IIT Working Paper IIT-00-0078-A. First version: March 2001; this version: September 2001.
IEEE Transactions on Power Systems, Vol. 17, No. 2, pp. 349-357, May 2002.

A market approach to long-term security of supply

Carlos Vázquez   Michel Rivier   Ignacio J. Pérez-Arriaga
Member, IEEE   Fellow, IEEE

Instituto de Investigación Tecnológica (IIT)
E.T.S.I. ICAI. Universidad Pontificia Comillas, Madrid, Spain
http://www.iit.upco.es

Abstract: The problem of ensuring that there is enough generation capacity to meet future demand has been an issue in market design since the beginning of the deregulation process. Although ideally the market itself should be enough to provide adequate investment incentives, there are several factors that prevent this result from being achieved, and some actual markets have already experienced problems related with a lack of generation capacity. A regulatory framework to address this question is presented. The procedure is based on an organized market where reliability contracts (based on financial call options) are auctioned, so both their price and their allocation among the different plants are determined through competitive mechanisms. This results in a stabilization of the income of the generators and provides a clear incentive for new generation investment, with a minimum of regulatory intervention. Additionally, the method represents a market-compatible mechanism to hedge demand from the occurrence of high market prices.

Keywords: Capacity markets, capacity payments, long-term guarantee of supply, generation adequacy, wholesale market design, electricity markets.

I INTRODUCTION

The most fundamental characteristic of the restructuring process that is taking place in numerous countries around the world is that market mechanisms have replaced the highly regulated procedures that were used in the decision-making processes under the traditional regulation. In particular, within a liberalized market there are no mandatory expansion plans that determine which generation units have to be installed in the system and when. Instead, market participants decide on their own, according to their business expectations, whether they want to build a certain facility or not. This is the reason for the regulators' concern about whether there will be enough installed capacity to meet demand in the long term. This has always been a contentious issue in market design since the beginning of the liberalization process [1].

Although economic theory [2], [3] shows that the spot market itself is enough to provide adequate investment signals and that the rational response of the generators to the prices in the spot market is to install facilities until the socially optimal level of investment is reached, some practical difficulties arise in real markets, from the generators' viewpoint, that prevent the straightforward realization of these theoretical results. The first one is the frequent application of price caps, which limit the revenue that generators can extract from the market and discourage investment.

The second one is the existence of risk aversion in the potential investors in new generation. This can be illustrated by the case of a potential peaking unit that would only generate a few hours a year, when the market price is highest. As a consequence, the peaking generator would receive no remuneration most of the time and a high income in a few occasions. Assume that the expected average income of this potential generator is enough to recover its variable and capital costs with a reasonable profit margin. Then, although in theory the decision of investing in this peaking unit would be economically justified, a risk-averse firm would feel that the high volatility of the generator's income renders the investment very risky, and consequently the firm would turn down the opportunity to invest. If either the regulator or the consumers want this unit to enter the market, they will have to diminish the risk through the provision of some extra income stability, or to increment the generator's expected income through some additional source of revenue (e.g. a capacity payment).

A third difficulty may be caused by an oligopolistic behavior of the incumbent utilities, which may under-invest in order to raise the market prices, when the barriers to entry are sufficient to block the contestability effect of potential new entrants.

Difficulties also exist from the consumers' side. In an ideal market, consumers seeking a better reliability would try to sign long-term contracts to protect themselves against high prices and service interruptions. This should encourage the entrance of more generation, since the contracts would bring the income stability that the generators are looking for. The level of reliability in the system would finally depend on how much capacity the consumers would be willing to commit in long-term contracts and how much they are willing to pay for their security of supply.

However, actual markets are not performing so efficiently. In most cases, the consumers are isolated from spot prices either by regulated tariffs or by average tarification procedures, so they do not feel the need for hedging against high prices and they see no advantage in long-term contracting. Even in those markets where demand is really exposed to the spot prices, long-term contracts are not taking place either. Most of the consumers are not mature enough to realize the risks involved and they tend to make their decisions using only very short-run criteria. This lack of demand-side response creates a malfunctioning of the long-term market that cannot be solved in the short run, and causes a lack of generation investment that paves the way for potential future shortages. It should be noted that the basic consumers' response that is needed is not demanding less energy from the market when prices are high -that is the typical goal of demand-side management programmes- but signing efficient hedging contracts to express their risk aversion and their need for a higher reliability level.

Three types of regulatory approaches have been employed or proposed to deal with this problem: a) leave it to market forces, and therefore wait for the consumers to learn and to perform efficiently in the long-term market; b) establish some regulatory mechanism of payments to encourage new investment and, in some cases, to stabilize the volatile income of generators; c) impose an obligation on the buyers, forcing them to trade in a long-term market some kind of capacity product, so that a prescribed level of generation adequacy is ensured.

The most orthodox solution to this problem would be the first one [4], [5]. If nothing specific is done, consumers would end up realizing the need for protecting themselves against high prices and blackouts, probably after experiencing a few of them. Unfortunately, this would probably be a long and painful learning period, which may include several rationing episodes and, according to what we have seen until now, it is likely that it would be considered more a market design problem than a problem with the consumers that are not hedging efficiently. Most likely, the market rules would be dramatically changed before consumers have had time to complete their learning process. This is therefore a grim outlook, both for the consumer and for the market.

Some of the recently liberalized systems (e.g. Australia, California, Scandinavia) have ignored the generation adequacy problem in their market designs, either because at the time competition was implemented there was an important overcapacity in the system that made irrelevant the threat of rationing, or because they believed the market could solve this problem by itself, mainly through the use of long-term contracts. Unfortunately, this approach has performed poorly in real markets and has lead to a lack of investment and price spikes in some systems (California [4], [5], Midwest USA). In some other cases (Finland, Norway, Australia), highly interventionist measures are being considered, such as requiring the System Operator to purchase peaking units, that otherwise would be mothballed or simply retired, after the realization that market prices were not providing the proper incentives, apparently.

The second approach to the problem, administrative payments [3], [6], has been followed in several countries, such as Argentina, Colombia or Spain, which have adopted an explicit remuneration for the installed capacity as an economic signal intended to augment the volume of installed and available generation. In theory, capacity payments would attract new investment, resulting in lower (and more stable) market prices, with this price reduction being compensated by the capacity payment itself. The regulated nature of the procedure has created strong disagreements regarding both the total volume of money to be paid to the generators and the allocation of it among the different facilities, especially when there are thermal and hydro units involved. Besides, it is unclear whether the method can effectively enhance the reliability of the system, and dissatisfied consumers argue that they are paying a capacity charge in exchange for nothing. An alternative implementation of the regulated approach is the reliability adder of the pool of England & Wales (this adder will disappear with the reform that is currently taking place), which artificially raises the pool price in the hours when it is detected (in the short term) that the system is close to rationing. The English model has been criticized for distorting the short-term market prices and for being particularly easy to manipulate.

An instance of the third alternative, -capacity markets [7], [8], [9]-, has been adopted in some of the pools in the Northeast of the USA (PJM, NYPP, NEPOOL) where the regulatory authorities determine the amount of firm capacity that each one of the consumption entities has to buy, as well as the maximum amount that each generator is allowed to sell. Organized markets are established to facilitate trading. As a consequence, the price that remunerates installed generation capacity is the outcome of a competitive capacity market. One problem with this scheme is that competition is only used to determine the price, but not the quantity that each generator can sell. When there are only thermal units in the system, the regulator can easily compute the firm capacity of any generator, so the question is not contentious, but when there are hydro units involved (or other energy-limited plants), this firm capacity is very difficult to calculate from a model simulation and its determination becomes a very controversial issue, as in the second approach. Another problem with this approach, as with the second approach too, is that it provides a weak incentive for reliability-oriented operation. This is derived from the fact that consumers see no real product in exchange for the capacity they have to purchase, since generators have very few responsibilities resulting from the deal and they can even decide to export instead of selling into the pool, even if the demand is being rationed, if external prices are attractive [8].

This paper proposes the use of some kind of financial products to define a particular type of reliability contracts that really commit the generators to be available when the system needs them because of scarcity of supply. Similar financial products have been put forth in the past in the electricity sector to deal with short-term demand side management questions (see for instance [10]). Here they will be used to create a reliability market that determines the price of the reliability product through competitive mechanisms. The proposed scheme results in a stabilization of the income of the generators, -specially for peaking units, which are the most conflicting ones from a reliability viewpoint-, and provides a clear incentive for new generation investment. Additionally, the method provides a maximum price insurance for consumers, stimulates a reliability-oriented operation of the generation plants and can be equally applied to thermal and hydro units. The proposed approach provides a satisfactory solution to the long-term guarantee of supply problem, with a limited amount of regulatory intervention. This approach was developed by the authors of this paper while working as consultants for the Association of Colombian Generators (ACOLGEN) in the revision of the current Colombian system of capacity payments.

The paper is organized as follows. After this introduction, section II describes the proposed market mechanism in detail, while section III analyzes its characteristics and its expected results. An illustrative example is presented in section IV.

Some additional regulatory considerations are commented upon section V. Finally, the main conclusions are summarized in section VI.

II THE PROPOSED MARKET MECHANISM

It is proposed to establish an organized market where the regulator requires the Market or the System Operator to buy a prescribed volume of reliability contracts from generators on behalf of the demand. These 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 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.

A.- The Product

The reliability contracts to be auctioned consist of a combination of a financial call option with a high strike price and an explicit penalty for non-delivery.

For the purpose of this paper, a call option [11] is defined as a financial contract that entitles the buyer of the option to receive from the seller, for each MW purchased under the option contract, any positive difference between the price *p* of the underlying asset (in our model, the underlying asset is defined to be the energy sold in the short-term market, so *p* is the spot price for energy) and a predetermined reference price *s*, known as the strike price. This is true for every hour within the time horizon for which the option is defined to be active (e.g. one year). In exchange for that, the seller receives a premium fee *P*.

For a consumer who is buying his demand from the spot market, purchasing one of these call options means that every time the spot price goes above the strike price, he will receive the difference. Thus, when the spot price *p* is higher than s, he will be paying *p* in the spot market and receiving (*p-s*) from his option, and his net payment will be equal to *s*. The call option is acting as a price cap which limits the price at which the consumer is buying from the short-term market at a maximum value of *s*. This can be interpreted as if the consumer had the right, but not the obligation, to buy at the strike price s. When the spot price is low, he decides not to exercise his right and buys directly at the spot market at a price *p*; when the price is higher than *s*, he makes use of his option and buys at a price *s*.

From the generator's point of view, selling an option means that the generator will receive an amount of money *P* in exchange for limiting to *s* the price he will obtain from selling his energy, therefore renouncing to the opportunity of selling at spot prices that are higher than *s*. The generator is exchanging an expected volatile income, associated to the part of the spot price above the strike price *s*, for a fixed payment *P*. The option is then stabilizing a fraction of the generator’s income, therefore reducing his risk. This would be very interesting for risk-averse firms, especially for peaking units whose income volatility is typically very high.

It should be noted that the option-selling generator has to compensate the consumer any time the spot price goes above the strike price, regardless whether the unit is producing or not. If the unit is generating when the price *p* is higher than *s*, it will be receiving the spot price from the pool, so the option just implies having a lower net payment s, but if the unit is not able to produce for some reason, then the option implies the obligation to pay the sum (*p-s*) while receiving nothing from the spot market, and that is a net loss *p* (ignoring production costs) for the generator. It might be argued that the net loss for the generator, when it is unavailable, is always *p*, regardless whether he has signed an option contract or not. But one thing is not being able to pocket some possible revenues *p* and another one having to pay a possible penalty *p-s*. With risk averse generators, having signed an option contract should enhance the stimulus of the units to be producing during those periods when *p>s*, therefore encouraging the generator to incorporate some extra reliability criteria in his operations planning.

In order to strengthen the incentive to be available during the critical periods when *p>s*, it is proposed to add to the option contract an explicit obligation associated to the physical delivery of the committed capacity. Whenever *p>s* and the unit is unable to honor his obligation to produce, the generator will have to pay, apart from the previous implicit penalty (*p-s*), an additional penalty *pen*. This explicit penalty is meant to discourage even more those bids that are not backed by reliable generation capacity.

The example in Fig. 1 and Table 1 illustrates the characteristics of the proposed product.

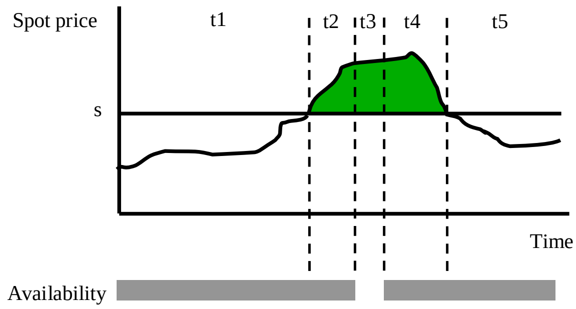

Fig. 1. Application of the reliability contracts

TABLE 1. APPLICATION OF THE RELIABILITY CONTRACTS

Time period		*t1*	*t2*	*t3*	*t4*	*t5*
Demand pays ($/kWh)		*p*	*p*-(*p*-*s*) = *s*	*s*	*s*	*p*
Generator receives ($/kWh)	without option	*p*	*p*	-	*p*	*p*
	with option	*p*	*p*-(*p*-*s*) = *s*	-(*p*-*s*)-*pen*	*s*	*p*

In this case, the spot price exceeds the strike price during three periods: *t2*, *t3* and *t4*, and the generator is able to produce during all periods, except for *t3*. Considering this, the net payments of the different agents are as follows: The demand, which is supposed to be completely hedged by call options, always pays the minimum between the spot price *p* and the strike price *s*. A generator that has not sold any options is not affected by the strike price so he always receives the spot price *p*. An options-seller generator receives the spot price during *t1* and *t5* when he is producing and the spot price is lower than s, he receives the strike price s during $t2$ and $t4$ when he is producing and the spot price is above the strike price, and he has to pay an implicit penalty (p-s) plus an explicit penalty *pen* during $t3$ because, being the spot price p higher than s, he is not able to produce.

B.- The Procedure

The mechanism would be implemented as follows:

1) Initially, the regulator sets the basic parameters (see II-C for a discussion on how to calculate them) for the auction: a) the strike price, which should be at least 25% above the variable cost of the most expensive generator expected to produce during the considered time horizon; b) the total amount of options to be bought, Q, which depends on the expected demand and on the adopted reliability criteria; c) the value of the explicit penalty *pen*; and d) the time horizon of the auction, typically one year.

2) The generators submit any number of bids to the reliability auction, just including a single pair of price and quantity per bid. The quantity in the bid expresses the capacity that the generator is willing to commit through the option, while the price represents the minimum premium fee that the generator is willing to receive for it. Each generator may submit several bids corresponding to several different fractions of its installed capacity.

3) Bids are ordered according to their price and the lowest ones are selected until the sum of all the accepted quantities equals the prescribed quantity Q. The price of the last accepted bid determines the price P in the reliability market, which is the per-unit premium fee that is paid to all the accepted generators.

4) A generator with an accepted bid of q MW receives a premium fee of $P \cdot q$. In exchange, he commits to compensate the demand with a sum $(p\text{-}s) \cdot q$ whenever the spot price p exceeds the strike price s. Additionally, he has to pay to the demand a sum $pen \cdot (q\text{-}g)$ if the spot price is above the strike price and his production g is lower than the committed capacity q. All of this applies for every hour of the auction horizon, and each hour is evaluated independently.

The premium fee P is paid on a day-by-day basis in order to avoid any distortion associated to cash-flow effects. Should the time horizon be very large, some indexation with inflation or some other parameter could be applied.

C.- Determination of the Basic Parameters

1) Strike Price s: Having the strike price predetermined by the regulator instead of allowing each generator to bid his own value for it, reduces the bids to a price-quantity pair that can be easily treated in a simple auction. If generators could submit both their strike price and their desired premium fee, bids would result difficult to compare and this would add significant complexities to the auction mechanisms. Besides, the value of the strike price is not critical since, from the generators' perspective, it only represents the frontier between the fraction of its income that is recovered from the spot market and the fraction that is recovered from the premium fee. If the strike price were slightly lower, then the remuneration that the generators receive from the spot market would be slightly lower and the premium fees required would be slightly higher (see III-A), but the total income would remain constant. So it is not a problem to set the strike price s administratively.

In order to determine a value for this parameter, the strike price s could also be considered as a frontier between the *normal* energy prices ($p<s$) and the *near-rationing* or *emergency* energy prices ($p>s$). Thus, the strike price should be set high enough to ensure that the reliability insurance mechanisms associated to the options only activate when the price is really high and, therefore, it does not interfere with the evolution of the market in normal conditions. It should be at least 25% above the variable cost of the most expensive generator expected to produce.

2) Volume of Capacity to be Auctioned Q: The total required volume of reliability contracts should be prescribed by the regulator, based on the desired reliability target, which can be estimated as a function of the expected peak demand and the available installed capacity, according to the reliability criteria and procedures that are commonly used in traditional power systems analysis, see [12].

3) Explicit Penalty *pen*: The main objective of the explicit penalty *pen* is to reinforce the risk aversion of generators and to discourage those bids with weak generation support. In III-C, after the bidding process is described, some criteria to set this value are presented.

4) Time Horizon: If the option contracts are desired to have a stabilizing effect on the generators' income, then the time horizon of the auction should be long enough to provide a stable economic signal. Short horizons make the premium fees follow too closely the pattern of the spot prices. For example, in a thermal system a monthly auction would tend to yield non-zero prices for capacity (i.e., non-zero premium fees) for only one or two months per year, when the demand is high and rationing is more likely, and would bring zero prices for the rest of the time, when there is spare capacity in the system. However, a yearly auction would provide a constant price for the entire year. A one-year horizon is recommended for a general case, unless specific hydro cycles (for example, El Niño in Colombia) require larger periods.

It is also useful to allow for a lag period so, if the auction is conducted on december 31th of year n, the considered period should not be the year $n+1$, but the year $n+2$, at least. This also helps to stabilize the outcome of the reliability market, by making it more difficult for the agents to incorporate into their bids the effects of the hydro cycles. Besides, it facilitates the participation of new entrants in the reliability market, since they can have their bids accepted before the most expensive part of the construction process starts. A lag period of two years is recommended in a general case.

In this design it is implicitly assumed that bids are valid for every hour in the considered period (one year, typically). Different sub-periods within the year could be identified (for instance, different bids for peak and off-peak hours) but this could lead to a concentration of the premium fee in some periods with the price of the rest being close to zero. A reduced number of reliability products to be traded facilitates trading in general, so the time frame should not be divided into sub-periods. However, under special circumstances this division may help the generators to prepare their bids. This is typically the case when there are two or more periods in the year with similar rationing probability but with potentially different generation availability (for instance, a summer peak with less demand than the winter peak but also with less hydro generation).

III BIDDING PROCESS AND MARKET OUTPUT

A.- Bidding Strategy of the Generators

An economically rational generator should calculate his desired premium fee in the auction (i.e., his bid price) as:

$$P_i = Er\left[\int_{p>s} (1-\lambda_i)\cdot(p-s)\,dt\right] + Er\left[\int_{p>s} \lambda_i\cdot(p-s+pen)\,dt\right] \quad (1)$$

where P_i is the premium fee required by bid i

p is the spot price

s is the strike price

pen is the explicit penalty

λ_i is the probability of generator i of not being able to produce the capacity committed in the option

Er is a mathematical operator that we could call *risk-affected expected value* that evaluates the future income of the generator, taking into account the uncertainty and the risks involved. For a risk neutral firm Er must coincide with the expected value, but it will be different, in general, for each firm.

The first term in (1) represents the income that a generator with a reliability option contract will not receive from the spot market as a consequence of his option, since for him the market price has a maximum value s. The integral extends to every period where the spot price is higher than the strike price s and is equal to the shaded area in Fig. 1. For a risk-averse generator, these are uncertain earnings that he is willing to exchange for a fixed premium fee so, for this firm, the Er operator would value this income below its expected value.

The second term in (1) represents the potential penalties to be paid. It comprises both the implicit penalty associated to the difference between the spot price and the strike price and the additional explicit penalty *pen*. A risk averse firm will use more-severe-than-average hypothesis to evaluate these penalties (for instance, worst-case conditions), so the penalties will be assessed higher than their mean expected value.

From (1) we can observe that, on one hand, the premium fee required by a certain block of capacity increases as its availability decreases and, on the other hand, is independent from the generator's production costs. This means that the more reliable a generator is, the more competitive it will be in this market, and that his competitiveness will not be affected by other criteria different from reliability considerations. These are precisely the characteristics one should look after in this kind of market.

If there were no risk aversion (i.e., if the Er operator were equal to the expected value for all market participants) and if there were no explicit penalty *pen,* then (1) would become

$$P_i = E\left[\int_{p>s} (p-s)\,dt\right] \quad (2)$$

where E stands for the expected value operator, and generators with very different reliability characteristics would bid exactly the same. However, under normal conditions, risk aversion would create a competitive advantage for the more reliable generators and their bids would be naturally preferred in the auction because they would tend to be lower. The explicit penalty *pen* allows the model to reinforce this competitive advantage and to create a difference among reliable and less-reliable generators, even in the absence of risk aversion.

An important advantage of this bidding process is that generators decide on their own how to divide their total capacity into different blocks and how to price each block. Thus, the final volume of capacity assigned to each generator is a market result and not the outcome of an administrative process, as it is the case in most methods currently being applied, where it is the source of endless litigation.

B.- Net income

The bids in the reliability market can be characterized within one of these three prototype cases (market power is ignored in this analysis): a) Firm energy blocks, which would bid the lowest prices, since they will include only the first term in (1), because their failure probability at the time when $p>s$ is negligible; b) Less-firm energy blocks, which would bid higher prices, because of the need to incorporate a significant risk of penalties when the block is not available at the time when $p>s$; c) New entrants (or existing generators considering to exit the system), which should include in their bid, apart from the loss of income in (1), an additional term reflecting their need to recover their total fixed and variable costs, so that the investment is attractive. If spot prices (including any price caps in the spot market) are not enough for this unit to recover its investment costs, then the difference between the expected spot-market remuneration and the required income should be recovered through the reliability market if the group is to be installed. Otherwise, it will decide not to invest.

This procedure serves to ensure that there will be any desired volume of available installed capacity in the system, since the price in the reliability contract market will rise until all the prescribed capacity Q is provided. However, it has to be noted that this capacity price is only guaranteed for the considered period (e.g. one year), so the investment risk is only reduced and the economic signal for investment is not perfect. A sound design of this reliability contract market must promote the reliance of the generators in the stability of the capacity price, so they could make their investment decisions counting on receiving more or less the same income during a number of years. Otherwise, their bids would become prohibitively high.

The typical bid curve in the reliability market would include first the bids corresponding to firm energy blocks, then a first group of less firm bids, then the potential new entrants and finally the group of least firm energy blocks. If there are no significant barriers to entry, this group of least firm blocks (typically the last blocks of the hydro units or very old thermal units) will never be accepted, because there will always be some new entrant to displace them. The four groups of bids are depicted in Fig. 2.

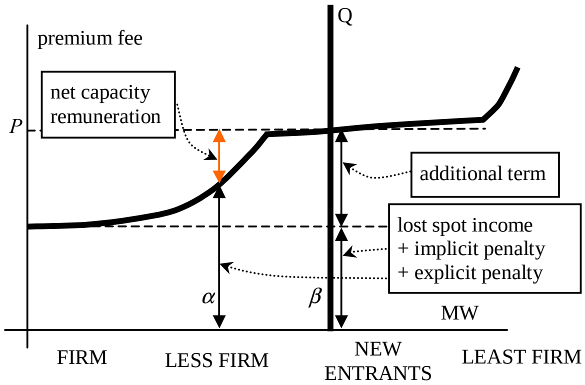

Fig. 2. Typical aggregated supply curve

In Fig. 2 what appears to be a typical aggregated supply curve is shown. Here a new entrant β happens to be the marginal generator and sets the market price for capacity *P*, which is paid to all the accepted bids. The bid price of this generator would include: a) the loss of market income *p-s* because of the option contract, plus the potential penalties it will have to paid (as computed from (1)), and b) an additional term. If this new entrant is a reliable generator, as it should normally be the case, part a) of the bid would be almost as low as the cheapest bids in the market. The additional term would then be equal to its total investment cost minus the income it expects to receive from the market. Any other generator α in Fig. 2 will receive the marginal price *P*, which compensates him for the costs associated to signing the option contract, as expressed in his bid, plus some extra income. This capacity remuneration is an incentive for new capacity to be built and it fulfils the same function as the capacity payments that have been implemented elsewhere. Here these payments are determined via market mechanisms, instead of being determined administratively. This net margin in the capacity remuneration is lower for the less reliable units and higher for the firm energy blocks, regardless of their variable costs or any other characteristic. It seems that the price of the reliability contract market has good properties as an economic signal.

C.- Meaning of the Penalty Function

One could say that typically there will be two kinds of generation competing at the margin in the reliability market: non-firm energy blocks, which would be subject to high penalties, and new entrants, who would be subject to low penalties but will require an additional revenue. If the value of the explicit penalty *pen* is for instance doubled, the bid of the new entrants will change very little, while the bids from the non-firm blocks will dramatically increase. Thus, the level of firmness that is required to be competitive with the new entrants will be significantly augmented. This illustrates how the value of *pen* can be used to determine the equilibrium point between the two types of generators. The regulator could use this to determine a certain level of firmness or availability level which is considered to be the lowest acceptable threshold to be a member of a certain market. Then, *pen* must be set so that any block with an availability below the threshold will be displaced by a new generator. This rule can be expressed in words as "it is preferable to build new facilities rather than accepting such an unreliable generator". These calculations to determine *pen* can only be crude approximations, since they require numerical information of a confidential nature about the costs of potential new entrants.

D.- Implementation

The proposed methodology is based on a market mechanism and, as in any other market, the price is not controlled, it depends on the decisions of the participants, it may end up being lower or higher than expected, and it is subject to the effect of external events.

One of the potential sources of abnormal results is market power, which can only be effectively controlled by structural measures such as the reduction of horizontal concentration, and whose analysis is beyond the scope of this paper. Another potential source of problems may happen during the initial stages of the market when the generators do not fully understand the implications of the auction or they have not developed the proper tools to calculate their bids. There is some space here for irrational bidding behavior and the ensuing market distortion. Since the considered auction period will be typically large (e.g. one year) any errors in the first auctions may have important economical implications. Moreover, the learning process will be necessarily slow. Therefore, much attention must be paid to the design of a set of rules to avoid abnormal behavior during the earlier stages of the market.

The following are the most relevant safeguard rules that were recommended for the application of the proposed method in the Colombian market: a) limit the maximum amount of capacity that a generator can bid into the auction to its nominal capacity; b) do not allow portfolio bidding or combining the bids in any form, so the generators would be more cautious knowing that they cannot rely on other generators from the same firm to provide the committed capacity; c) do not allow the winning bids to transfer their obligation of physical delivery to other generators in secondary markets, in order to reinforce their risk aversion, so that they will bid in the market just the capacity they know almost for sure that will be available when needed.

The need for these safeguard rules is not obvious and it depends greatly on the characteristics of the system. Most of them should be removed after a few years, when the market becomes more mature.

IV CASE EXAMPLE

The following example illustrates the proposed approach.

A.- System Data

The system consists of two thermal generators, T1 and T2, two hydro generators, H1 and H2, and a number of potential new entrants, all of them thermal, that are represented by a single letter N. T1 comprises just one unit, while T2 has three units. Each hydro generator has several capacity blocks, each one with its own water availability level. Additionally, both generators have some small probability of a turbine forced outage. The characteristics of the generators are described in Table 2.

TABLE 2. GENERATORS' CHARACTERISTICS

	Capacity (MW)	Forced Outage Rate (%)		Capacity (MW)	Water Availability (%)
T1	400	7.0	H1	195	100
T2	150	7.5		300	84
	150	7.5		425	68
	150	8.0		500	52
N	1000	7.0		Forced Outage Rate: 0.02 %	
			H2	60	100
				100	70
				185	40
				250	12
				Forced Outage Rate: 0.04 %	

The regulator must establish the auction parameters, whose values are shown in Table 3.

TABLE 3. AUCTION PARAMETERS

Total required capacity Q	1500 MW
Strike price s	150 $/MWh
Explicit penalty pen	350 $/MWh

It is assumed that the market has a price cap of 300 $/MWh.

B.- Bids

Based on these data, each market participant must calculate its bids using (1). Each generating unit will be divided into blocks, and the availability of each block will be obtained from the technical characteristics of the equipment (see Table 2). For simplicity, it is assumed here that whenever the spot price p is higher than the strike price s, it equals the market price cap, so p-s is always equal to 150 $/MWh. Therefore, the critical parameter is the number of hours that each generator estimates that the spot price will be higher than the strike price. In order to simulate somehow the effects of risk aversion on the bids, the estimated duration of the critical period when $p>s$ is assumed to be perceived by each generator as a linear function of its own availability, ranging from 120 hours for the blocks that will never be able to produce when $p>s$ (these generators will tend to overestimate the duration of the critical periods), to 80 hours for the entirely available capacity blocks. Now it is possible to calculate (1) for each block, and the results are presented in Table 4.

TABLE 4. GENERATORS' BIDS

	Quantity (MW)	Availab. (%)	Estimated critical periods (h/year)	Lost spot income ($/MW)	Implicit penalty ($/MW)	Explicit penalty ($/MW)	Extra income ($/MW)	Total price ($/MW)
T1	400	93.00	83	11551	869	2029	0	14449
T2	150	99.96	80	11997	5	13	0	12015
	150	98.24	81	11892	213	498	0	12604
	150	77.00	89	10303	3077	7181	0	20561
H1	195	99.98	80	11999	2	6	0	12007
	105	83.98	86	10885	2076	4845	0	17806
	125	67.98	93	9464	4458	10401	0	24322
	75	51.98	99	7735	7146	16674	0	31555
H2	60	99.96	80	11998	5	11	0	12014
	50	69.96	92	9656	4146	9675	0	23477
	75	39.96	104	6235	9368	21858	0	37460
	65	11.96	115	2067	15215	35503	0	52785
N	1000	93.00	83	11551	869	2029	7500	21949

The additional income that the generators require is zero for all blocks except for the new entrant. For this generator a figure, representing the difference between his total investment and operating costs and his total income from the energy and the capacity markets, has been estimated. Note that the only data that are submitted to the reliability auction are contained in the first and the last columns in Table 4, quantity and price, respectively. Based on these bids, an aggregate supply curve is formed and the market is cleared by crossing it with the specified demand, as shown in Fig. 3.

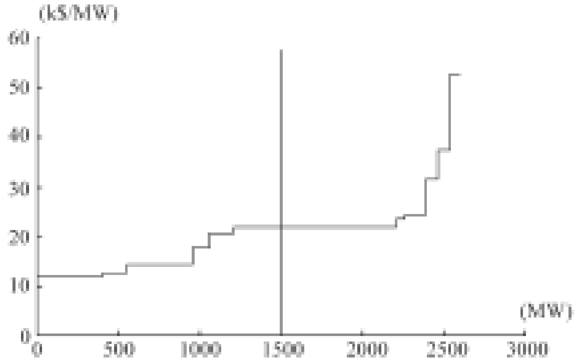

Fig. 3. Market Clearing

C.- Market Output

The results of the market are described in Table 5.

TABLE 5. MARKET RESULTS

	Assigned quantity (MW)	Total income ($/MW)	Net income ($/MW)
T1	400	21949	7500
T2	150	21949	9934
	150	21949	9345
	150	21949	1388
H1	195	21949	9942
	105	21949	4143
	0	21949	0
	0	21949	0
H2	60	21949	9935
	0	21949	0
	0	21949	0
	0	21949	0
N	290	21949	7500

In this example, the new entrant sets the market price. The net income obtained by the generators from this reliability market is equivalent to the capacity payment in the administrative methods that have been implemented in several countries. Note that the allocation of this payment is in agreement with sound efficiency criteria, see [3], since the most reliable blocks receive a larger net income than the less reliable ones.

V REGULATORY CONSIDERATIONS

The proposed approach is based on the answers to three fundamental questions:

A.- Who is Ultimately Responsible for Reliability?

As mentioned before, long-term contracting between consumers and generators must ideally ensure any desired level of reliability for consumers since, in this situation, every consumer would freely decide how much he wants to pay for having a firm energy supply and which is the level of firmness he is willing to pay for. But when markets are not so mature, it seems reasonable that some market authority (the regulator or the system operator) acts on behalf of the consumers purchasing some product that ensures the desired level of long-term guarantee of supply, at least during a transitory period. In this proposal the regulator, maybe assisted by the system operator, determines the characteristics and organizes an auction of reliability contracts.

The proposed scheme can be seen as a market compatible price cap. Often, consumers are not aware of the possibility of having high prices in the system and they expect the regulator to provide some kind of implicit protection against them. At the same time, regulators typically do not want to see that consumers are subject to very high prices, and they are often tempted to impose strict price caps on the market. The problem is that this discourages investment. If the regulator decides that consumers should be protected against high prices, a call option can do so without eliminating the investment economic signal (and probably enhancing it, because of the stabilizing effects of the contract on the generators' revenues). Besides, the reliability options proposed here, as any other long term contract, may have an effect in reducing market power in the short run, for spot prices higher than the strike price.

One may think that the regulator should only represent the small regulated customers, while the larger consumers, who are supposed to manage properly their electricity supply, could make contracts by themselves. Unfortunately, this solution will create free-riding problems that could result in one of the two groups of customers unfairly benefiting from the other's long-term commitments. For instance, this may happen if the regulator specifies a very strong reliability requirement for the captive consumers, while some qualified consumers decide not to spend any money in reliability-related insurance and they just enjoy the guarantee of supply that the rest of the consumers are providing to the system. This situation is not acceptable, but the rules to quantitatively evaluate and organize the reliability support that the qualified consumers may receive are very complex. Until these problems are satisfactorily solved, it is decided that the entire demand should be equally treated.

B.- What is the Appropriate Format of a Reliability Contract? Which is the Product?

One of the fundamental problems with the regulated capacity payment and the capacity market schemes is that the remuneration they provide to the generators depends only on the physical characteristics of the equipment, but it is barely related to the actual performance of the generator during the scarcity situations, i.e., it is a payment for the *potential* contribution to reliability as estimated *ex-ante* by the regulator, but not for the *actual* contribution. Hence, there are almost no incentives for generators to make operational decisions (reservoir management, maintenance scheduling, etc.) that improve their availability during the critical periods and increase the reliability of the system as a whole. From the consumers point of view, there is an amount of money that is being paid to the generators and that does not imply any obligation for them, so the consumers are receiving less than what they could receive in exchange for their capacity payments.

Incentives for reliability-oriented operation can be provided through several kinds of mechanisms. Here, financial options (calls) have been proposed as the product that consumers must buy with their capacity charges. Options are well-known contracts that have been studied in depth in other contexts. Compared with other financial products that can also bring adequate incentives, such as futures, the proposal for implementation of an options market has the advantage of interfering very little with the spot energy market, since options only activate when the system is close to rationing, while futures are always active. This is a very relevant issue, since an administrative decision to commit in long-term futures-like contracts would probably deteriorate the short-term market and even basically replace it. When the decision about the contracting time frame is freely taken by the consumers, then it can be expected that competitive forces would provide an optimal result, but if these long-term commitments are mandatory, then the structure of the long-term contracts should be minimal. The call option that has been proposed is only active in the near-rationing range of prices and it does not affect the spot market under normal conditions.

Accordingly, voluntary long-term contracts (e.g. contracts for differences) would still have a place in the market, coexisting with the mandatory options. From the consumers' point of view, the option would provide a hedge for the part of the price between the strike price and the price cap, while the long-term contract would protect him against price fluctuations between zero and the strike price. Thus, in order to calculate the value of a future contract under the proposed scheme, the part of the spot price that is above the strike price (the shaded area in Fig. 1) should not be considered.

The product here can be seen as an improvement on the capacity market approach. The capacity market has been replaced by another market where the traded product is "energy produced during the critical periods". This is very closely related with firm capacity, but it includes also a short-term demand of the product that makes it not only dependent on investment decisions, but also on operation decisions. Here there is a strong incentive for the generators to operate in a reliability oriented way. At the same time, consumers receive something (a maximum-price hedge) in exchange for all the capacity that they are contracting. And, since there is a clear definition of the product traded in this market, there is no need for the regulator to calculate the firm capacity of each unit. The volume of reliability contracts assigned to each generator results naturally from the market.

C.- Bilateral Contracts or Organized Markets?

In PJM and other pools in the Northeast USA, the system operator determines the total amount of capacity that is needed, and then allocates it among the different consumers, who have to buy firm capacity from the generators in order to fulfill their requirements. Although some organized markets have appeared, the contracting process is essentially bilateral. However, this design has been influenced by an important degree of vertical integration within the electricity companies, at least in the earlier stages of the market.

The alternative model is an organized exchange where people buy and sell standardized products. This has the advantage of reduced transaction costs -since the agents do not have to shop around to find convenient trades-, or lower entry barriers for new generators. On the other hand, standard contracts have been criticized for their inflexibility. Considering the special nature of the buyers in this market, it has been proposed that the reliability contracts should be negotiated with the system operator in a transparent centralized exchange.

VI CONCLUSIONS

A regulatory framework to address the issue of long-term guarantee of supply in deregulated power systems has been presented. The proposed scheme provides investment incentives with a minimum of regulatory intervention. It may be considered as a form of capacity market where the specific reliability product permits a market-based determination of both the price and the allocation of the payments to the generators. Additionally, the financial instruments that have been designed provide the consumers with a market-compatible cap on the price they pay for their energy and they create a strong incentive for generators to operate and maintain their plants so they are available during the most critical periods for the system.

VII ACKNOWLEDGEMENTS

This paper incorporates the results and ideas of several research and consultancy projects that have been developed at IIT during the last decade. The present form of these ideas was the outcome of a research project developed by the authors for the Association of Colombian Generators (ACOLGEN). Juan Carlos Enamorado actively participated in the project and contributed with invaluable assistance. Useful ideas have been obtained from discussions with numerous people: Diana Campos, Francisco Ochoa and Federico Echavarría, among other people from ACOLGEN; Carlos Batlle, Andrés Ramos, Tomás Gómez and Claudia Meseguer, from IIT; Angel Landa, Carlos Sallé, Carlos Solé, Carmen Illán, Alberto de Frutos and, particularly, Miguel Ángel Fernández Ordóñez, all of them, for some time at least, at the Spanish Electricity Regulatory Commission; Larry Ruff; Pedro Marín and his team at the Carlos III University in Madrid; Shmuel Oren, from the University of California at Berkeley; Hugh Outhred, from the University of New South Wales.

VIII REFERENCES

[1] I. J. Perez-Arriaga, "Reliability in the new market structure", *IEEE Power Engineering Review*, December 1999, pp 6-10.

[2] M. C. Caramanis, "Investment decisions and long-term planning under electrical spot pricing", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-101, No. 12, December 1982.

[3] I. J. Perez-Arriaga, C. Meseguer, "Wholesale marginal prices in competitive generation markets", *IEEE Transactions on Power Systems*, Vol. 12, No. 2, May 1997.

[4] S. Oren, P. T. Spiller, "Wild prices out west: what can be done?", *Public Utilities Fortnightly*, November 1, 2000.

[5] S. Borenstein, "The trouble with electricity markets (and some solutions)". University of California Energy Institute. POWER Working Paper PWP-081, January 2001. Available at http://www.ucei.berkeley.edu/ucei/

[6] A. Ford, "Cycles in competitive electricity markets: a simulation study of the western United States", *Energy Policy*, Vol. 27, No. 11, 1999.

[7] R. A. Rosen, F. Sverrison, J. Stutz, "Can electric utility restructuring meet the challenges it has created?", Tellus Institute, November 2000. Available at http://www.tellus.org.

[8] S. Stoft, "PJM's capacity market in a price-spike world", University of California Energy Institute. POWER Working Paper PWP-077, May 2000. Available at http://www.ucei.berkeley.edu/ucei/

[9] G. L. Doorman, "Peaking capacity in restructured power systems", Ph.D. Thesis, Sintef Energy Research, June 2000.

[10] H. P. Chao and R. Wilson "Priority service: pricing, investment, and market organization", *American Economic Review*, Vol. 77, No. 5, December 1987.

[11] S. Stoft, T. Belden, C. Goldman, S. Pickle "Primer on electricity futures and other derivatives", Lawrence Berkeley National Laboratory Working Paper LBNL-41098, January 1998. Available at http://www.lbl.gov/

[12] J. Endrenyi, "Reliability modeling in electric power systems", John Wiley & Sons, New York, 1978.

IX BIOGRAPHIES

Carlos Vázquez obtained his Electrical Engineer degree in 1993 from the Universidad Pontificia Comillas (ICAI), Madrid, Spain. He is currently a researcher in this university at the Instituto de Investigación Tecnológica (IIT), where he is completing his doctoral dissertation "Auction models for electricity markets". His experience is focused on economics of electric power systems, with emphasis on wholesale markets design and transmission pricing and management, and it includes research projects for companies and regulatory agencies in Spain and Latin America.

Michel Rivier obtained his Electrical Engineer degree in 1986 and a Ph.D. degree in Electrical Engineer in 1998, both from Universidad Pontificia Comillas, Madrid, Spain. He is a Research Staff Member at the Instituto de Investigación Tecnológica (IIT) and Professor at the Electrical Engineering Department of this University. His areas of interest include operations, planning, economy and regulation of electric energy systems. He has participated in several consultancy and research projects concerning these subjects for Spanish and foreign firms related with the electricity industry.

Ignacio J. Pérez-Arriaga obtained the Ph.D., and M.S. in Electrical Engineering from the Massachusetts Institute of Technology (USA) and the Electrical Engineer degree from the Universidad Pontificia Comillas, Madrid, Spain. He is full Professor of Electrical Engineering, founder and Director for eleven years of the Instituto de Investigación Tecnológica, and has been Vice-Rector for Research in the Universidad Pontificia Comillas. For five years he has served as Commissioner at the Spanish Electricity Regulatory Commission (CNSE). He has worked and lectured extensively on power system dynamic analysis, monitoring and diagnosis of power system devices and systems, intelligent computer design of industrial systems, planning and operation of electric generation and networks, and regulation of the electric power sector. In this later topic he has been consultant for governments, international institutions, industrial associations and utilities in more than 15 countries.