Long-term reliability of generation in competitive wholesale markets: A critical review of issues and alternative options

Ignacio J. Pérez-Arriaga
Instituto de Investigación Tecnológica (IIT)
Universidad Pontificia Comillas, Madrid, Spain
http://www.iit.upco.es

IIT Working Paper IIT-00-098IT
June 2001

1. Introduction

The quality of service received by the end consumer of electricity, as the outcome of the global reliability performance of the power system, results from a chain of activities, where the most critical ones are generation, transmission, distribution and system operation.

The recent changes in the regulation of the electric power industry worldwide have modified the traditional reliability issues and approaches drastically. In the vertically integrated utility under cost-of-service regulation, reliability was seen as a major ingredient in the global exercise of centralized utility planning, at all levels: generation, transmission and distribution. In the traditional approach each activity, -i.e. transmission, distribution or generation-, was usually examined separately, but with the global objective of providing a reliable service at minimum cost.

Under the new market-oriented regulation, each one of these activities, because of their intrinsically different characteristics, -some of them can be performed in competition while others are natural monopolies that have to be regulated-, have been unbundled. For each one of the activities the new regulation must make sure that the appropriate economic incentives exist so that the quality of supply is maintained also at socially optimal levels.

A key feature of the new competitive environment for the power industry is that the customer must be at the center of any business strategy. As a final objective it has to be remembered that in a mature market the customer must have a saying also in reliability matters, mostly by making decisions on the reliability level of supply that he or she wants to receive and is willing to pay for.

This document only concerns reliability of generation, where the change has been more pronounced since in the new regulation generation is fully opened to competition. The document contains two major sections. The first one presents the major issues to be addressed when considering reliability of supply of generation. The second part describes the major approaches to this topic that have been either applied or seriously considered at international level. The document indicates the weak and strong points of each one of the approaches, but it does not provide any final recommendation.

2. The questions

Question # 1: Is there a problem?

The issue under discussion is whether the deregulated activity of generation of electricity in competitive wholesale power markets does or does not need regulatory intervention, -and if it does of what kind-, in order to provide a satisfactory level of reliability of supply. According to basic principles of economic theory, if there is scarcity in electricity supply, -or in the supply of any other commodity provided by a market-, the price of the commodity will increase enough to attract new investment, as well as to encourage more production from the existing plants, until the normal level of supply and prices is reestablished. Is there anything special with electricity, so that this basic economic scheme may not work satisfactorily in power markets?

Several features of the electric power industry have often been presented as a justification for a distinct regulatory approach regarding reliability of supply. In the first place, electricity cannot be stored, therefore the resources to produce it have to be ready whenever the demand requires them. Besides, electricity provision requires very large investments, the production plants take significant time to be installed and operational (this time has been very much reduced recently, with two to three years being presently a characteristic figure for the very popular combined cycle gas turbines) and have a long economic life (about thirty years or even more). Finally, electricity is an essential good, without an easy replacement in modern society; shortages of electricity have significant social and political implications, what makes politicians, regulators and system operators particularly aware of reliability of electricity supply.

The pioneering countries in designing competitive wholesale electricity markets, -Chile, England & Wales and Argentina-, opted for some type of regulatory intervention to promote reliability of supply, basically consisting of capacity payments under different formats. Other countries, -mostly in Latin America-, followed suit, while others –Nordic European countries, Australia or California-, relied totally on the market and adopted a policy of wait-and-see, although some of them are revising this policy, -leaving the Californian case apart-, in particular regarding incentives to maintain in operation existing peaking units or the promotion of investment in new ones. Other efforts have been addressed to improve the efficiency of the regulatory intervention by implementing it via market mechanisms. This is the case of the pools of PJM, New England and New York in the U.S.A., as well as the on-going regulatory revisions in Argentina or Colombia. On the other hand, England & Wales has abolished its particular approach to a capacity payment in the new trading agreements or NETA.

The experience with existing markets is still limited in time and does not allow drawing definitive conclusions on a topic such as reliability of supply, which requires a long time of observation. However, useful conclusions may already be obtained from some of the difficulties and mistakes that have been identified when implementing and applying the very diverse methods in existence today, as it will be shown below. And, as the recent experiences of electricity shortages in Chile, California or Brazil and some Eastern USA systems show, -as well as some near misses in Norway, Australia, Alberta and some Eastern USA systems and the growing concern in Argentina, Colombia, Peru or Spain-, lack of generation supply in restructured competitive wholesale markets is not any more a theory but a painful reality or at least a very uncomfortable threat.

Question # 2: Who has the ultimate responsibility for the reliability of supply?

Most people would agree that this is the regulator's responsibility. In particular in a topic as critical and immature as long-term guarantee of supply in generation, where the rules are currently under revision in most countries. The regulator must establish the rules and modify them in case the reality proves that they are not working properly. The rules may just rely on purely market forces, or they may establish some kind of capacity payment or an ad hoc capacity market; they may even transfer to the distribution & supply companies, -as it has been done in Argentina-, the immediate responsibility of the reliability of supply. But, in any case, the regulator must always supervise that each piece of the regulatory framework is working properly and it is the regulator who has to introduce the appropriate changes when necessary.

This role of the regulator may be understood in two different ways, which very much condition the specific regulation of each country:

- One line of thought is that the regulator must intervene just to neutralize specific market shortcomings (such as lack of consumer response, risk aversion of potential investors in generation because of large volatility in market price, or inadequate regulation that eliminates the incentives for consumers or other "load serving entities" to purchase at the lowest possible price or to hedge against high market prices) with the minimal *ad hoc* measures that respond to these deficiencies.

- The other major line of thought is that the regulator must actively introduce the appropriate regulatory measures that guarantee a prescribed level of quality of service in generation supply, where this level is set by the regulator himself, therefore acknowledging that the market, -because of some unavoidable imperfections-, is not able to achieve what the regulator considers to be a satisfactory level.

Regulators throughout the world have opted for one solution or the other, very much depending on the specific characteristics of their systems, such as the volume and uncertainty of the hydro component in the generation mix, or the existence or not of excess generation capacity when the market was established, as well as their preference for purely market based or more administratively oriented regulatory solutions. The borderline separating both approaches is fuzzy, thus, in some instances it is difficult to classify one particular approach into one or the other group.

Question # 3: Is reliability a long-term or a short-term issue?

The National Electric Reliability Council in the U.S.A. defines **reliability** as “the degree to which the performance of the elements of the electrical system results in power being delivered to consumers within accepted standards and in the amount desired”. Therefore, the ultimate measure of the reliability of the generation activity is the level of quality of supply provided to the load by generation at the wholesale level. Although the quality of supply only materializes in real time, its provision encompasses a number of deregulated activities that have to be performed in different time ranges, from several years to seconds, such as investment in new facilities, scheduled plant maintenance, fuel acquisition and management (particularly of hydro resources) and provision of operation reserves of different types (cold, tertiary, secondary and primary reserves).

Any serious regulatory approach to reliability of generation supply, needs to address all and each one of these dimensions. For the sake of simplicity, it will be termed **adequacy** issues all questions related with the *existence of enough installed available capacity of the appropriate characteristics to be capable of meeting the estimated demand* –i.e. investment in new facilities and retirement or life extension of the existing ones, as well as long term operating decisions affecting the availability of a unit to be a part of the operation of the system at a given time-. **Security** issues are all questions concerning the *readiness of the installed available capacity to respond to the different operation requirements that allow the system to meet the actual load*. Reliability comprises both adequacy and security.

Adequacy and security are therefore distinct products that are provided by the generators, although both necessarily converge on the actual quality of service that generation provides in real time. It seems that some kind of consensus is being reached regarding *security* in generation. The design of most recent wholesale electricity markets contemplates the existence of ad hoc markets for the provision of certain quantities of the several required kinds of operating reserves. These quantities are mandated by the Independent System Operator (ISO). This scheme appears to be working satisfactorily and it can be considered to be a hybrid between the purely market and the interventionist approaches.

On the other hand, *adequacy* in generation is still a very open issue, even conceptually, and no consensus exists on the best method to approach it. The regulatory mechanisms that will be discussed in this document mostly address adequacy, although they also have some security implications, since the interplay between both cannot be ignored.

The distinction between security as a short-term issue and adequacy as a long-term one does not imply that only long-term economic signals are involved in the provision of a desired level of adequacy. As it will be seen below, the different possible mechanisms for the provision of adequacy may generate both short and long-term economic signals. For instance, capacity payments are long-term economic signals, as they are typically determined once a year for each generator. However, it is impossible to separate these payments from the actual operation, for a number of reasons: dependence of the payment on the availability of the unit or the design of the mechanism of allocation of the payments. Alternative approaches to obtaining adequacy may use long-term auctions for the desired adequacy product, with the possibility of short-term clearing mechanisms or incentives to deliver the product. In these cases both short and long term signals are generated in the provision of adequacy, and both are equally relevant. Even if adequacy is entirely left to the market, the specific adopted procedures of computation of the market prices, -e.g. the existence of a price cap, or the interference that mandatory quantities of operating reserves may have on market behavior-, will have a definite influence on the long-term investment decisions of potential new entrants.

Question # 4: Are the revenues obtained from short-term marginal prices enough to cover the total (fixed & variable) costs of all generators, the peaking units in particular?

In a competitive market where demand responds to prices, microeconomic analysis of a power system shows that, -in the absence of economies of scale in generation-, the resulting market price is sufficient to remunerate the total costs of those generators whose investment is well adapted to the existing demand and to the existence of the remaining generation plants. This complete cost recovery condition applies to all generators if each one of them meets the preceding condition.

Note that the peaking units, -i.e. those with the highest variable costs-, do receive a revenue that allows them to recover their fixed costs, since these generators may bid a price above their variable cost when no other less expensive available units may displace them. This price will be as high as the consumers will permit by reducing their consumption until the usual equilibrium of supply and demand is reached.

This ideal situation where supply and demand always reach an equilibrium and therefore define an unambiguous market price, which results in complete cost recovery of all well adapted production units, requires a number of conditions to be met, which is usually not the case in practice:

- Elasticity of demand to market prices –which requires that prices must be perceived by demand in real time- and the participation of the demand in the determination of these prices.
- When the preceding condition fails, at least it is required the existence of adequate pricing mechanisms to be applied in the event that the market fails to provide enough supply to meet the demand. Note that the existence of price caps for the market prices and the value of these price caps will affect considerably the income of the peaking generating units.

- Correctness of the mechanism of determination of the market price, in particular the possible influence of mandatory levels of operating reserves in depressing the energy prices.
- The correct allocation of economic risks to buyers and sellers that is implicit in a specific market design. This allocation will affect the use of financial hedging mechanisms and will influence the behavior of risk averse potential investors.

Question # 5: How to allow the consumers the freedom to choose their desired level of reliability of supply?

One can think of an advanced power system where each consumer, when the price in the wholesale market is high enough, disconnects his load either partially or totally. Alternatively, this conceptually could be done by the utility, depending of the type of contract that the customer has signed[1].

Another possibility, which avoids the difficulty of having the means to physically discriminate in the curtailment of loads, is that the consumers may sign financial contracts (options) to protect them against the risk of market prices above a certain value. If the market price cap is eliminated, the other party in the option contract (a generator typically) will be exposed to a large economic risk if it does not have the energy to supply the consumer. Therefore an economic incentive exists for the generator to be available when needed. However, this scheme renders the consumer inelastic to prices above the strike price of the option contract. If most of the elastic demand has signed these contracts, the possibility exists of market breakdown. However, the generators exposed to economic risk (the unavailable ones) would be willing to pay the consumers for being disconnected and these will accept if paid their cost of non supplied energy. This will result in voluntary disconnection and the avoidance of forced curtailment[2]. Of course, this approach requires putting in place all these trading mechanisms, which is not the situation right now.

An additional difficulty is the coexistence in many systems of qualified consumers, -those that can choose supplier and negotiate the terms of the contract, including quality of service-, and captive consumers, who have to accept the conditions of the regulated tariff. If the regulator imposes a level of quality of service for the captive demand (e.g. by creating a mandatory market for capacity, just to cover the captive demand), there is the possibility that some of the qualified consumers decide to “free ride”, i.e. to benefit from the guarantee of supply that the generators serving the captive demand provide anyhow to the entire system, therefore avoiding to sign agreements with the remaining generators to secure a firm supply and reasonable prices. If this is

---

[1] Note that the capability of the utility to selectively disconnect consumers, -except for the very large ones- is usually very limited.

[2] It may be argued that curtailment based on economic preferences will end up with well-off consumers having a better quality of service than less affluent consumers. This is true with every other commodity, but it may be politically difficult to accept for the supply of electricity, which has been considered a public service for so long, except for medium and large consumers. The regulator might act on behalf of the small customers, determining a suitable value for the strike price of the option for all of them.

the case, it would be necessary to establish the economic conditions under which each consumer may use the electricity produced by a generator with which he has not signed any contract.

Question # 6: Which is the product that the generators are committed to deliver regarding guarantee of supply and how can the individual responsibilities be established precisely?

The problem with most of the existing regulatory approaches to the provision of generation adequacy is that they fail to satisfactorily specify the product that the generators have to deliver in exchange for some kind of economic compensation, let it be determined administratively or through a market mechanism. The critical point here is that long-term commitments by generators must refer to their availability in real time conditions, when there is an actual threat of lack of generation supply.

3. The answers

Conceptually, one may classify the existing or proposed regulatory options with regard to generation adequacy in competitive markets into the six broad categories that are described below.

Option # 1: No regulatory intervention or "leave it to the market"

Description:
No further description is needed.
This approach has been used in California where, coupled to serious market design flaws and other circumstances, has resulted in serious shortages of generation. It is also being used in Australia, where generation shortages have been avoided so far and no specific measures are being considered immediately, and in the Nordic European countries (the NORDEL group) where there is a growing concern that the initial large surplus of generation capacity is dangerously dwindling. In fact, Sweden and Finland have already started to apply option # 3, which has been also considered by Norway.

Motivation:
Obviously it is the simplest of all options. It follows the basic principle of minimizing the interference with the market.

Implementation:
In principle there is no need for any implementation. However, strictly speaking, there are issues such as determining the participation of demand in the market, establishing price caps or the design of risk allocation that constitute fundamental regulatory decisions with an influence on the final level of reliability that the actual market will achieve. Consistency with the "leave it to the market" approach would require the elimination of any price caps, allowing full participation of the demand in the market and letting each market agent to fully experiment the volatility of the market prices. No existing market is presently operating under such rules.

Evaluation:
This approach ignores the existence of failures in actual markets. The two most important ones are the passivity of a significant part of the demand and the risk aversion of investors, when they must face a volatile or insufficient income, as indicated in question # 4 above. These market failures may result in undesirable adequacy levels, with episodes of electricity shortages and high market prices that consumers would have gladly paid in advance some extra money to avoid.

The policy of no regulatory intervention would work perfectly, however, in an ideal and mature market. In this market all the demand and all the generation would in principle be fully exposed to the volatility of the market price, which would not be limited by any price cap. Under these circumstances, the rational response of the demand would be to find some way of hedging against very high prices. Strictly speaking shortages would never happen, as each individual demand would voluntarily disconnect once the price reaches a certain value, which is different for each block of load of each customer. Generators would be the natural counterparts in these risk hedging contracts, since the rational behavior of generators would be to hedge the risk of the volatile revenues and to obtain a more stable revenue instead.

These price risk hedging contracts could adopt many different formats, since they would be freely agreed between any generator and any electricity purchasing entity. Probably the reference format for these contracts would be something similar to the reliability option contracts that are described as option # 5 below.

As indicated before, the passivity of the demand, -due to different reasons-, renders this scenario of generalized hedging contracts an unrealistic one. The generators, without their natural contracting counterpart will not be able to hedge their risk and risk aversion of the potential new investors may result in a shortage of generation.

Option #2: Capacity payments

Description:
Pay some administratively determined amount to each generator in proportion to its estimated (with a mathematical model, using historical operation data to asses the “firm capacity” of each unit or by some other means) contribution to the reliability of the system. The amount to be perceived by each generator is computed in advance annually, but the payments can be performed on a monthly or even daily basis. Once in actual operation, a generator only receives the corresponding daily payment when it is available on that day.

This method was implemented in Chile for the first time in 1981 and, with different variations, -mostly related to the criteria and the algorithm that are used to determine how much money should be paid to each generator-, it is also used in several Latin American countries, such as Argentina, Colombia, Peru, Brazil and some others in Central America and also in Spain. The ex ante method of computation of market prices in the pool mechanism of England &

Wales (before it was changed to NETA) indirectly also results in a capacity payment, by overestimating the probability of loss of load for the next day[3].

Motivation:

There are two principal motivations for this approach, which could be considered either independently or jointly. The first motivation for capacity payments is to partly stabilize the volatile income of generators, in particular of the peaking units, by acknowledging the loss of remuneration that results from the existence of a regulated low price cap for the market price and compensating the corresponding loss of income of generators by a stable remuneration of the available generating capacity. This may be considered a "soft" regulatory intervention. The second motivation is to directly promote an extra level of generation adequacy by establishing a capacity payment for the available generation capacity, -and correspondingly an extra cost for the demand-, with the purpose of stimulating new investments and discouraging early retirement of otherwise unprofitable units. This is a "harder" regulatory intervention.

Implementation:

The practical implementation of this approach may vary. In general, the global amount (or the per unit value) of capacity payments is established annually on the basis of the prescribed price cap value (first motivation) and/or of the target adequacy level (second motivation). Then, the payments to the individual generators are broadly based on some estimated measure of their a priori contributions to the system adequacy for the considered year. The final payments are conditioned to the actual availability of each generator. This mechanism has a theoretical justification: In the same way that it can be demonstrated that market prices fully pay for the total (fixed plus variable) costs of the well adapted generation mix in a perfectly competitive market, it can be also proved that, for any given extra amount of desired generation adequacy above the one provided by the market, there is a capacity payment that is due to each generator as a function of its theoretical contribution to the reliability of the system.

Evaluation:

This capacity payment approach normally succeeds in providing a fairly stable economic signal to generators (but not always, as this depends on the adopted algorithm for allocation of the capacity payments), but it has some important drawbacks. First, the economic signal, if not carefully implemented, may introduce distortions in the generators' behavior in the short-term market, as it has been mostly the case in Argentina. Second, it is very complex to find a convincing way of determining the volume of the payments and of allocating them to the different generators. This has been the source of endless disputes among generators, -particularly between hydro and thermal generators-, and with the regulator. Finally, and most important, this approach lacks the definition

---

[3] Strictly speaking, the former method of price computation in the pool of England & Wales was a purely marginal market price computed ex ante and, therefore, it contained a term that accounted for the probability of the pool price being equal to the cost of non served energy. However, this probability was grossly overestimated therefore resulting in an extra payment to generators.

of an identifiable commercial product for which the generators obtain their remuneration. Therefore, there is nothing tangible that is really being paid for or that can be traded, there is no specific commitment from the generators' side, and the desired level of adequacy cannot be guaranteed.

Option #3: Purchases of peaking units by the System Operator

Description:
The System Operator, following the instructions of the regulator, purchases a certain number or all of the peaking units of the system. Depending on the circumstances, the System Operator may also commit to purchase new peaking units. The incurred costs are charged to the consumers, as an extra cost associated to reliability enhancement. Purchasing of peaking units is currently being used in Sweden and Finland.

A scheme that has been considered in the design of the Italian wholesale market is closely related to the purchase of peaking units. The idea is to allocate, -by means of an ad hoc auction-, some capacity payments to a subset of the generation units, the so called "strategic capacity reserves", which meet some specific requirements: to have an aggregated capacity that is sufficient to cover the expected demand of peaking capacity in conditions of stress for the system, and to belong to the type of technology that provides capacity at the margin in the particular system: typically peaking or very old and inefficient units. After winning the auction these units happen to be under full operational control of the System Operator, who will use them to cover any deficits of generation that may appear while trying to meet the system demand.

In England & Wales, under the new trading arrangements that are known as NETA, the System Operator can purchase in advance any desired amount of operating reserves. By forcing this mechanism, it is possible to have some impact on generation adequacy. Norway is also presently securing long-term operating reserves, but only with one year of anticipation.

Motivation:
This approach is meant to avoid that the generation units that provide capacity at the margin, -peaking units typically, but maybe also old and/or very inefficient units-, may decide to leave the market when their revenues are too volatile or they are insufficient (perhaps only transitorily) to cover at least their total operating costs. The approach also seeks to attract new investments in peaking generation capacity.

Implementation:
Purchasing peaking units does not require further explanation. The Italian scheme of "strategic capacity reserves" requires running a specific type of auction. In both cases the System Operator is free to use these units in the way it considers most appropriate to enhance the reliability of the power system.

Evaluation:
This approach, as well as the option # 4 below, is strongly interventionist and may seriously interfere with the proper functioning of the market. The

generators non belonging to the System Operator, as well as the potential new entrants may judge that the market prices depend too much on the purchasing decisions that are dictated by the regulator.

An additional concern with the proposed scheme of strategic capacity reserves (which otherwise seems to be well adapted to the specific problem that it is addressed to solve) is the breakdown of the market into two different parts: the competitive market and the units under the control of the System Operator as strategic reserve units. This approach may be suitable for a limited amount of time, -for instance while some initial surplus of capacity disappears because of the natural demand growth-, but the separation of the market between two parts may result in undesirable patterns of behavior of the agents. Note, for instance, that market power grows much when there is scarcity (California has provided a good example of this principle). Removing some plants from the market will increase much the possibility of artificial scarcity, therefore facilitating the possibility of very high bids by the units that remain in the market. This difficulty may be overcome by requiring the System Operator to bid at all times in the spot market with the strategic reserve units at the variable cost of these units.

Option # 4: Regulatorily determined competitive bidding

Description:
Although free entrance of new generation is allowed into the system, the regulator or some other administrative authority supervises that there is no threat of insufficient generation adequacy, according to some pre-established criterion. If it is considered that there is a lack of entry of new generation, this authority may start a competitive bidding process for the addition of the required extra generation.

This approach was adopted by countries as France or Portugal in the implementation of the European Directive 96/92/EC on the Internal Electricity Market. It is interesting to note that the proposal of a new European Directive to amend Directive 96/92/EC that was presented at Stockholm by the European Commission in April 2001, explicitly forbids the tendering procedure to acquire new generation, except when used as an exceptional measure for reasons of security of supply. The proposal of Directive also establishes a double surveillance of generation adequacy, both at Member State and European Union level. The Green Paper of the European Commission “Towards a European strategy for the security of energy supply” of November 2000 emphasizes this same position. This scheme is no doubt much influenced by the recent California crisis.

Motivation:
Lack of confidence that the market may provide enough generation adequacy by itself and reluctance to make use of other more indirect or market-based mechanisms to encourage new investment.

Implementation:
At least some kind of reliability model, -even if it is very rudimentary-, is needed in order to check whether the generation adequacy that is estimated that the

market will provide meets the prescribed criterion or not. More sophisticated generation planning models may be used to determine the volume and type of generation technology that would be optimally required. Once the volume (and maybe the technology) of additional generation has been determined, a competitive auction can be organized to determine who will build it.

Evaluation:
The evaluation of option # 3 is also valid here. This approach is strongly interventionist and may seriously interfere with the proper functioning of the market. The market agents, as well as the potential new entrants, may judge that the market prices depend too much on the purchasing decisions that are dictated by the regulator.

Option # 5: Capacity markets
Description:
Every year each purchasing entity (e.g. large consumer, retailer, trader, etc.) is required, -under the threat of a heavy fine-, to purchase enough firm generation capacity to cover its expected annual peak load plus a regulated margin. The regulator determines the amount of firm capacity that each generation unit can provide. The generators must have the committed firm capacity available whenever they are required to produce.

Motivation:
The motivation for this approach is to guarantee a regulated generation adequacy level for the system by defining specific commitments of purchase of firm production capacity to all the consuming entities. The approach also tries to specify the commitment that each generation unit must maintain in real time, i.e. the commercial product associated to the concept of generation adequacy.

Implementation:
The practical implementation of the approach consists of requiring mandatory levels of contract coverage of firm generation capacity to al consuming entities in the system: qualified consumers and any retailers purchasing power on behalf of either qualified or captive consumers. The level of mandatory coverage (e.g. 15% above the estimated annual peak load of the consuming entity) may be proposed by the System Operator and authorized by the regulator. Interruptible load may also qualify as a form of providing firm capacity to the system. The transactions between buyers and sellers of firm capacity may be facilitated via organized long-term auctions, -from months to one or more years ahead of real time-. The committed capacity has to be available at the time of delivery, or otherwise it will be subject to a heavy fine. However, the commitments of firm capacity may be traded in the short-term, if there is uncommitted capacity still available. This approach has been adopted in PJM and other regional entities of the Eastern USA, such as the ones of New York and New England. The Secretary of State for Energy in Argentina has considered the possibility of abandoning the capacity payments and adopting some variation of the capacity market scheme.

Evaluation:
In capacity markets there is an identifiable commercial product associated to generation adequacy and a commitment by the agents to purchase and to deliver the product, although the concept of firm capacity and the conditions of delivery remain somewhat ambiguous, particularly when hydro units are involved. The regulator determines the total amount of desired firm capacity and also the rules (which cannot avoid some degree of arbitrariness) to determine the firm capacity that is provided by each unit. No doubt these rules could be easily contested in case of coexistence of thermal and hydro units of varied reservoir capabilities.

Consumers remain fully exposed to the potential high prices in the energy market. Also, the commitment of the generators to supply power at times of need is not precisely defined (for instance, how is defined the obligation, -the availability- of a hydro unit with storage capacity?).

Market mechanisms determine the price of capacity, which may be very volatile, depending on the tightness of the margins of installed capacity over the system peak load and the anticipation of the auctions with respect to real time. The volatility may be reduced by increasing the time horizon of the auctions. Otherwise, the uncertainty in the remuneration of the generators may not gain much in terms of stability with this approach.

A modification of the standard design of this method could allow the consumers to freely choose their reliability level. The modification simply consists of leaving the qualified consumers the freedom to choose whether to contract firm capacity and, if so, how much. Consumers without contracted firm capacity would have less priority of supply in case of a shortage of power. The problem with this modification is that it opens the way to free riding by those consumers that do not contract firm capacity but enjoy a good quality of supply since it is paid by the remaining consumers. Besides, except for large consumers, the utility will not be able in general to cut the supply selectively to the consumers.

Option # 6: Reliability contracts

Description:
An organized market is established where the regulator requires the System Operator to purchase a prescribed volume of reliability contracts from generators on behalf of all the demand. This volume must be such that a satisfactory level of generation adequacy is obtained, according to the rules presented below. The reliability contracts allow the consumers to obtain a price cap on the market price in exchange for a fixed remuneration for the generators. Additionally the consumers obtain a satisfactory guarantee that that there will be enough available generation capacity whenever it is needed. Otherwise the generators will be penalized. The generators are also compensated economically for this service: the higher the contribution to the reliability of the system, the higher the compensation will be (this is automatically built-in in the procedure, it is not an administrative decision).

Motivation:
The motivation for this approach is multifold. It is recognized that somebody (the regulator) must act on behalf of the demand, specify the desired generation adequacy level and establish contracts with the generators. The objective is that consumers can obtain a well defined commercial product in exchange for their money. The method limits the risk for consumers, by capping the market price at a value much lower than the typical value of non served load. Besides, the method guarantees a regulated level of generation adequacy and creates a strong incentive for the generators to have available their committed capacities whenever the market price exceeds a threshold value. As far as the generation side is concerned, the risk of generators is also reduced, -in particular for the peaking units-, by providing a stable income in exchange for the most volatile fraction of their revenues (i.e. the very high but also very volatile revenues when there are spikes in the market prices).

Besides, market mechanisms are now used to determine the price to be paid to the committed capacity and also the value of each generator's committed capacity, since how much capacity to bid is now each generator's decision.

Implementation:
The key point of the approach is the precise specification of the mutual obligations of the consumers and the generators by means of a particular type of option contract. The implementation of the method comprises the following steps:

- The regulator (perhaps with the technical support of the System Operator) specifies an adequate level of required available capacity Q for the system, ready at any time when it might be needed, for an entire year, several years in advance (e.g. 3 years), so that new entrants may have the opportunity of bidding in the auction of reliability contracts.
- The scheme is based on the following type of contract for any generating plant k:
    a) for a future period of time T, e.g. 1 entire year, 3 years from now,
    b) for a certain capacity Qk,
    c) the plant must be producing in the day-ahead market the contracted capacity Qk whenever the pool price P exceeds a value S (strike price) that is defined by the regulator. Otherwise the plant must pay a penalty PEN, also defined by the regulator, per MW of production below the committed level Qk.
    d) the plant k (and all other committed plants) receives a fixed payment PAY (€/MW) for the contracted capacity Qk during the time period T of the contract,
    e) the plant k must pay back the amount (P-S).Qk whenever P>S.
- These contracts will be assigned in a competitive auction, where a participant generation plant k will bid the quantity Qk to be contracted and a value PAYk of requested remuneration.
- The counterpart in the contracts will be System Operator. Any net payments or benefits will be charged to consumers.
- All winning plants in the auction will be paid according to the marginal price PAY of the auction.

- Only physical generators can bid in the auction, i.e., no purely speculators are allowed to bid. Besides, economic guarantees would be required from generators, as in any other organized market.
- The price cap of the market may be eliminated or it may be set very high (note that the demand is totally hedged against prices above S).

Evaluation:

Strong points of the method are:

- There is no need to evaluate the firm capacity of each generator administratively, as each generator bids the amount (or amounts, if it decides to bid several blocks at different prices) and price that it considers to balance the loss of revenues and the risk of penalization with the stable revenue of the contract premium.
- The premiums of the reliability contracts are also determined by the market.
- A prescribed level of generation adequacy is attained and, besides, the committed generators have strong incentives to be available at those times when they are needed.
- The consumers are fully protected from any high prices in the energy market.

The method has also some weak points, most of them also shared by the alternative procedures that have been presented above:

- The method does not promote an active demand response when the market price exceeds the strike price S of the reliability contracts.
- The remuneration of capacity (the premium fee of the auction) may be volatile. This volatility may be reduced by extending the number of years of anticipation of the auction.
- The mechanism may not be enough to attract new entrants, if each auction only covers a time horizon of one year. At least it must be admitted that the methods is an improvement over the do-nothing situation with respect to new investments. The incentive for potential new entrants can be increased by extending the time horizon of the auction over a number N of years. The problem with this is that it forces to run the auction once every N years or to run it every year for 1/N of the total target capacity. This last solution may aggravate any existing market power problem in the capacity market.
- As any other market-based mechanism, it has the potential for market power abuse if there exists a significant level of horizontal concentration.

4. Other topics

Two relevant topics will be briefly commented below. The first one is market power and how to deal with it when applying market-based mechanisms in the provision of capacity adequacy. The second one is the need for coordination of any schemes for the provision of capacity adequacy in a regional market.

4.1. Market power

Only those approaches to generation adequacy that rely on markets, namely: “leave-it-to-the market”, “capacity markets” and “reliability contracts” may suffer from market power related problems. Fully regulated solutions, such as "capacity payments", "purchase of peaking units" or "regulatorily determined competitive bidding" are immune to market power.

The reaction of oligopolistic agents to a "leave-it-to-the market" scheme is to withdraw capacity and to raise the energy prices. Without any barriers to enter the market, the oligopolistic situation would be contested soon by new entrants, and the effects of market power would be gradually reduced. But in actual markets some entry barriers always exist, so a level of generation adequacy somewhat below the one that a perfectly competitive market would provide is to be expected in an market with some horizontal concentration.

The markets of capacity in options # 5 ("capacity markets") and 6 ("reliability contracts") can be subject to the exercise of market power if significant horizontal concentration exists. Here, an approach will be commented to handle transitorily the situation within the context of the "reliability contracts" approach. A similar approach to the "capacity markets" method is also possible.

The proposed modifications to the "reliability contracts" method make it possible: a) to capture most of the benefits of the original approach; b) to reduce market power significantly; c) to have a smooth transition from this initial scheme to the fully competitive one, as the conditions allow it to happen. The proposed scheme, while significant market power in generation exists, introduces the following modifications with respect to the original permanent approach described as option # 6 above:

- The regulator determines, as before: a) the total amount of required available capacity Q; b) the reference or strike price S for the contracts; c) the value of the penalization PEN. But now it also determines: d) the value of the contract fee PAY to be paid to all MWs of contracted capacity; e) the initial value of the contracted capacity Qk for each plant k (the plant may be divided into several capacity blocks, with each one of them being separately treated).
- Each plant k may modify its initially assigned contracted capacity Qk by participating in an organized exchange (run by the Market or the System Operator) of increments and decrements about the Qk of all plants who want to participate.
- The value of the Qks must be determined in some non-sophisticated way, since there is the option of going to the organized exchange. Some rough measure of availability (firm capacity) could be utilized to allocate Q to all existing plants. A fraction of Q should be left open for the organized market, in order to allow the possibility of bidding by new entrants.

4.2. Regional markets

There is a generalized trend towards the organization of commercial rules encompassing more than a single centrally-operated system[4]. The Internal

[4] Frequently the power system under the control of a single system operator coincides territorially with a country. This is for instance the case of Italy, Argentina or New Zealand. However, Germany has six system operators, Switzerland has three, and in the USA or Canada a system operator may just cover one state, as it is the case of California or Ontario, which on the other hand are larger than many European countries.

Electricity Market of the European Union, Mercosur in South America, the Central American Electricity Market or the Regional transmission Groups in the USA are some representative examples. The basic commercial rules that allow the functioning of regional markets concern open access to the regional transmission network, transmission tariffs for cross-border trade, management of congestions and other network constraints, harmonization of the categories of consumers who are free to choose supplier, independence of the system operators from the agents of their systems, etc.

A topic that has been mostly ignored so far in the initial rules of regional markets is harmonization of the mechanisms, if any, to provide a satisfactory level of generation adequacy. Only in the design of the rules of the Central American Market has been this issue discussed at some length. Within the Internal Electricity Market of the EU the need for harmonization was first raised by the Portuguese and Spanish regulators in exploratory meetings to define an Iberian Electricity Market. Only recently, as an aftermath of the Californian crisis, the issue has been addressed in a very preliminary fashion by the proposal of a new European Directive and in the Green Book on security of supply, as indicated before.

With respect to generation adequacy, the approaches in the 15 EU countries are widely varied and there is no perspective of harmonization in this respect. Most countries have not created a competitive wholesale market and they still may rely on centralized planning procedures combined with competitive tendering for the construction of new generating plants. This is for instance the case of France and also partly of Portugal. Other countries with markets, such as Norway, Sweden or Finland have initially adopted a “leave-it-to-the-market” approach, although Sweden and Finland have recently chosen a mechanism of regulatory intervention to keep enough peaking units in operation, consisting in the purchase of single cycle gas turbines by the System Operator, and Norway is reconsidering its initial position as the initial large margin of installed capacity over peak demand is becoming smaller. Regulated capacity payments are used in Spain and, in an ad hoc manner also in the former pool arrangements of England & Wales, which have recently replaced by NETA, which belongs to the “leave-it-to-the market” family. The German and Dutch markets do not seem to have adopted any regulatory mechanism regarding generation adequacy.