterized by experiments with RNM and optimal power flow modeling.

Reserve requirements are modeled as day-ahead commitments of capacity to regulation and spinning/nonspinning contingency reserves, to capture the commitment of capacity necessary to robustly resolve short-term uncertainty in load and renewable energy forecasts and power plant or transmission network failures.

A.6.3 Experiments run in GenX for the Utility of the Future study

The primary use of the model was to consider the optimal long-run equilibrium of central station generators and DERs under a variety of potential parametric conditions (technology costs, network characteristics, etc.) to explore when and how DERs can compete effectively with conventional resources. The model can also examine the impact of different network charges on the equilibrium capacity mix; the impact of allowing DERs to participate in wholesale ancillary services markets; the role of DERs in meeting carbon dioxide emissions constraints; and other policy or market design-relevant questions.

Outputs, per zone and voltage level, include: capacity investments in each resource, distribution reinforcement costs, operational decisions for each resource, power flows, zonal shadow prices, and revenues of generators and DERs.

References

De Sisternes, F. and M.D. Webster. 2013. "Optimal Selection of Sample Weeks for Approximating the Net Load in Generation Planning Problems." ESD Working Paper Series 2013-03. Cambridge, MA: Massachusetts Institute of Technology. January 2013.

Gómez, T., C. Mateo, A. Sánchez, P. Frías, and R. Cossent. 2013. "Reference Network Models: A Computational Tool for Planning and Designing Large-Scale Smart Electricity Distribution Grids." In S. Khaitan and A. Gupta, eds., *High Performance Computing in Power and Energy Systems.* 247–79. Berlin and Heidelberg, Germany: Springer. link.springer.com/chapter/10.1007/978-3-642-32683-7_9.

Mateo, C., T. Gomez, A. Sánchez-Miralles, J. Peco, and A. Candela. 2011. "A Reference Network Model for Large-Scale Distribution Planning with Automatic Street Map Generation." *IEEE Transactions on Power Systems* 26 (1): 190–97. doi:10.1109/TPWRS.2010.2052077.

MERGE Project. 2011. "Functional Specification for Tools to Assess Steady State and Dynamic Behavior Impacts, Impact on Electricity Markets and Impact on High Penetration of EV on the Reserve Level." www.ev-merge.eu/images/stories/uploads/MERGE_WP2_D2.2.rar.

MIT. 2015. *The Future of Solar Energy*. Cambridge, MA: Massachusetts Institute of Technology. mitei.mit.edu/system/files/MIT%20Future%20of%20Solar%20Energy%20Study_compressed.pdf.

Nahmmacher, P., E. Schmid, L. Hirth, and B. Knopf. 2014. "Carpe Diem: A Novel Approach to Select Representative Days for Long-Term Power System Models with High Shares of Renewable Energy Sources." SSRN Scholarly Paper ID 2537072. Rochester, NY: Social Science Research Network. papers.ssrn.com/abstract=2537072.

Palmintier, B.S. 2013. "Incorporating Operational Flexibility into Electric Generation Planning: Impacts and Methods for System Design and Policy Analysis." Cambridge, MA: Massachusetts Institute of Technology. dspace.mit.edu/handle/1721.1/79147.

Silverstein, A. 2011. "Transmission 101." Presented at the NCEP Transmission Technologies Workshop. www.naruc.org/grants/Documents/Silverstein%20NCEP%20T-101%200420111.pdf.

Tabors, R., A. Rudkevich, and R. Hornby. 2014. "Rate Impact on LIPA Resident and Commercial Customers of 250MW Offshore Wind Development on Eastern Long Island." www.aertc.org/docs/Rate%20Impact%20LIPA%20Cust.pdf.

TWENTIES Project. 2013. *TWENTIES Project: Final Report*. www.twenties-project.eu/system/files/Twenties%20final%20report_web_v2.pdf.

Vallés, M., P. Frías, C. Mateo, R. Cossent, and J. Reneses. 2014. "Economic Benefits of Active Demand for Stakeholders." Project: ADVANCED/WP6 – T6.3/D6.3. Funded by European Commission within "FP7 - Cooperation/Energy."

Van den Bergh, K., E. Delarue, and W. D'haeseleer. 2014. "DC Power Flow in Unit Commitment Models." TME Working Paper-Energy and Environment. www.mech.kuleuven.be/en/tme/research/energy_environment/Pdf/wpen2014-12.pdf.

Van Hertem, D., J. Verboomen, K. Purchala, R. Belmans, and W.L. Kling. 2006. "Usefulness of DC Power Flow for Active Power Flow Analysis with Flow Controlling Devices." The 8th IEE International Conference on AC and DC Power Transmission. London, United Kingdom: ACDC 2006, 58–62.

Appendix B

A Review of Business Models for Distributed Energy Resources

B.1 Introduction

The electric utility business model is in a state of profound transition. A 2013 survey found that 94 percent of the senior power and utility executives surveyed "predict complete transformation or important changes to the power utility business model" by 2030 (PWC 2013). These changes are being driven primarily by the influx of distributed energy resources (DERs)—including solar photovoltaics and other distributed generation, thermal and electrical energy storage, and more flexible and price-responsive management of electricity demand—and information and communication technologies (ICTs). Many predict that the changes driven by DERs will be highly disruptive to the electricity sector and that, without adaptation, incumbent utilities[1] risk falling into a "death spiral" that threatens their financial viability (PWC 2015; Kind 2014).

1 In the US context, "utility" typically refers to the distribution system owner/operator, whether in a traditional or restructured environment. In European or other contexts, the term "utility" is often interpreted more broadly and refers to generators, network companies, and other power sector firms involved in the supply of electricity. We adopt a broad definition of the utility and use the term to describe any company engaging in the provision of electricity services.

Electricity infrastructure is considered uniquely critical due to its role as an enabler of other economic functions and sectors (Office of the Press Secretary 2013). The financial stability of electric utilities is key to the effective management, maintenance, and expansion of the trillions of dollars of global electricity assets that are critical (Kind 2014). Further, a well-crafted business model will have important impacts on the financial performance of a firm (IBM Global Business Services 2006; Zott and Amit 2007). Understanding the business models that are emerging in the power sector is therefore important, not only to incumbent utilities and new market entrants, but also to the public at large.

To shed light on this discussion, this appendix presents a novel empirical review and analysis of the business models for three of the most widely deployed DERs: solar photovoltaics, electricity and thermal storage, and demand response. We define the key "value capture" and "value creation" components of 144 distributed energy business models.[2] We use an ontological approach[3] similar to that of Osterwalder and Pigneur (2010) to define distributed energy business models. We also create a structured framework with which to analyze and classify distributed energy business models (Richter 2012, 2013).

For each business in our data set, we define the electricity services provided, revenue streams captured by the provision of these services, customers targeted, and key DERs used. We use data about electricity services, revenue streams, customer segments, and key DERs to define a small set of business model "archetypes" that describe common classes of many business models. While differences exist among the business models in each archetype, each archetype shares a common set of features. For each archetype, concrete examples of active business models are provided.

This appendix proceeds as follows. First, we provide a brief review of the current literature on utility business models. Second, we present an overview of our data and methodology. Third and finally, we define business model archetypes for the three largest DER categories: demand response (DR) and energy management systems (EMS); electrical and thermal storage; and solar photovoltaics (PV). We also describe some of the interesting nuances that exist within each archetype.

B.2 Literature Review

Little academic literature describing utility business models exists. Of the publications that do exist in the academic, trade, and industry analyst literature, few are as rigorous as the present study. First, paradoxically, studies of business models often do not define either the utility business model or a business model more broadly[4] (Newcomb et al. 2013; Lehr 2013; Rocky Mountain Institute 2013). Second, many studies define and explore a single business model or a small set of business models associated with a single technology without exploring how these models may be competitively positioned against other business models (Huijben and Verbong 2013; Weiller et al. 2015; Bell et al. 2014; Ruester et al. 2012; Behrangrad 2015). Finally, a number of studies perform analyses of a technology providing a limited set of electricity services, without exploring the full range of services that the technology is providing or may provide (Huijben and Verbong 2013; Weiller et al. 2015). Traditional engineering- or economics-driven business model analyses tend to assume that business models are superfluous, since in those studies suppliers capture economic rents through the sale of services at competitive, market-based rates. These studies assume that if value exists, suppliers will automatically deliver it, and customers will automatically pay for it—when in reality, market perfections often "conceal" value that can only be uncovered through business models.

Only a small subset of business model studies have analyzed utility business models using an ontological approach that is similar to the one presented in this appendix, and none has done so using quantitative empirical methods. Several of these studies focus on a subset of business models that utilize a particular technology (e.g., Schoettl and Lehmann-Ortega 2011; Okkonen and Suhonen 2010). Richter (2012, 2013) uses case studies and surveys in combination with an ontological approach to develop an understanding of utility business models that employ a variety of renewable energy technologies. This appendix builds upon the existing literature by using a data-driven approach to circumscribe and glean insights from the current distributed-energy business-model landscape.

B.3 Data and Methodology

Our analysis includes a sample of 144 regionally diverse companies whose core business operations are associated with one or more of three DER technology categories—demand response (DR) and energy management systems (EMS); electrical and thermal storage; and solar PV. Many of the companies in our sample heavily rely on information and communication technologies (ICTs) to enable communication and control

2 "Value capture" is the means by which a business monetizes its product or service. "Value creation" is the means by which a business generates welfare for its customer(s).

3 An "ontology" is defined as an explicit specification of a conceptualization (Gruber 1993). It can be understood as a formal description of the concepts and relationships in a specific domain (in our case, business model research).

4 Many of the early authors of business model literature failed to define the term. Of the studies surveyed by Zott et al. 2011, 37 percent did not promulgate a definition of a business model, "taking its meaning more or less for granted."

of the DER resource of interest. However, given their ubiquitous nature, we do not include ICTs as a standalone category in this analysis.

Data for the companies used in this analysis were collected from publicly available news, academic, and industry publications between February 2014 and October 2015. In addition, to ensure that the sample was representative of the "universe" of DER business models that exist today, we sampled from the Cleantech Group's i3 database, a commercial database that contains information on more than 24,000 "clean tech," DER, and sustainability-focused businesses (i3 Connect n.d.). We created three sets of companies from the i3 database—one for each of the DER technology categories. The i3 database categorizes businesses by core focus; we used this feature to create sets of all business models that were categorized as "ground-mounted PV" and "rooftop PV" (which together comprised our solar PV set); "grid energy storage"[5]; and "demand response." We then drew stratified random samples from each set, such that the distributions of companies in our final sample were similar to those of the i3 database in terms of company headquarter region and founding year. We sampled 50 companies in each of our three DER technology categories. A small number (six) of the companies did not fit our coding criteria, and thus were not included in our final sample.

Tables B.1 and **B.2** show the number of companies in our sample in each founding-year bracket and in each region. These percentages are compared with the percentages that are found in the larger i3 database. As Tables B.1 and B.2 show, the distribution of companies in our sample deviated from the relevant distribution of companies in the i3 sample by no more than 5 percent.

Table B.1: Founding Year of Companies in Sample

YEAR FOUNDED	DEMAND RESPONSE AND ENERGY MANAGEMENT SYSTEMS			ELECTRICAL AND THERMAL STORAGE			SOLAR PV		
	Sample Number	Sample Percent	i3 Percent	Sample Number	Sample Percent	i3 Percent	Sample Number	Sample Percent	i3 Percent
1990 or Earlier	4	9%	7%	3	6%	6%	6	12%	9%
1991-1995	1	2%	3%	1	2%	2%	2	4%	5%
1996-2000	6	13%	12%	3	6%	6%	4	8%	8%
2001-2005	9	20%	20%	5	10%	11%	9	18%	19%
2006-2010	20	43%	46%	22	47%	47%	23	46%	47%
2011-2015	6	13%	13%	14	29%	28%	6	12%	12%
Total	46	100%	100%	48	100%	100%	50	100%	100%

5 Note that despite the name, this category also includes "behind-the-meter" energy storage companies, as discussed in Section.

Table B.2: Headquarter Region of Companies in Sample

REGION	DEMAND RESPONSE AND ENERGY MANAGEMENT SYSTEMS			ELECTRICAL AND THERMAL STORAGE			SOLAR PV		
	Sample Number	Sample Percent	I3 Percent	Sample Number	Sample Percent	I3 Percent	Sample Number	Sample Percent	I3 Percent
Africa	0	0%	0%	0	0%	1%	1	2%	1%
Asia Pacific	0	0%	4%	4	8%	8%	4	8%	9%
Central and South America	0	0%	0%	0	0%	0%	0	0%	0%
Europe and Israel	10	22%	22%	12	25%	24%	19	38%	38%
Middle East	0	0%	0%	0	0%	0%	0	0%	1%
North America	36	78%	74%	32	67%	67%	26	52%	52%
Total	46	100%	100%	48	100%	100%	50	100%	100%

B.4 Business Models for Demand Response and Energy Management Systems

Figure B.1 depicts the DR and EMS business model landscape. The figure shows the customers targeted (horizontal axis), services provided (vertical axis), and revenue streams leveraged (color) by business models. The size of each circle represents the number of business models within a given category.

From Figure B.1, we can identify three major business model clusters, or archetypes, that share similar characteristics (i.e., that target similar customers and provide similar services). A brief description of each of those business model archetypes follows.

B.4.1 Market-based capacity and reserve demand response

The majority of DR and EMS business models have emerged in restructured power markets (Cappers et al. 2010; Hurley et al. 2010; Shariatzadeh et al. 2015). The market rules that determine the exact structure of products procured from DR business models vary from market to market (Behrangrad 2015; Shariatzadeh et al. 2015). However, a number of common themes can be found among business models that fall within an archetype that we have defined as "market-based capacity and reserve DR." The generic structure of this business model archetype is depicted in **Figure B.2**.

Figure B.1: Business Model Taxonomy for Demand Response and Energy Management Systems

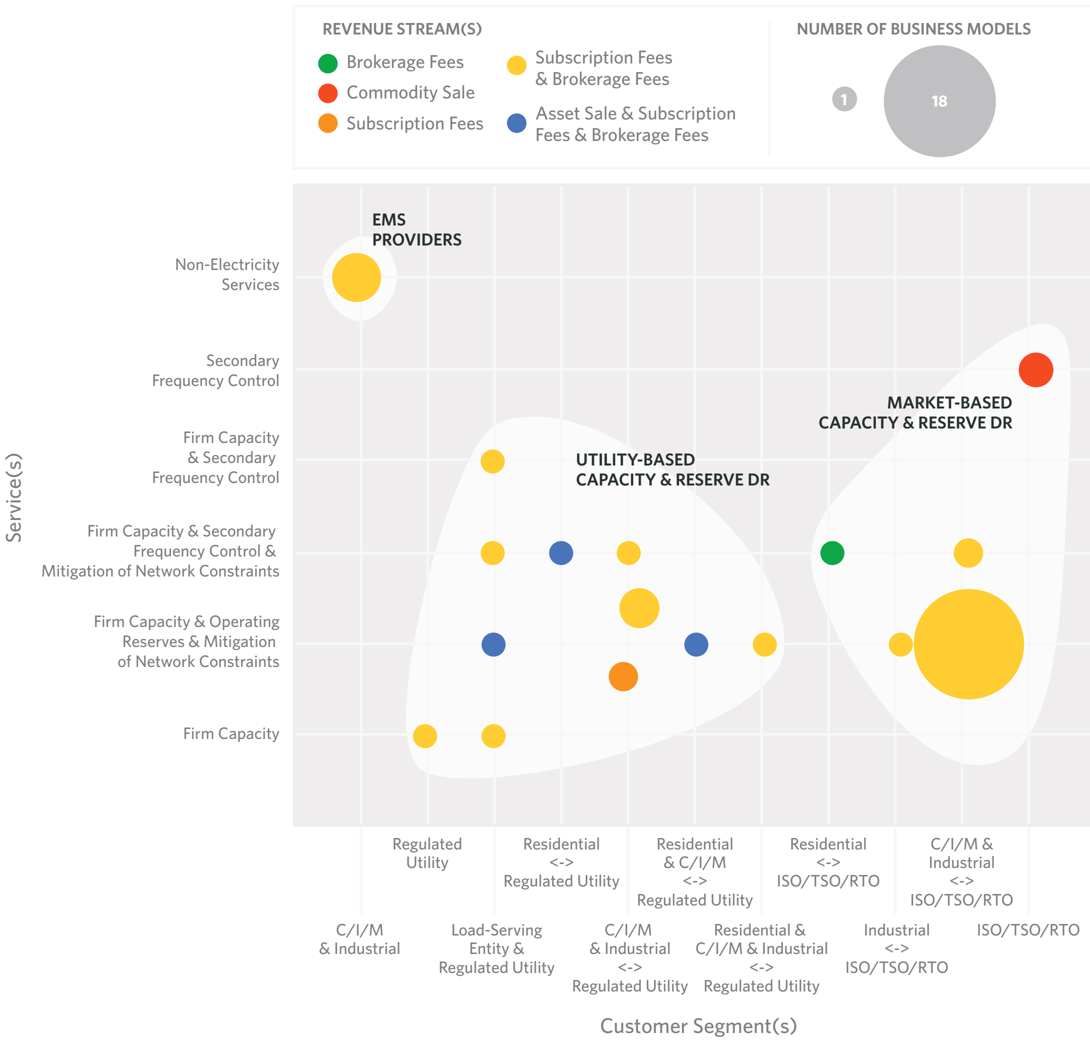
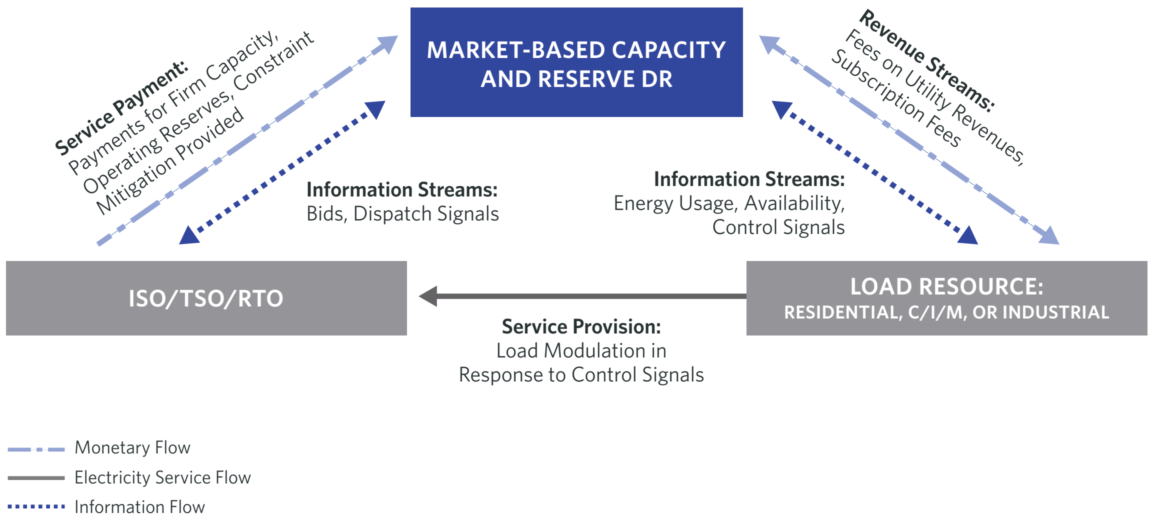


Figure B.2: Generic Business Model Structure for Market-Based Demand Response



Businesses within this archetype most commonly target large commercial, institutional, or municipal (C/I/M) and industrial customer segments. This is due to a number of factors, including market rules such as minimum bid-size requirements, transaction costs, and customer acquisition costs (Behrangrad 2015). These businesses often provide an EMS (or similar product) to the targeted customer to optimize the customer's energy consumption (and production in the case of customers with on-site distributed generation); the EMS also enables customers to participate in DR programs offered by an independent system operator (ISO), with market interaction facilitated by the DR business. In certain cases, the businesses do not provide an EMS-like product, and loads are simply controlled through alternative measures (e.g., phone calls instructing customers to manually respond). These business models most commonly leverage customer loads such as lighting; heating, ventilation, and air conditioning (HVAC) units (chillers and fans); refrigeration units; other variable-frequency drive units; idiosyncratic industrial process loads; and customer-sited generation such as backup diesel or gas units, fuel cells, or batteries (Palensky and Dietrich 2011; Hansen et al. 2014). Among the various types of operating reserves, DR is most commonly deployed for secondary reserves (i.e., contingency reserves); however, the provision of primary reserve services is becoming increasingly common (Dehghanpour and Afsharnia 2015). Secondary reserves are favored among DR providers, as secondary reserves are dispatched less frequently and the required response time is typically lower than it is for primary reserves.

These businesses typically make a profit by taking a portion of the revenues generated from the sales of these services—i.e., by brokering market revenues (brokerage fees)—and/or by charging for the use of the energy management software that enables the demand control (subscription fees). Note that the business model's revenue is a brokerage fee, rather than a commodity sale; the business distributes the revenues associated with commodity sales to the DR resources under contract.

Table B.3 includes several examples of business models that fall under this archetype. EnerNOC and REstore are typical examples of businesses selling to C/I/M and industrial customers. Ohmconnect is an interesting exception in that it explicitly targets residential customers through a user-friendly app. By sending signals to homeowners directly or communicating with home area network – connected devices, Ohmconnect enables residential consumers to participate in power system markets. Encycle targets mostly small and midsize C/I/M facilities, which are commonly overlooked among DR businesses.

Table B.3: Market-Based Demand Response Business Model Examples

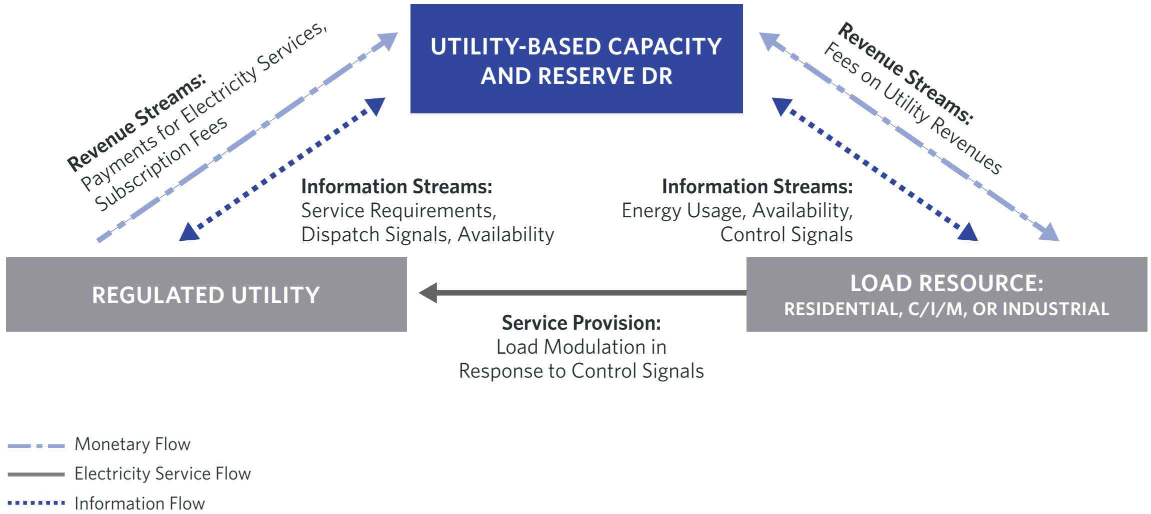
LARGE DR RESOURCE BUSINESS MODELS **Typical Customers:** • C/I/M & Industrial <-> ISO **Typical Services:** • Firm Capacity, Operating Reserves, and Constraint Mitigation	EnerNOC (United States) Innovari (United States) REstore (Europe)
SMALL DR RESOURCE BUSINESS MODELS **Typical Customers:** • Residential and Small C/I/M <-> ISO **Typical Services:** • Firm Capacity, Operating Reserves, and Constraint Mitigation	Ohmconnect (United States) Encycle (United States) Lichtblick (Europe)

B.4.2 Utility-based capacity and reserve demand response

The second-largest cluster of DR and EMS business models, which we have called "utility-based capacity and reserve DR" companies, sell demand response products directly to regulated utilities. The utility customers of these DR businesses tend to operate in vertically integrated or partially restructured markets (i.e., markets without competitive retail supply). Regulated utilities will contract with DR providers to procure (most commonly) firm capacity, operating reserves, and mitigation of network constraints. These utilities will either operate these programs under compulsion from their regulators or, in certain cases, proactively seek regulatory approval (see, for example, Consolidated Edison's program [NYPSC 2014], or Tuscon Electric Power's program [ACC 2013]). The generic structure of this business model is depicted in **Figure B.3**.

In these cases, distribution utilities seek an explicit (although sometimes unlimited) capacity of qualifying DR resources. DR businesses procure DR resources at prices determined by negotiation with the utility and the regulator. In many cases, a single DR business will operate the DR program on behalf of the utility (Hurley et al. 2014). Participating load resources obtain a share of the revenues earned by the DR aggregators. This represents a strict departure in strategy from the market-driven DR business model archetype described above; the DR provider's focus is on selling products to the utility and working with the utility to connect with (most commonly) C/I/M and industrial customers. In some cases, the utility will help the DR provider target specific customers or customer classes (DR/customer engagement opportunity identification is a stand-alone service provided by certain ICT companies as well). These DR businesses tend to earn revenues through subscription fees (i.e., payments from the utility linked to the provision of the DR management software, etc.) or

Figure B.3: Generic Utility-Based Demand Response Business Model Structure



brokerage fees (i.e., keeping a share of the revenue earned from the sale of the DR resource to the utility). Many of the businesses operating in market environments also operate in these regulated environments; examples include EnerNOC and Comverge.

In market-driven environments, transaction and customer acquisition costs have driven DR business models to target larger C/I/M and industrial customers, yet the regulated environment allows for greater participation of residential loads. However, the technical requirements of coordinating very fast responses from residential loads has limited the majority of the business models in the archetype to providing only capacity and secondary reserves (Mathieu et al. 2012). The most common loads that are used by residential DR companies are HVAC units. Examples of business models of this type include Comverge and EcoFactor. Nest, a residential smart thermostat provider, offers a similar service through its Rush Hour Rewards program.

An exception to the above model is the "behavioral" model. Businesses that use this model tend not to provide explicit control or dispatch signals, but rather provide "nudges" and targeted incentives to create a response (Allcott 2011; Gillingham and Palmer 2014). These businesses sell their services directly to the regulated utility and do not engage with the consumer outside of the context of the behavioral program. The revenue model is typically based upon subscription fees and shared savings (brokerage fees) charged to the utility. Examples of business models of this type include Opower and Tendril. **Table B.4** includes several examples of business models within this archetype.

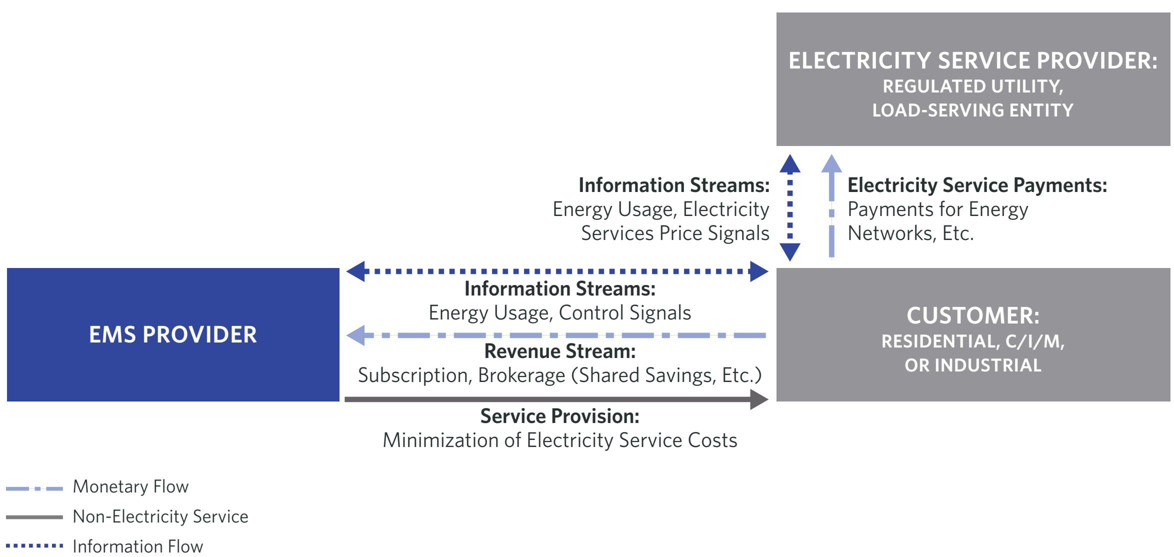
Table B.4: Business Model Examples for Utility-Based Demand Response

TRADITIONAL UTILITY DR RESOURCE BUSINESS MODELS **Typical Customers:** • C/I/M & Industrial <-> Regulated Utility **Typical Services:** • Firm Capacity, Operating Reserves, and Constraint Mitigation	EnerNOC (United States) Comverge (United States)
RESIDENTIAL-FOCUSED UTILITY DR BUSINESS MODELS **Typical Customers:** • Residential & C/I/M <-> Regulated Utility **Typical Services:** • Firm Capacity, Secondary Operating Reserves	EcoFactor (United States) Comverge (United States)
BEHAVIORAL UTILITY DR BUSINESS MODELS **Typical Customers:** • Residential <-> Regulated Utility **Typical Services:** • Firm Capacity	Opower (United States) Tendril (United States)

B.4.3 EMS providers

Finally, there are a set of businesses providing energy management systems with a focus on managing on-site operations without market interaction. These business models, which we have called "EMS providers," are focused primarily on the optimization of local energy usage in response to energy prices and local needs. A generic structure of the EMS business model is depicted in **Figure B.4**.

Figure B.4: Generic Energy Management Service Business Model Structure



These businesses tend to target C/I/M and industrial customers (Palensky and Dietrich 2011). Given that the focus of these business models is primarily on optimizing the consumption of energy services (rather than providing them), we have not considered them as providers of electricity services. Instead, they act as enablers of energy service provision by electricity consumers themselves. These businesses tend to earn revenues from shared savings arrangements (a type of brokerage fee), subscription fees (for the software provided), and asset sales of monitoring and control equipment. Examples of business models of this type for C/I/M and industrial customers include Gridpoint Energy and MeteoViva, and examples for residential customers include Nest and Wiser (a Schneider Electric company). **Table B.5** provides examples of EMS businesses.

Table B.5: Energy Management Systems Business Models

EMS PROVIDERS **Typical Customers:** • Residential, C/I/M, Industrial **Typical Services:** • Non-electricity Services	C/I/M and Industrial Providers	Gridpoint Energy (United States) Blue Pillar (United States) MeteoViva (Europe)
	Residential Providers	Ceiva Energy (United States) Rainforest Automation (Canada) Wiser/Schneider Electric (Europe)

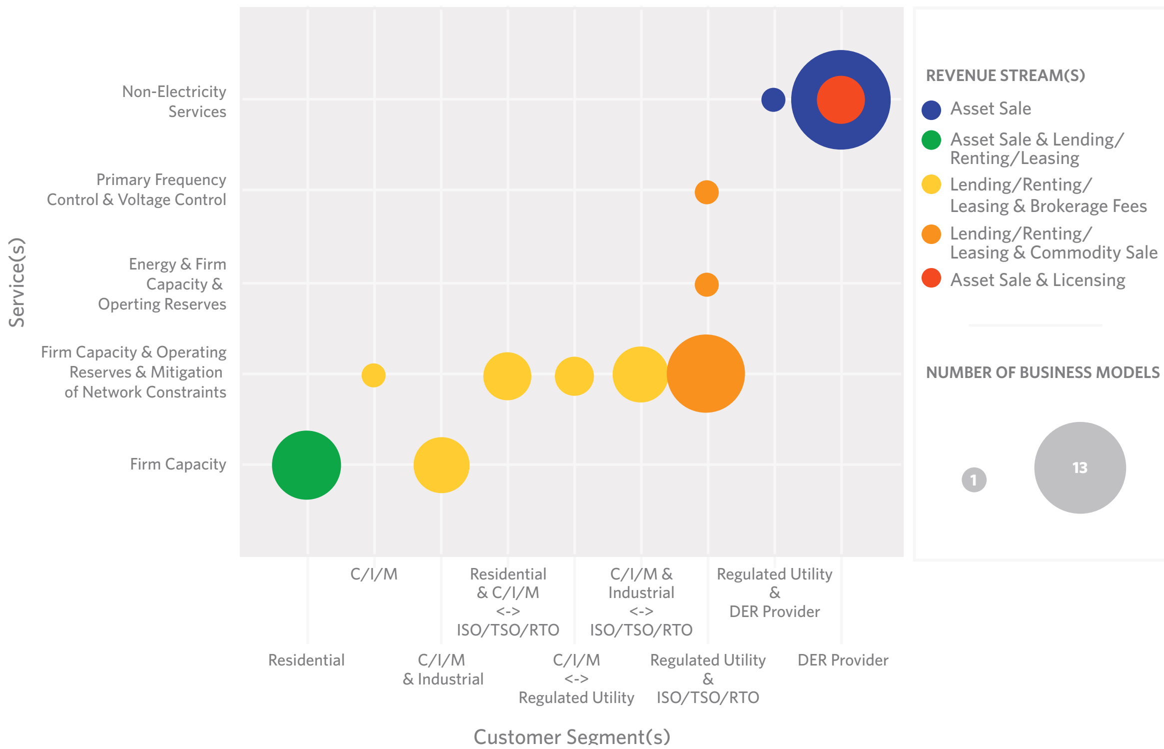
B.5 Electricity and Thermal Storage Business Model Archetypes

Electricity and thermal storage technologies[6] are often lauded as critical components of a clean energy future. It is therefore not surprising that energy storage deployments have been rapidly increasing (DOE 2015). Pumped hydro energy storage and molten salt thermal storage account for the vast majority of installed energy storage capacity to date, but these technologies are poorly suited to distributed applications (DOE 2015). Lead-acid technologies make up the bulk of distributed energy storage installations globally, although lithium-ion (Li-ion) and other advanced technologies are gaining traction (Agnew and Dargusch 2015; Schmalensee et al. 2015).

Energy storage technologies are diverse in their uses (Ruester et al. 2012; Schoff 2015). However, despite such diversity, the related business models can be clustered into three major archetypes. **Figure B.5** shows the business model landscape for electrical and thermal storage. Note that business models that deploy energy storage technologies in conjunction with solar PV are discussed in the solar PV section that follows.

6 Note that thermal storage technologies in this case exclude the storage of thermal energy in the heated or cooled space of a building, as this is categorized as demand response.

Figure B.5: Business Model Taxonomy for Electrical and Thermal Storage



From **Figure B.5** we see three major (non-manufacturer) archetypes. The primary defining feature of each archetype is its level of integration within power system operations. A number of business models are focused on the provision of ICT-based optimization and control services for energy storage technologies. Some of these businesses actively deploy projects and provide energy services to power system operators, while others simply provide optimization and control products to other businesses or end users. Business models that fall into the latter category are classified, along with technology manufactures, as providing non-electricity services.

B.5.1 Energy storage for network services

Numerous studies have highlighted the value of energy storage technologies for network and system applications, including various network capacity and ancillary service benefits (Denholm et al. 2013; EPRI 2013; Akhil et al. 2013). Certain states have begun legislatively requiring utilities to procure storage assets; for example, California's Assembly Bill 2514 requires the state's largest utilities to procure 1.3 gigawatts of storage by 2020 (California State Assembly 2010). A cluster of business models that we call "energy storage for network services" has emerged to meet this market. **Figure B.6** shows the general structure of this business model archetype; dotted lines are used between the financing function and the electrical and thermal storage (ETS) resource management/deployment function to indicate that this function could be performed internally or by partners that are external to the business.

Figure B.6: Generic Business Model Structure for Network Services Electricity and Thermal Storage

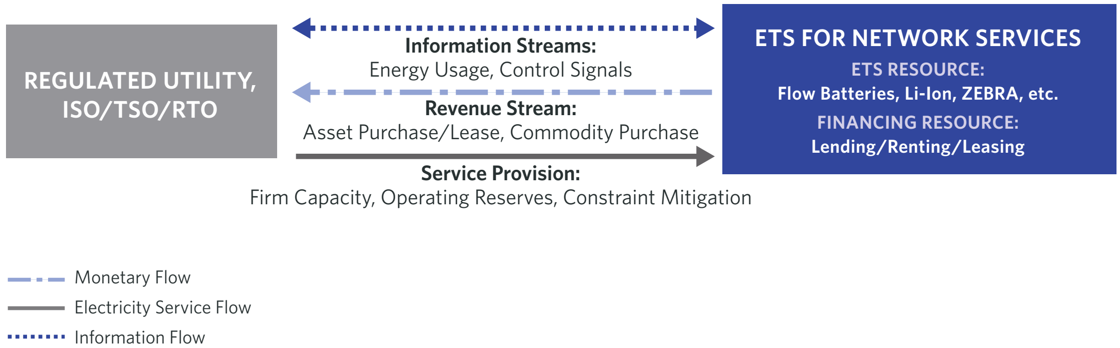


Table B.6: Network Service Business Models for Electricity and Thermal Storage

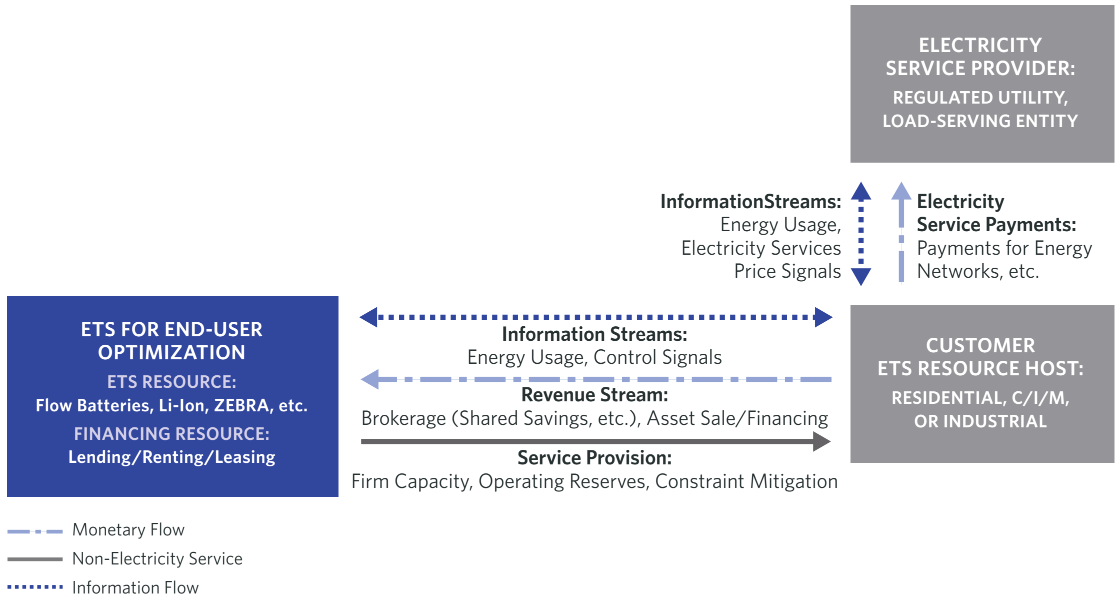
NETWORK SERVICES BUSINESS MODELS **Typical Customers:** • Industrial, Regulated Utility, Independent System Operator **Typical Services:** • Firm Capacity, Operating Reserves, Network Constraint Mitigation	Technology-Agnostic Developers	Invenergy (United States) SunEdison (United States)
	Proprietary Technology, Outsourced Manufacturing Developers	Younicos (Europe) AES Energy Storage (United States)
	Vertically Integrated Developers	Ecoult (Australia) Ambri (United States)

The majority of the businesses in this archetype are either technology-agnostic project developers, project developers with outsourced manufacturing but proprietary technology, or technology developers with downstream integrated project development arms. Businesses in this archetype tend to either serve vertically integrated regulated utilities or system operators, or they install batteries at industrial sites with the intent of providing network services. These businesses tend to earn revenues either from the sale or financing of the storage assets or from the sale of electricity services (typically firm capacity and operating reserves valued at market prices).

B.5.2 End-user optimization for energy storage

End-user systems involve installing storage assets "behind the meter" at customer sites. These behind-the-meter systems have historically been deployed at customer facilities to manage peak demand charges and arbitrage between low and high energy price hours under time-of-use or real-time pricing tariffs (actions that are collectively referred to as "bill management") (Neubauer and Simpson 2015; Masiello and Roberts 2014). These customers may be eligible to participate in bulk system markets but may desire not to. Alternatively, these customers may be in regions that do not allow distributed assets to participate in markets. The general structure for this group of companies, a business model that we call "energy storage for end-user-optimization," is depicted in **Figure B.7**.

Figure B.7: Generic Energy Storage for End-User Optimization Business Model Structure



To date, the primary motive for the deployment of residential energy storage systems has been to increase the profitability of solar PV systems through increasing "self-consumption" (i.e., minimizing the export of energy produced on site) (Masiello and Roberts 2014; Hoppmann et al. 2014; Johann and Madlener 2014). For commercial and industrial customers, the primary motive has been the avoidance of demand-based consumption charges (i.e., charges per kilowatt of peak demand) (Neubauer and Simpson 2015). As technology costs have fallen, providing backup power to residential customers and critical commercial and industrial loads has also emerged as a driver (Nair and Garimella 2010).

Business models within this archetype typically earn revenues either through a form of shared or guaranteed savings arrangements[7]—e.g., brokerage fees—or through the sale and financing of storage assets (St. John 2016). Shared or guaranteed savings arrangements are less common at the residential level. Industry analysts expect that the financing options that contributed to the rapid rise of distributed solar (discussed below) will have the same effect on energy storage (Wesoff 2015).

Table B.7: Energy Storage for End-User Optimization Business Model Examples

ENERGY STORAGE FOR END-USER OPTIMIZATION **Typical Customers:** • Residential, C/I/M, Industrial **Typical Services:** • Firm Capacity	C/I/M and Industrial Electric Thermal Storage (ETS) Providers	Stem (United States) Green Charge Networks (United States) Younicos (Europe)
	Residential ETS Providers	SolarCity (United States) Sungevity (United States) Sonnen (Europe)

7 Guaranteed savings arrangements or other variations of performance contracts are brokerage fees with a hedge against nonperformance. In these cases, the monetized value is the access to the savings. This brokerage fee differs slightly from those leveraged in demand response business models, where the monetized value is the access to new revenue streams.

B.5.3 Energy storage for end-user and system co-optimization

Regulations and market design changes have emerged (or are emerging) to allow distributed resources to be aggregated and bid into power system markets (see, for example, the European Smart Grid Task Force Expert Group 3 recommendations, the New York Reforming the Energy Vision proceeding, the California Independent System Operators' 2015 telemetry rulings, and ERCOT's "DREAM" Task Force) (MDPT Working Group 2015; Smart Grid Task Force 2015; California ISO 2015a).

In response to these emerging market-based opportunities, businesses are attempting to deploy storage technologies behind the customer meter that simultaneously attempt to lower costs to the end user and participate in power system markets. These businesses most commonly attempt to connect C/I/M and industrial customers with ISO markets, providing firm capacity, operating reserves, and mitigating network constraints. These businesses tend to earn revenue on the sales of the storage assets and/or through brokerage fees on market-based revenues (e.g., fees levied for managing market interaction on behalf of the battery host). Certain business models earn revenues on a shared savings basis (an alternative type of brokerage fee) (Green Charge Networks 2014). A general structure for this business model archetype is presented in **Figure B.8,** and examples of companies within the archetype are given in **Table B.8**.

A subset of businesses has emerged to aggregate customer-sited storage assets and either allow direct control of these assets by the utility or provide control on behalf of the distribution utility. These business models have historically targeted commercial customers and have tended to operate on an asset-sale and brokerage-fee basis.

Finally, there exists a small subset of companies that are attempting to provide operating reserves to system operators through the use of thermal storage in bricks and water heaters. Given the large potential for the use of residential water heaters, these businesses have targeted not only commercial and industrial loads but also residential loads. These companies tend to operate on a brokerage-fee basis similar to their larger storage counterparts.

Figure B.8: Business Model Structure for Generic End-User and System Co-optimization

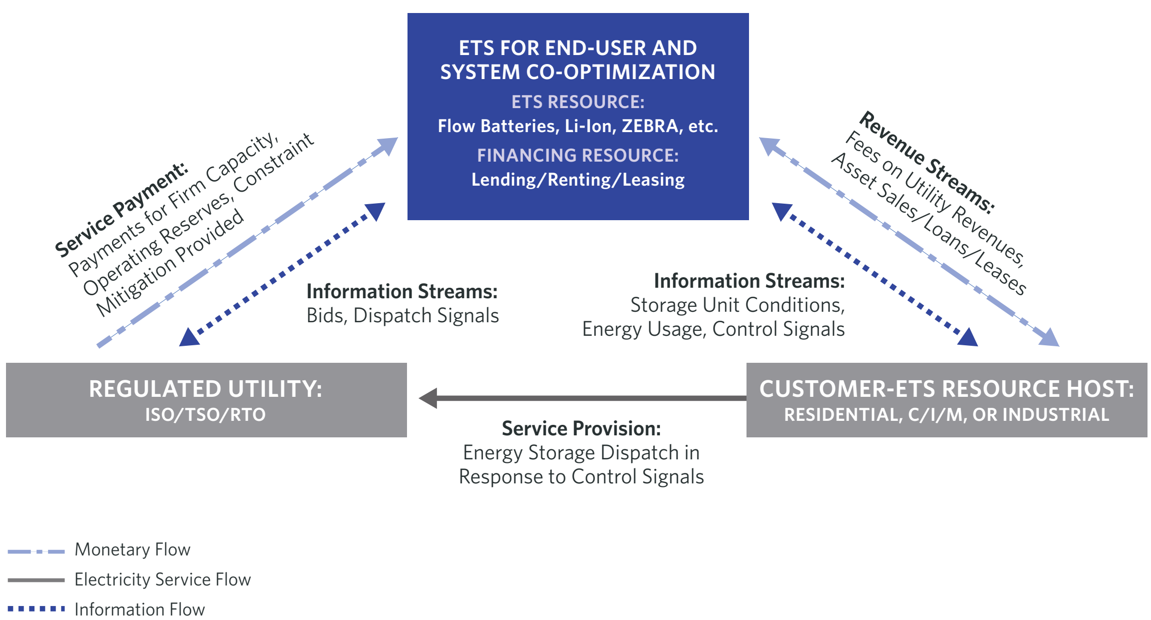
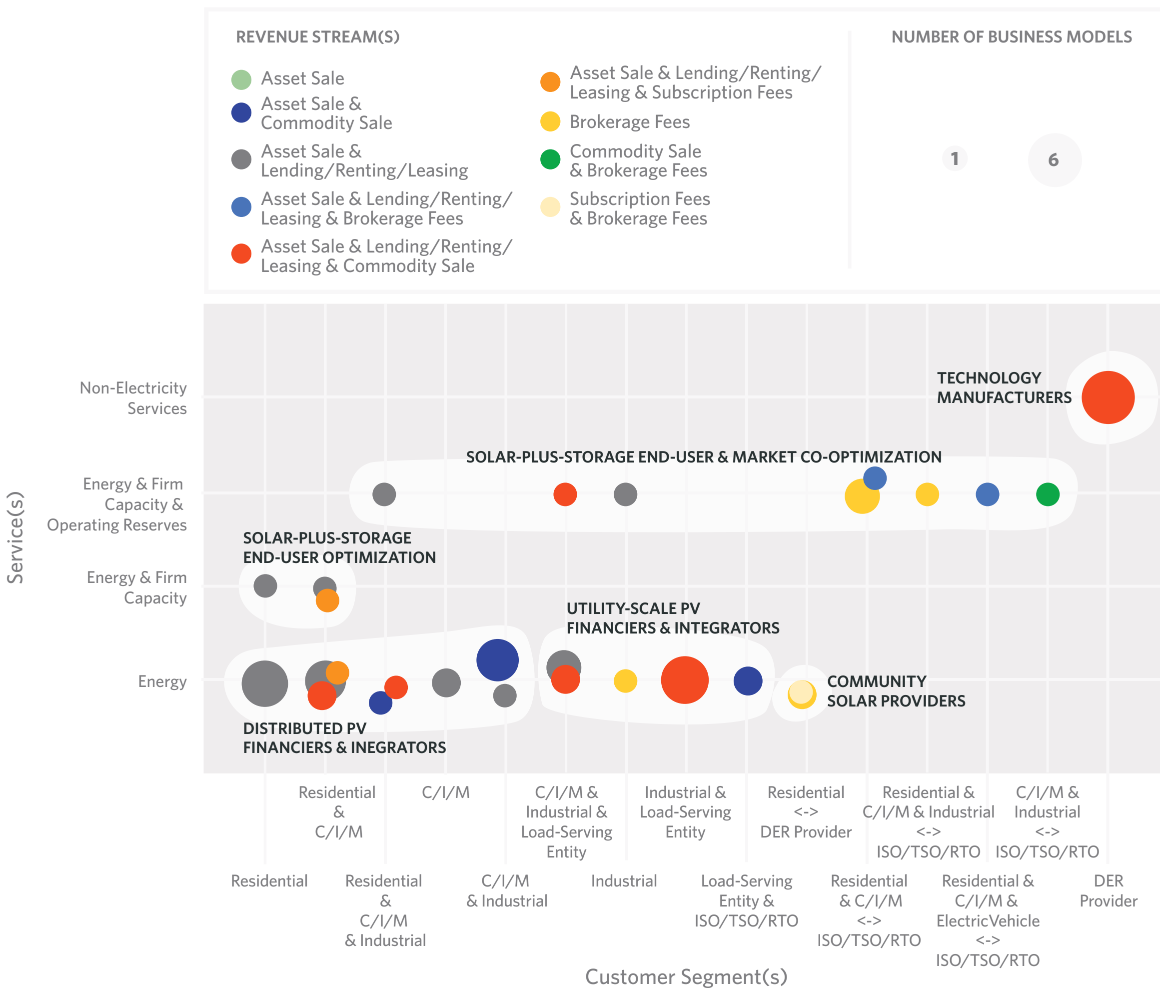


Table B.8: End-User and System Co-optimization Business Model Examples for Storage

END-USER AND SYSTEM CO-OPTIMIZATION FOR STORAGE **Typical Customers:** • Residential, C/I/M, Industrial <-> Regulated Utility, ISO/TSO/RTO **Typical Services:** • Firm Capacity, Operating Reserves, Network Constraint Mitigation	C/I/M and Industrial Providers	Stem (United States) Green Charge Networks (United States)
	Direct-Utility-Control Providers	Ice Energy (United States) Advanced Microgrid Solutions (United States)
	Residential and Small C/I/M Providers	Vcharge (United States) Steffes Corp. (United States)

Figure B.9: Solar PV and Solar-Plus Storage Business Model Taxonomy



B.6 Solar PV and Solar-Plus-Storage Business Model Archetypes

Solar photovoltaic (PV) technologies can be classified into two broad categories: wafer-based and thin film (Jean et al. 2015).[8] Crystalline silicon (c-Si) modules currently account for roughly 90 percent of global manufacturing capacity, while thin-film modules, led by cadmium telluride (CdTe), make up the remainder of the global manufacturing capacity (Metz et al. 2014). Due to a number of factors, including their higher efficiency, c-Si modules have historically dominated the relatively small distributed generation market (Davidson et al. 2013). Outside of niche distributed applications such as building-integrated PV and transparent PV, thin-film technologies are largely used in utility-scale plants.

Figure B.9 depicts the solar PV and solar-plus-storage business model landscape, as well as clusters within it. Solar PV and solar-plus-storage business models target a diverse group of customer segments and use many different revenue models. Below we describe the solar PV and solar-plus-storage business model archetypes and explore some of the diversity within each archetype.

B.6.1 Solar-plus-storage business model archetypes

Solar-plus-storage systems have similar deployment trends to their storage-only counterparts; certain integrators are focused on connecting distributed PV and storage assets with bulk power system markets or system operations, while others are focused primarily on maximizing the system owner's financial returns without integrating with markets or system operations.

Distributed energy storage paired with solar PV has become the focus of intense academic and industry study. Academic studies have focused on the potential technical benefits of distributed energy storage and solar PV (Alam et al. 2013; Hill et al. 2012; Zahedi 201; Moshövel et al. 2011) and, to a lesser degree, the economics of deploying these systems (Hoppmann et al. 2014; Khalilpour and Vassall 2015). Industry and trade organizations have focused on the economic attractiveness and systemwide economic implications of PV and storage systems for network service providers and vertically integrated utilities (Bronski et al. 2015, 2014; Byrd et al. 2014; EPRI 2014; CSIRO 2013). Of particular interest is that these systems enable the system host to significantly reduce or eliminate total consumption of energy from the bulk power system, thereby reducing network congestion and deferring investments in network reinforcements (but also commonly shifting sunk network costs from the system host to other network users) (California State Assembly 2010; Bronski et al. 2015, 2014; CSIRO 2013).

B.6.1.1 Solar-plus-storage end-user and system co-optimization

Solar-plus-storage systems deployed at customer sites are subject to many of the same market integration regulations and market rules as the storage systems discussed above. Certain system operators, such as the California ISO, have different rules for aggregations of multiple DER types (e.g., solar and storage) than they do for aggregations of single DER types (e.g., storage only) (California ISO 2015b).   demonstrates the generic structure of the business model archetype that we call "solar-plus-storage end-user and system co-optimization." Again, dotted lines are used between the financing function and the solar-plus-storage resource management/deployment function to indicate that this function could be performed either internally or by partners that are external to the business.

8 Note that solar thermal technologies, including concentrated solar power technologies, are excluded from this review. Solar PV technologies account for more than 98 percent of installed solar capacity and nearly 100 percent of the global distributed solar capacity (REN21 2014).

Figure B.10: Generic Solar-Plus-Storage for End-User and System Co-optimization Business Model Structure

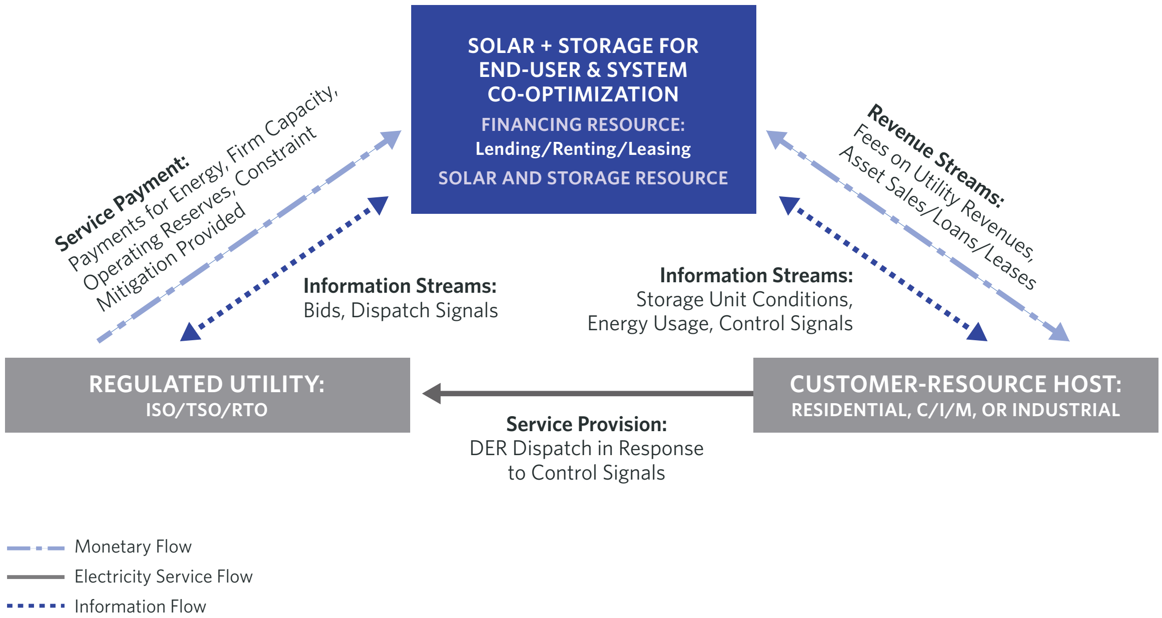


Table B.9: Solar-Plus-Storage for End-User and System Co-optimization Business Model Examples

SOLAR-PLUS-STORAGE FOR END-USER AND SYSTEM CO-OPTIMIZATION **Typical Customers:** • Residential, C/I/M, Industrial <-> Regulated Utility, ISO/TSO/RTO **Typical Services:** • Energy, Firm Capacity, Operating Reserves	Solar-Plus-Storage Developers	Sunverge (United States) Solar Grid Storage/ SunEdison (United States)
	Virtual Power Plant Developers	Lichtblick (Europe) DONG Powerhub (Europe)

A number of business models have emerged that attempt to bring "firm" solar PV resources to market by pairing solar PV and storage technologies. The aggregations of PV and storage (and, in some cases, other technologies, such as demand response and distributed generators) are often termed "virtual power plants," or "VPPs." Revenue streams are structured around the sale and financing of the assets and fees for brokering market interactions on behalf of the system hosts. In certain cases, businesses will own the projects and earn revenues on the sales of energy (most often under long-term power purchase agreements), operating reserve, and capacity services (i.e., commodity sales revenues). **Table B.9** presents some examples of businesses currently operating within this archetype.

B.6.1.2 Solar-plus-storage end-user optimization

As noted above, solar-plus-storage systems have been most commonly deployed at customer sites to increase self-consumption (most often in the face of explicit incentives to do so), provide backup power, and minimize electricity demand charges. **Figure B.11** presents a general structure of this business model archetype, which we have called "solar-plus-storage end-user optimization."

Figure B.11: Generic Solar-Plus-Storage for End-User Optimization Business Model Structure

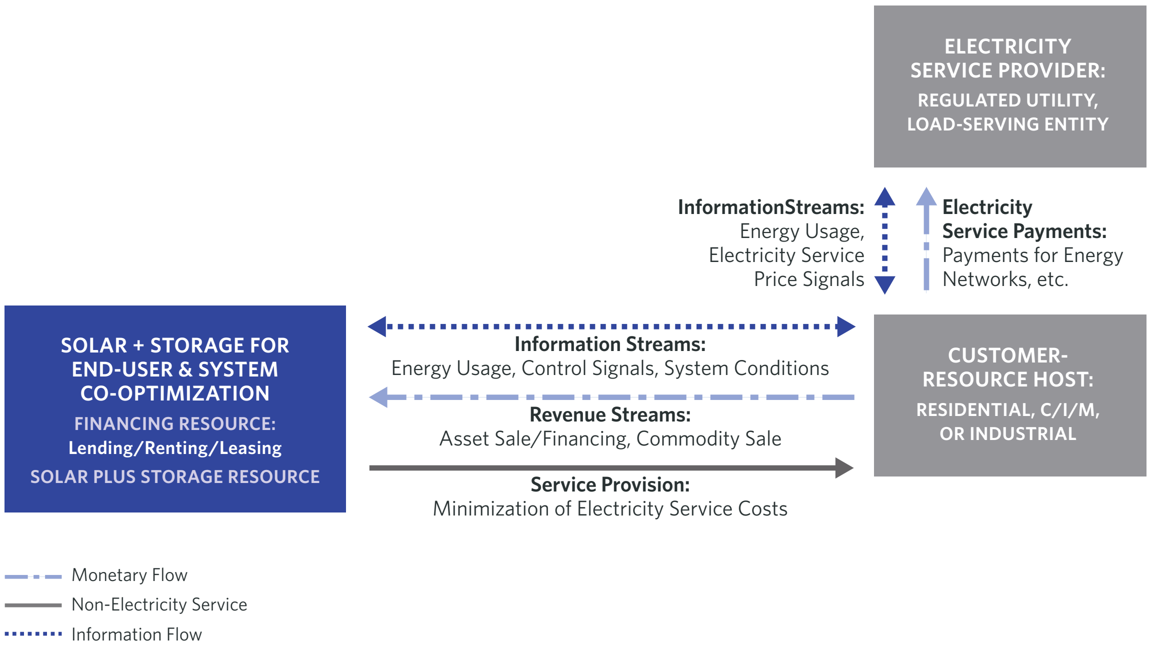


Table B.10: Solar-Plus-Storage for End-User Optimization Business Model Examples

SOLAR-PLUS-STORAGE FOR END-USER OPTIMIZATION **Typical Customers:** • Residential, C/I/M, Industrial **Typical Services:** • Energy, Firm Capacity	SolarCity (United States) Juicebox Energy (United States)

Many US states have explicit subsidies for energy storage technologies. In recent years, governments, including the German government, have begun offering subsidies for solar PV paired with energy storage (KfW 2013). These businesses operate very similarly to their pure-play storage or solar counterparts. They tend to sell products directly to residential and C/I/M customers and structure revenue streams around the sales and financing of solar and storage assets. Examples of companies within this business model archetype are presented in **Section B.6.2**.

B.6.2 Solar photovoltaics business model archetypes

PV system integrators fall into three key archetypes, each with significant internal nuances, which are explored below. All grid-connected PV systems operate with a DC/AC inverter. These inverters have the capability to modulate their power factor, providing or consuming active power (during producing hours) and reactive power (at all hours). Modulating the power factor of distributed PV and storage systems has been shown to be effective at maintaining distribution voltage in certain cases (Turitsyn et al. 2010; Moreno et al. 2015). Furthermore, Volt/VAR control through PV inverters has been shown to enable conservation voltage reduction (CVR) and line loss reduction, leading to significant power savings (Farivar et al. 2011). However, the economic opportunity of voltage support and CVR is currently considered to be rather small (Akhil et al. 2013). Furthermore, PV inverters have historically been required to operate at a power factor[9] of unity in the United States. Standards are currently being developed to allow for power factor control in the distribution system (IEEE Standards Association n.d.). Germany has mandated that all PV inverters be capable of providing voltage support through reactive power control since 2009 (SMA Solar Technology 2012). Despite these recent developments, however, few frameworks have been developed to decide whether or not (and if so, how) to remunerate distributed PV systems for providing voltage support. For these reasons, no prominent business models exist to enable PV systems to provide these services, and all PV business models are currently focused primarily on energy provision.

B.6.2.1 Distributed PV finance and installation

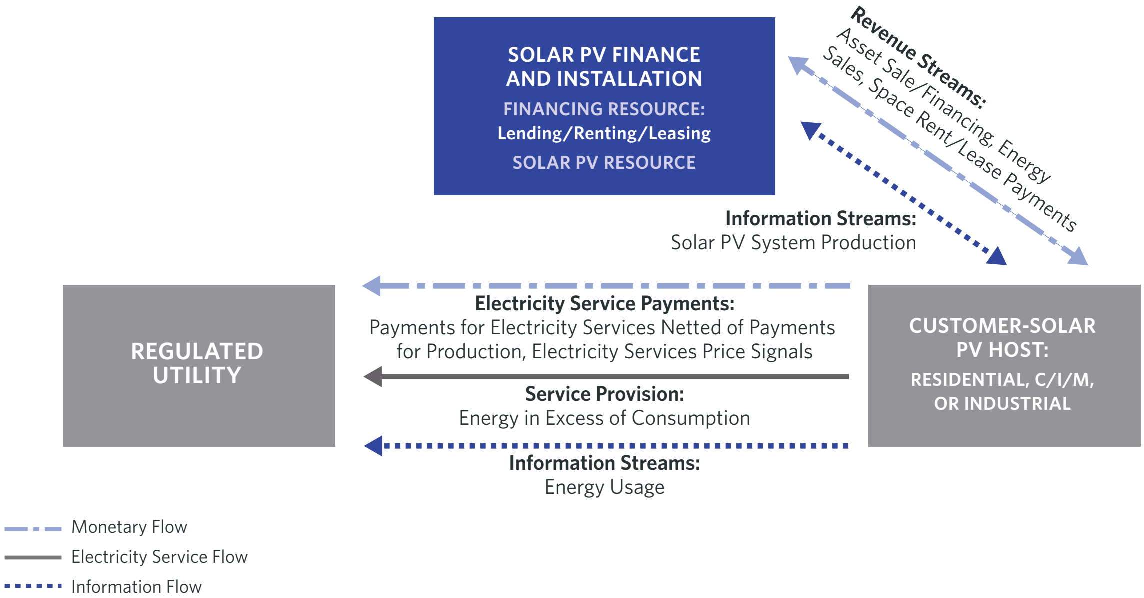
Historically, high capital costs have been a major impediment to PV adoption. Over the past decade, however, technology costs have fallen, and financing solutions have emerged to combat the challenge of high initial costs. Such "distributed PV finance and installation" companies comprise the largest solar PV business model archetype. The two dominant methods of financing distributed PV are direct ownership (through direct purchase or a debt product) and third-party ownership models[10] (Bolinger and Holt 2015; Speer 2012). In 2014, third-party ownership financing structures were employed in 60 percent to 90 percent of installations in the largest residential PV markets in the United States (Bolinger and Holt 2015). Indeed, available financing has been shown to be a major driver of distributed PV investment (Drury et al. 2013). Historically, third-party ownership structures have been more popular in the United States than in the European Union. However, some industry analysts are predicting growth in these types of financing models as feed-in tariff policies wane in Europe (Sharma et al. 2015).

The exact structure of the financing option depends on the policy environment, the degree of technological development, and many other factors. In the United States, the largest single explicit subsidy for distributed PV systems is the investment tax credit (ITC). The ITC is a credit applied to the income taxes of the ITC claimant that is based on the capital cost of the installed system (for a basic review of the ITC, see SEIA [2015]). However, many homeowners and small business owners do not have the "tax appetite" (i.e., do not pay enough in taxes) to fully benefit from this subsidy (Speer 2012). Business models have emerged to capture this subsidy and enable the customer to procure the PV system for low or no upfront costs. Typically the installer and a third party (i.e., neither the installer nor the PV system host) own the PV system and monetize the ITC (Speer 2012). The PV system host then gives the installer/owner a lease payment or a payment for the energy produced (the two payment methods are functionally identical) (Speer 2012). In feed-in tariff (FIT) and net-metering environments, the business' revenue stream can be generated directly from the sales of energy at the FIT or net-metered rate; in these cases, the business will typically rent or lease the real estate (including rooftops) from the system host. Finally, in European markets and with increasing regularity in the United States, loan products offer a path to reducing upfront investment barriers (Bolinger and Holt 2015; Seel et al. 2014). **Figure B.12** reflects the generic structure of this business model archetype.

9 The power factor is defined as the ratio between the active power (kW) and the apparent power (kVA).

10 Niche financing options such as "property assessed clean energy" are emerging, but the dominant methods remain direct and third-party ownership.

Figure B.12: Generic Solar PV Finance and Installation Business Model Structure



The exact structure of the project deployment financing may dramatically change the economics of projects. However, the financing structure does not change the basic components of the business model (i.e., the business is still earning revenues through a lease-like payment). For a detailed description of deployment financing methods for solar PV, see Lutton (2013) and Speer (2012). Furthermore, there are a number of methods for selling bundles of solar PV projects after the projects have been installed; these options fall into the categories of securitization and bonding. These methods can significantly lower the cost of capital for the businesses installing the assets (Motyka et al. 2015). However, these post-installation bundling and sales methods do not dramatically change the electricity services aspects of the business models discussed. That is, the business must still identify/locate a customer, install a system, and deliver an electricity service that generates a revenue stream sufficient to justify the cost of capital.

A significant amount of variation exists in the role and the degree of vertical and horizontal integration of the solar PV integrator. For example, some small installers partner with larger financiers; others outsource the installation of the systems. Still others perform the installation and financing functions themselves. SolarCity has announced plans to further vertically integrate into manufacturing. Certain businesses have undergone horizontal integration. For example, a number of load-serving entities (LSEs) have begun to provide solar PV. In these cases, the LSE earns revenues on the sale or financing of the asset as well as on subscription fees that are typically levied on retail customers. Some companies that have traditionally focused on home security services, such as Vivint Solar and Alarm.com, have also horizontally integrated into solar PV installation.

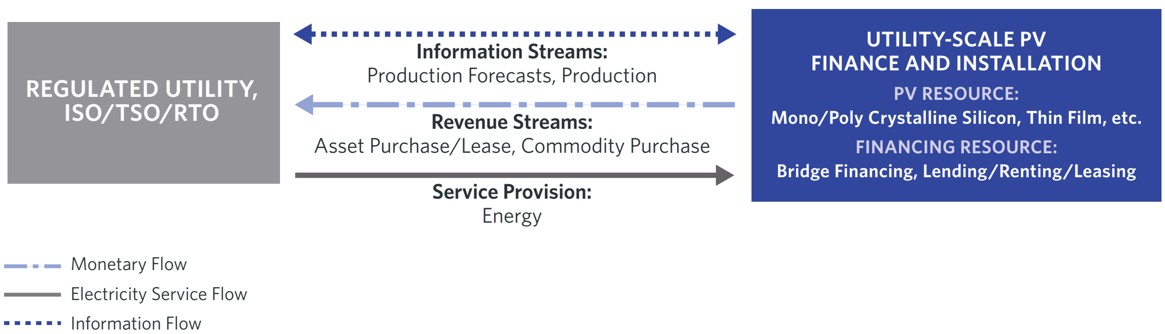
Finally, a significant number of business models have emerged that provide support services, such as generating sales leads, performing pure financing functions, or performing system maintenance. These services are categorized as non-electricity services.

Table B.11 presents examples of some of the companies in this diverse solar PV business model archetype.

Table B.11: Distributed Solar PV Finance and Installation Business Model Examples

DISTRIBUTED SOLAR PV FINANCE AND INSTALLATION **Typical Customers:** • Residential, C/I/M, Industrial **Typical Services:** • Energy	Vertically Integrated Providers	SolarCity (United States)
		Solairedirect (Europe)
	Non-vertically Integrated Providers	SunRun (United States)
		Clean Power Finance (United States)
	Horizontally Integrated Providers	Vivint Solar (United States)
		Vector (New Zealand)

Figure B.13: Generic Utility-Scale Solar PV Finance and Installation Business Model Structure



B.6.2.2 Utility-scale PV finance and installation

It is difficult to demarcate the scale of installation that should be considered “utility scale,” as many “utility-scale” plants are connected at distribution voltages (SMA Solar Technology 2012); for example, as of March 2013, 95 percent of solar PV capacity in Germany was connected to low- and medium-voltage networks, despite the significant number of multi-megawatt-scale plants (von Appen et al., 2013). Nonetheless, many businesses focus exclusively on developing large, multi-megawatt-scale PV plants. The primary driver behind these types of installations in the United States has been the fulfillment of “renewable portfolio standards” that require utilities or other agents (commonly the load-serving entity) to procure a certain quantity of solar PV (or credits associated with solar) by a certain date (Steward and Doris 2014). Therefore, large-scale solar businesses commonly sell to multiple parties: The energy is sold into wholesale markets or directly under long-term power purchase agreements (PPAs, i.e., commodity sales) to industrial or regulated utility customers, while the renewable energy credits (RECs) associated with the energy are sold to another party (again, most often the load-serving entity). In certain cases, these businesses sell large projects to commercial or industrial customers and then sell credits associated with the energy produced to utilities (Wiser et al. 2010). These business models require financing structures that are often quite different from those deployed at smaller scale plants. Utility-scale PV business models tend to be focused on the establishment of PPAs. **Figure B.13** presents a general structure of such “utility-scale PV finance and installation” business models.

There are many supporting roles and distinctions within the category of utility-scale PV providers. Certain businesses are focused entirely on procuring the rights to land, ensuring that PPAs are signed, and sourcing contractors to perform construction. These “developers” will often eschew any ownership of the project after construction. Other engineering, procurement, and

Table B.12: Business Model Examples for Utility-Scale Solar PV Finance and Installation

UTILITY-SCALE SOLAR PV FINANCE AND INSTALLATION **Typical Customers:** • Industrial, Regulated Utility, ISO/TSO/RTO **Typical Services:** • Energy	SunEdison (United States) First Solar (United States) Juwi Solar (Europe) Sainty Solar (China)

construction companies (EPCs) focus on exactly what that name suggests—engineering, procuring supplies for, and constructing the projects. EPCs tend not to take ownership stakes in the projects in which they participate.

Business models within this archetype utilize securitization methods similar to those deployed by the distributed PV business models discussed above (Motyka 2015). As with the distributed business models, the method of securitization, bonding, or direct sale of the assets after installation functions primarily to lower the cost of capital for the installation and does not significantly change the manner in which the electricity services are provided. **Table B.12** presents examples of companies within this archetype.

B.6.2.3 Community solar providers

Many residences and commercial buildings are not proper sites for distributed PV installations because of shading, building ownership challenges, and other factors. "Community solar providers" have emerged to capitalize on economies of unit scale or to enable consumers located in unsuitable areas to procure solar PV. Community solar involves installing large solar PV plants located away from the customer site. Customers can purchase the rights to a portion of the output of the solar plant, or can purchase an equity stake or share in revenues from a portion of the plant outright (Coughlin et al. 2010). The business earns revenues by charging the customer for access to the PV system outputs (brokerage fees). The community solar provider will typically sell the plant's output under a long-term PPA and distribute the associated revenues to the project's shareholders (Coughlin et al. 2010). The community solar provider approach has been particularly popular among regulated utilities that see it as a way to leverage their strengths and provide a value-added solar service (Siegrist et al. 2013). The community solar market is still relatively small (tens to hundreds of megawatts in the United States) and geographically restricted to policy-friendly environments, but it is expected to grow over the next decade (Munsell 2015). An interesting and related model is that of solar "crowd funding" startup, Mosaic. Mosaic allows individuals (e.g., homeowners or business owners as opposed to banks) to offer funds (typically in the form of debt) to finance the construction of solar projects. In this way, Mosaic acts as a bridge between financiers (individuals) and system owners (other individuals), and charges a fee for offering this service (brokerage). Mosaic has used this same model to allow individuals to own parts of centralized plants—a form of community solar. **Figure B.14** provides provides a general structure of this business model archetype.

The community solar provider approach has been particularly popular among regulated utilities that see it as a way to leverage their strengths and provide a value-added solar service (Siegrist et al. 2013). The community solar market is still relatively small (tens to hundreds of megawatts in the United States) and geographically restricted to policy-friendly environments, but it is expected to grow over the next decade (Munsell 2015). An interesting and related model is that of solar "crowd-funding" startup Mosaic. Mosaic allows individuals (e.g., homeowners or business owners as opposed to banks) to offer funds (typically in the form of debt) to finance the construction of solar projects. In this way, Mosaic acts as a bridge between financiers (individuals) and system owners (other individuals), and charges a fee for offering this service (brokerage). Mosaic has used this same model to allow individuals to own parts of centralized plants—a form of community solar.

Table B.13 provides examples of companies operating within the community solar provider business model archetype.

Figure B.14: Generic Community Solar Business Model Structure

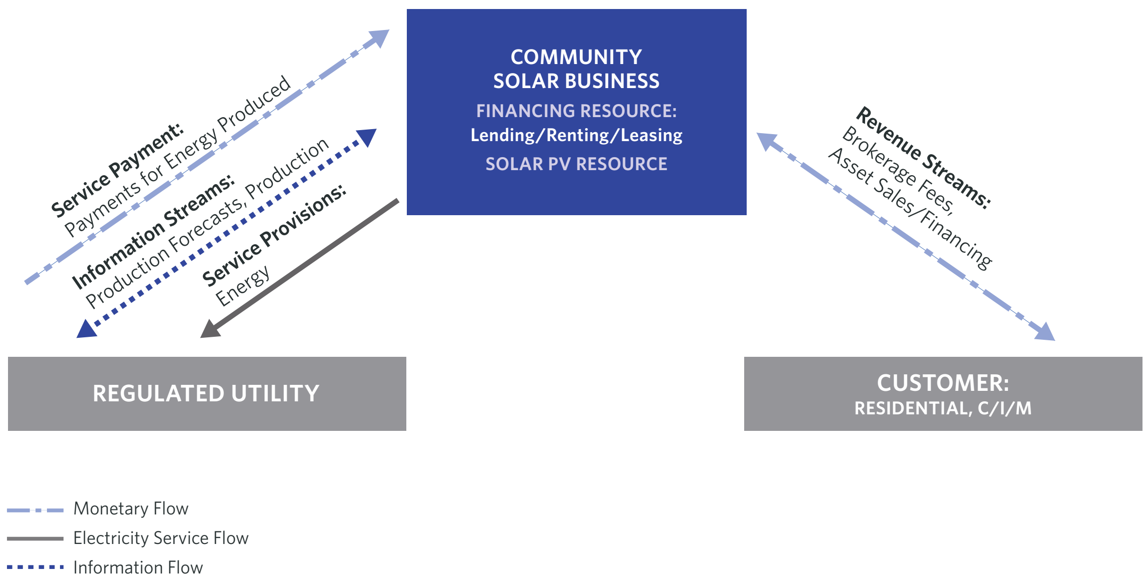


Table B.13: Community Solar Business Model Examples

COMMUNITY SOLAR **Typical Customers:** • Residential <-> DER Provider **Typical Services:** • Energy	Nexamp (United States) Next Step Living (United States)* Blue Wave Renewables (United States) Mosaic (United States)

*No longer in business.

References

Electricity Innovation Lab. 2013. "New Business Models for the Distribution Edge: The Transition from Value Chain to Value Constellation." Boulder, CO: Rocky Mountain Institute.

ACC. 2013. "In the Matter of the Application of Tucson Electric Power Company for the Establishment of Just and Reasonable Rates and Charges Designed to Realize a Reasonable Rate of Return on the Fair Value of Its Operations Throughout the State of Arizona." Vol. E-01933A-1. Tuscon, AZ: Arizona Corporation Commission.

Agnew, S. and P. Dargusch. 2015. "Effect of Residential Solar and Storage on Centralized Electricity Supply Systems." *Nature Climate Change* 5: 315–8. doi:10.1038/nclimate2523.

Akhil, A.A., G. Huff, A.B. Currier, B.C. Kaun, D.M. Rastler, S.B. Chen, et al. 2013. *Electricity Storage Handbook in Collaboration with NRECA*. Albuquerque, NM: US Department of Energy/Electric Power Research Institute. doi:SAND2013-5131.

Alam, M.J.E., K.M. Muttaqi, and D. Sutanto. 2013. "Mitigation of Rooftop Solar PV Impacts and Evening Peak Support by Managing Available Capacity of Distributed Energy Storage Systems." *IEEE Transactions on Power Systems* 28: 3874–84. doi:10.1109/TPWRS.2013.2259269.

Allcott, H. 2011. "Social Norms and Energy Conservation." *Journal of Public Economics* 95: 1082–95. doi:10.1016/j.jpubeco.2011.03.003.

Amusa, M. 2013. "How the Biggest U.S. Utilities Are Planning for the Future." *Utility Dive*. www.utilitydive.com/news/how-the-biggest-us-utilities-are-planning-for-the-future/130302/.

Behrangrad, M. 2015. "A Review of Demand Side Management Business Models in the Electricity Market." *Renewable Sustainable Energy Reviews* 47: 270–83. doi:10.1016/j.rser.2015.03.033.

Bell, M., J. Creyts, V. Lacy, and J. Sherwood. 2014. "Bridges to New Solar Business Models: Opportunities to Increase and Capture the Value of Distributed Solar Photovoltaics." Boulder, CO: Rocky Mountain Institute.

Bolinger, M. and E. Holt. 2015. "A Survey of State and Local PV Program Response to Financial Innovation and Disparate Federal Tax Treatment in the Residential PV Sector." Berkeley, CA: Lawrence Berkeley National Laboratory.

Bronski, P., J. Creyts, M. Crowdis, S. Doig, J. Glassmire, H. Energy, et al. 2015. "The Economics of Load Defection: How Grid-Connected Solar-Plus-Battery Systems Will Compete with Traditional Electric Service, Why It Matters, and Possible Paths Forward." Boulder, CO: Rocky Mountain Institute.

Bronski, P., J. Creyts, L. Guccione, M. Madrazo, J. Mandel, B. Rader B, et al. 2014. "The Economics of Grid Defection: When and Where Distributed Solar Generation Plus Storage Competes with Traditional Utility Service." Boulder, CO: Rocky Mountain Institute.

Byrd, S., T. Radcliff, S. Lee, B. Chada, D. Olszewski, Y. Matayoshi, et al. 2014. "Solar Power & Energy Storage Policy Factors vs . Improving Economics." Morgan Stanley Blue Paper.

California ISO. 2015a. "ISO Board Approves Gateway to the Distributed Energy Future." Folsom, CA: California Independent System Operator.

California ISO. 2015b. "Expanded Metering and Telemetry Options Phase 2: Distributed Energy Resource Provider." Draft Final Proposal. Folsom, CA: California Independent System Operator.

California State Assembly. 2010. AB-2514 Energy Storage Systems. Sacramento, CA.

Cappers P., C. Goldman, and D. Kathan. 2010. "Demand Response in U.S. Electricity Markets: Empirical Evidence." *Energy* 35: 1526–35. doi:10.1016/j.energy.2009.06.029.

Coughlin J., J. Grove, L. Irvine, F.J. Janet, S.J. Phillips, L. Moynihan, et al. 2010. "A Guide to Community Solar: Utility, Private, and Non-Profit Development." Prepared by Bonneville Environmental Foundation and Northwest SEED. Golden, CO: National Renewable Energy Laboratory.

CSIRO. 2013. "Change and Choice—The Future Grid Forum's Analysis of Australia's Potential Electricity Pathways to 2050." Newcastle, Australia: The Commonwealth Scientific and Industrial Research Organisation.

Davidson, C., T.L. James, R. Margolis, R. Fu, and D. Feldman. 2014. "U.S. Residential Photovoltaic (PV) System Prices, Q4 2013 Benchmarks: Cash Purchase, Fair Market Value, and Prepaid Lease Transaction Prices." Golden, CO: National Renewable Energy Laboratory.

Dehghanpour, K. and S. Afsharnia. 2015. "Electrical Demand Side Contribution to Frequency Control in Power Systems: A Review on Technical Aspects." *Renewable and Sustainable Energy Reviews*: 41: 1267–76. doi:10.1016/j.rser.2014.09.015.

Denholm P., J. Jorgenson, T. Jenkin, D. Palchak, B. Kirby, and M.O. Malley. 2013. "The Value of Energy Storage for Grid Applications." Golden, CO: National Renewable Energy Laboratory.

Drury, E., P. Denholm, and R. Margolis. 2013. "Sensitivity of Rooftop PV Projections in the SunShot Vision Study to Market Assumptions." Golden, CO: National Renewable Energy Laboratory.

E.ON. 2014. "Our Strategy: 'Empowering Customers. Shaping Markets.'" EON.com.

EEI. 2013. "Key Facts about the Electric Power Industry." Washington, DC: Edison Electric Institute.

EIA. 2015. Electric Power Sales, Revenue, and Energy Efficiency Form EIA-861 Detailed Data Files. US Energy Information Administration.

EPRI. 2013. "Cost-Effectiveness of Energy Storage in California: Application of the EPRI Energy Storage Valuation Tool to Inform the California Public Utility Commission Proceeding R. 10-12-007." Palo Alto, CA: Electric Power Research Institute. doi:3002001162.

EPRI. 2014. "The Integrated Grid: Realizing the Full Value of Central and Distributed Energy Resources." Palo Alto, CA: Electric Power Research Institute.

Farivar, M., C.R. Clarke, S.H. Low, and K.M. Chandy. 2011. "Inverter VAR Control for Distribution Systems with Renewables." *2011 IEEE International Conference on Smart Grid Communications* 457–62. doi:10.1109/SmartGridComm.2011.6102366.

Gillingham, K. and K. Palmer. 2014. "Bridging the Energy Efficiency Gap: Policy Insights from Economic Theory and Empirical Evidence." *Review of Environmental Economics and Policy Journal* 8: 18–38. doi:10.1093/reep/ret021.

Green Charge Networks. 2014. "The City of Lancaster Aims to Become the First Net-Zero City in the U.S. with Energy Storage System from Green Charge Networks." Lancaster, CA: Museum of Art and History.

Gruber, T.R. 1993. "A Translation Approach to Portable Ontology Specifications." *Knowledge Acquisition* 5: 199–220.

Hansen J., J. Knudsen, and A. Annaswamy. 2014. “Demand Response in Smart Grids: Participants, Challenges, and a Taxonomy.” IEEE 53rd Annual Conference on Decision and Control. Los Angeles, CA: IEEE. 4045–52. doi:10.1109/CDC.2014.7040018.

Hill, C.A., M.C. Such, D. Chen, J. Gonzalez, and W.M. Grady. 2012. “Battery Energy Storage for Enabling Integration of Distributed Solar Power Generation.” *IEEE Transactions on Smart Grid 3*: 850–7. doi:10.1109/TSG.2012.2190113.

Hoppmann, J., J. Volland, T.S. Schmidt, and V.H. Hoffmann. 2014. “The Economic Viability of Battery Storage for Residential Solar Photovoltaic Systems: A Review and a Simulation Model.” *Renewable and Sustainable Energy Reviews* 39: 1101–18. doi:10.1016/j.rser.2014.07.068.

Huijben, J.C.C.M. and G.P.J. Verbong. 2013. “Breakthrough Without Subsidies? PV Business Model Experiments in the Netherlands.” *Energy Policy* 56: 362–70. doi:10.1016/j.enpol.2012.12.073.

Hurley, D., P. Peterson, and M. Whited. 2013. “Demand Response as a Power System Resource: Program Designs, Performance, and Lessons Learned in the United States.” Synapse Energy Economics, Inc.

i3 Connect. n.d. Cleantech Group. i3connect.com.

IBM Global Business Services. 2006. “Expanding the Innovation Horizon: The Global CEO Study 2006.” Somers, NY: IBM.

IEEE Standards Association. n.d. P1547.8—IEEE Draft Recommended Practice for Establishing Methods and Procedures that Provide Supplemental Support for Implementation Strategies for Expanded Use of IEEE Standard 1547.

Jean, J., P.R. Brown, R.L. Jaffe, T. Buonassisi, and V. Bulovic. 2015. “Pathways for Solar Photovoltaics.” *Energy and Environmental Science* 8: 1200–19. doi:10.1039/C4EE04073B.

Johann, A. and R. Madlener. 2014. “Profitability of Energy Storage for Raising Self-Consumption of Solar Power: Analysis of Different Household Types in Germany.” *Energy Procedia* 61: 2206–10. doi:10.1016/j.egypro.2014.12.110.

KfW. 2013. “KfW and Federal Environment Ministry Launch Programme to Promote Use of Energy Storage in Solar PV Installations.” Kfw.de.

Khalilpour, R. and A. Vassallo. 2015. “Leaving the Grid: An Ambition or a Real Choice?” *Energy Policy* 82: 207–21. doi:10.1016/j.enpol.2015.03.005.

Kind, P. 2013. “Disruptive Challenges: Financial Implications and Strategic Responses to a Changing Retail Electric Business.” Washington, DC: Edison Electric Institute.

Lacey, S. 2015. “NRG Will Spin Off Its Clean Energy Business: Investor Pressure, or a Bet on a Green Premium?” *Greentech Media*.

Lehr, R. 2013. “New Utility Business Models: Utility and Regulatory Models for the Modern Era.” *Electricity Journal* 26:35–53. doi:10.1016/j.tej.2013.09.004.

Lutton, J. 2013. “Tax Equity 101: Structures.” Woodlawn Associates.

Masiello, R.D., B. Roberts, and T. Sloan. 2014. “Business Models for Deploying and Operating Energy Storage and Risk Mitigation Aspects.” *Proceedings of the IEEE* 102: 1052–64. doi:10.1109/JPROC.2014.2326810.

Mathieu, J.L., M. Dyson, and D.S. Callaway. 2012. "Using Residential Electric Loads for Fast Demand Response: The Potential Resource and Revenues, the Costs, and Policy Recommendations." Proceedings of the ACEEE Summer Study on Energy Efficient Buildings. Pacific Grove, California: American Council for an Energy-Efficient Economy. 189–203.

MDPT Working Group. 2015. "Report of the Market Design and Platform Technology Working Groups." Market Design and Platform Techology, New York State Department of Public Service.

Metz, A., M. Fischer, G. Xing, L. Yong, and S. Julsrud. 2014. "International Technology Roadmap for Photovoltaic: 2013 Results." Semi, PVGroup.

Moreno, R., R. Moreira, and G. Strbac. 2015. "A MILP Model for Optimising Multi-Service Portfolios of Distributed Energy Storage." *Applied Energy* 137: 554–66. doi:10.1016/j.apenergy.2014.08.080.

Moshövel, J., K.-P. Kairies, D. Magnor, M. Leuthold, M. Bost, S. Gährs S, et al. 2015. "Analysis of the Maximal Possible Grid Relief from PV-Peak-Power Impacts by Using Storage Systems for Increased Self-Consumption." *Applied Energy* 137: 567–75. doi:10.1016/j.apenergy.2014.07.021.

Motyka, M., B. Boufarah, S. Sanborn, and T. Stevens. 2015. "U.S. Renewable M&A Revs Up Despite Headwinds." *Electrronic Journal* 28: 80–9. doi:10.1016/j.tej.2015.05.006.

Munsell, M. 2015. "US Community Solar Market to Grow Fivefold in 2015, Top 500MW in 2020." *Greentech Media*.

Nair, N.K.C. and N. Garimella. 2010. "Battery Energy Storage Systems: Assessment for Small-Scale Renewable Energy Integration." *Energy and Buildings* 42: 2124–30. doi:10.1016/j.enbuild.2010.07.002.

Neubauer, J. and M. Simpson. 2015. "Deployment of Behind-the-Meter Energy Storage for Demand Charge Reduction." Golden, CO: National Renewable Energy Laboratory.

Newcomb, J., V. Lacy, L. Hansen, and M. Bell. 2013. "Distributed Energy Resources: Policy Implications of Decentralization." *Electricity Journal* 26: 65–87. doi: 10.1016/j.tej.2013.09.003

NYPSC. 2014. "Tariff Filing by Consolidated Edison Company of New York, Inc to Make Revisions to Its Demand Response Programs Rider S - Commercial System Relief Program and Rider U - Distribution Load Relief Program Contained in P.S.C. No. 10 - Electricity." Vol. 13-E-0573. Albany, NY: New York Public Service Commission.

Office of the Press Secretary. 2013. Presidential Policy Directive—Critical Infrastructure Security and Resilience.

Okkonen, L. and Suhonen, N. 2010. "Business Models of Heat Entrepreneurship in Finland." *Energy Policy* 38: 3443–52. doi:10.1016/j.enpol.2010.02.018.

Osterwalder, A. and Y. Pigneur. 2010. *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. Hoboken, NJ: John Wiley and Sons.

Palensky P. and D. Dietrich. 2011. "Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads." *IEEE Transactions on Industrial Informatics* 7: 381–8. doi:10.1109/TII.2011.2158841.

PwC. 2013. "Energy Transformation: The Impact on the Power Sector Business Model." London, United Kingdom: PricewaterhouseCoopers.

PwC. 2015. "A Different Energy Future: Where Energy Transformation Is Leading Us." London, United Kingdom: PricewaterhouseCoopers.

REN21. 2014. "Renewables 2014 Global Status Report." Paris, France: REN21 Secretariat.

Richter, M. 2012. "Utilities' Business Models for Renewable Energy: A Review." *Renewable Sustainable Energy Reviews* 16: 2483–93. doi:10.1016/j.rser.2012.01.072.

Richter, M. 2013. "Business Model Innovation for Sustainable Energy: German Utilities and Renewable Energy." *Energy Policy* 62: 1226–37. doi:10.1016/j.enpol.2013.05.038.

Rocky Mountain Institute. 2013. "New Business Models for the Distribution Edge: The Transition from Value Chain to Value Constellation." Boulder, CO.

Ruester S., X. He, J. Vasconcelos, and J.-M. Glachant. 2012. "Electricity Storage: How to Facilitate Its Deployment and Operation in the EU." Florence, Italy: European University Institute.

Schmalensee, R., V. Bulović, R. Armstrong, C. Batlle, P. Brown, J. Deutch, et al. 2015. *The Future of Solar Energy*. Cambridge, MA: Massachusetts Institute of Technology.

Schoettl, J.-M. and L. Lehmann-Ortega. 2011. "Photovoltaic Business Models: Threat or Opportunity for Utilities?" *Handbook of Research on Energy Entrepreneurship*, R. Wustenhagen and R. Wuebker, eds. Cheltenham, United Kingdom: Edward Elgar Publishing. 145–71.

Schoff, R. 2015. "Energy Storage: Technology Applications and Grid Integration." MIT Energy Initiatiative Associate Member Symposium: Storage, Renewables, and the Evolution of the Grid. Cambridge, MA: Massachusetts Institute of Technology. 1–16.

Seel, J., G.L. Barbose, and R.H. Wiser. 2014. "An Analysis of Residential PV System Price Differences between the United States and Germany." *Energy Policy* 69: 216–26. doi:10.1016/j.enpol.2014.02.022.

SEIA. 2015. "The Solar Investment Tax Credit." Washington, DC: Solar Energy Industries Association.

Shariatzadeh, F., P. Mandal and A.K. Srivastava. 2015. "Demand Response for Sustainable Energy Systems: A Review, Application, and Implementation Strategy." *Renewable Sustainable Energy Reviews* 45: 343–50. doi:10.1016/j.rser.2015.01.062.

Sharma A., K. Melkonyan, F. Xie, W. Wilkinson, J. Berg, J. Jin, et al. 2015. "Top Solar Power Industry Trends for 2015." IHS.

Siegrist, C.R., B. Barth, B. Campbell, B. Krishnamoorthy, and M. Taylor. 2013. "Utility Community Solar Handbook: Understanding and Supporting Utility Program Development." Washington, DC: Office of Energy Efficiency and Renewable Energy, Department of Energy.

SMA Solar Technology. 2012. "Technology Compendium 3.4: PV Grid Integration." Niestetal, Germany.

Smart Grid Task Force—EG3. 2015. "Regulatory Recommendations for the Deployment of Flexibility." Brussels, Belgium: Expert Group 3.

Speer, B. 2012. "Residential Solar Photovoltaics: Comparison of Financing Benefits, Innovations, and Options." Golden, CO: National Renewable Energy Laboratory.

St. John, J. 2016. "The Next Wave of Energy Storage Financing: Green Charge Closes $50M Non-Recourse Debt Funding." *Greentech Media*.

Steward, D. and E. Doris. 2014. "The Effect of State Policy Suites on the Development of Solar Markets." Golden, CO: National Renewable Energy Laboratory.

Teece, D. 2010. "Business Models, Business Strategy and Innovation." *Long Range Planning* 43: 172–94. doi:10.1016/j.lrp.2009.07.003.

Turitsyn, K., P. Sulc, S. Backhaus, and M. Chertkov. 2010. "Local Control of Reactive Power by Distributed Photovoltaic Generators." First IEEE International Conference on Smart Grid Communications. Los Angeles, CA: IEEE. 79–84. doi:10.1109/SMARTGRID.2010.5622021.

US DOE. 2015. DOE Global Energy Storage Database. Energy Storage Exchange. US Department of Energy.

von Appen, J., M. Braun, T. Stetz, K. Diwold, and D. Geibel. "Time in the Sun: The Challenge of High PV Penetration in the German Electric Grid." *IEEE Power Energy Magazine* 11: 55–64. doi:10.1109/MPE.2012.2234407.

Weiller C., T. Shang, A. Neely, and Y. Shi. 2015. "Competing and Co-existing Business Models for EV: Lessons from International Case Studies." *International Journal of Automotive Technology and Management* 15: 126–48. doi:10.1504/IJATM.2015.068543.

Wesoff, E. 2015. "Financing Energy Storage: Taking Cues from the Success of PV." *Greentech Media*.

Wiser, R., G. Barbose, and E. Holt. 2010. "Supporting Solar Power in Renewables Portfolio Standards: Experience from the United States." *Energy Policy* 39: 3894–3905, doi: 10.1016/j.enpol.2010.11.025

Zahedi, A. "Maximizing Solar PV Energy Penetration Using Energy Storage Technology." *Renewable and Sustainable Energy Reviews* 15: 866–70. doi:10.1016/j.rser.2010.09.011.

Zott, C. and R. Amit. 2007. "Business Model Design and the Performance of Entrepreneurial Firms." *Organization Science* 18: 181–99. doi:10.1287/orsc.1060.0232.

MAILING ADDRESS

MIT Energy Initiative
77 Massachusetts Avenue, E319-307
Cambridge, MA 02139

energy.mit.edu

MIT Energy Initiative

INSTITUTO DE INVESTIGACIÓN TECNOLÓGICA