NÚMERO 332

VÍCTOR G. CARREÓN-RODRÍGUEZ
JUAN ROSELLÓN

Incentives for Supply Adequacy in Electricity Markets: An Application to the Mexican Power Sector

NOVIEMBRE 2005

CIDE

www.cide.edu

Las colecciones de **Documentos de Trabajo** del **CIDE** representan un medio para difundir los avances de la labor de investigación, y para permitir que los autores reciban comentarios antes de su publicación definitiva. Se agradecerá que los comentarios se hagan llegar directamente al (los) autor(es).

• D.R. ® 2005. Centro de Investigación y Docencia Económicas, carretera México-Toluca 3655 (km. 16.5), Lomas de Santa Fe, 01210, México, D.F.
Tel. 5727•9800 exts. 2202, 2203, 2417
Fax: 5727•9885 y 5292•1304.
Correo electrónico: publicaciones@cide.edu
www.cide.edu

Producción a cargo del (los) autor(es), por lo que tanto el contenido así como el estilo y la redacción son su responsabilidad.

Abstract

This paper studies the electricity market design, long run problem of ensuring enough generation capacity to meet future demand (resource adequacy). Reform processes worldwide have shown that it is difficult that the market alone provides incentives to attract enough investment in capacity reserves due to market and institutional failures. We study several measures that have been proposed internationally to cope with this problem including strategic reserves, capacity payments, capacity requirements, and call options. We then construct a model to analyze the structure of incentives for the expansion of electricity supply in the spot market and capacity in the long run electricity reserve market. Through a game-theory model, we analyze price convergence in three markets: the peak spot market, the non-peak spot market, and the long-run capacity reserve market. We finally carry out a simulation for Mexican power generation.

Resumen

Este artículo analiza el problema a largo plazo del diseño de un mercado eléctrico que garantice la generación suficiente para satisfacer la demanda futura (resource adequacy). Los procesos de reforma en todo el mundo han demostrado que es difícil que el mercado, por sí solo, proporcione incentivos para atraer inversión suficiente en capacidad de reservas debido tanto a fallas de mercado como institucionales. Estudiamos varias medidas propuestas internacionalmente para hacer frente a este problema incluyendo reservas estratégicas, pagos por capacidad, requisitos de capacidad, y "call options". Más adelante, construimos un modelo que analiza la estructura de incentivos para la expansión de la oferta de electricidad en el mercado "spot" y la capacidad de reserva de largo plazo del mercado eléctrico. Mediante un modelo de teoría de juegos, analizamos la convergencia de precios en tres mercados: el mercado spot "pico," el mercado spot "no-pico," y el mercado de capacidad de reserva de largo plazo. Finalmente, realizamos una simulación para la generación de electricidad en México.

Introduction

The recent electricity power crises in California, New York, Italy, Norway, Sweden, Brazil, Argentina, Chile and New Zealand have dramatically showed the importance of a reliable electricity supply.[1] As of 2000, generation reserves have declined in most markets since liberalization.[2] Average reserves have also decreased in most OECD markets except for the UK. An extreme case is Australia where there was significant initial overcapacity but reserves drop significantly after the reform. In the cases of UK, Sweden and PJM reserves in 2000 kept similar to those observed at the time of the original reform, but in Norway there was a decrease of 2% from 1991 to 2000, and in California of 7.5% from 1990 to 1998.

The change in reserve margins has occurred in most cases from a starting point of large reserves so that current reserves generally remain above 16%, which seems acceptable for reliability purposes. Likewise, several of the examples of electricity crises have been in systems that heavily depend upon hydropower (like Brazil). However, there is a growing concern on whether liberalized markets will be able to provide adequate incentives for sufficient investment in generation capacity. This is particularly problematic due to some intrinsic characteristics of electricity markets such as: a) a short-term inelastic demand that implies that the (long-term) supply-demand balance cannot be achieved through a market-clearing price; b) a lack of forward electricity markets beyond one or two years; c) the favorable arena for strategic behavior due to the difficulty to get market clearing prices in tight situations, and d) final consumers do not feel the need to engage in long-term contracts because they are usually isolated from spot prices by regulated tariffs (see Bouttes, 2004, and Vázquez *et al.*, 2002). De Vries and Neuhoff (2003) carry out an extensive analysis of the market and institutional failures in the electricity industry that impedes the development of long-term contracts including: lack of generators' counter-parties to sign long-term contracts, producers' imperfect information of the demand function, regulatory uncertainty on whether the regulator will impose price caps in periods of price spikes, investment cycles due to long-lead times for new generation facilities, generators' market power, and so forth.

Our paper studies the electricity market design long run problem of ensuring enough generation capacity to meet present and future demand (*resource adequacy*). Reform processes worldwide have shown that it is

[1] Reliability in electricity markets is usually understood as the sum of adequacy and security standards. Adequacy (security) is generally associated with the long run (short run). Security describes the ability of the system to deal with contingencies, while adequacy refers to the ability of the system to meet the aggregate consumer energy requirements at all times. Security includes the so called ancillary services (voltage support, regulation capacity), spinning reserves, black start capability, etc.). See Singh (2002), and Oren (2003).

[2] Annex presents data on generating reserves for IEA countries

difficult that the market alone provides incentives to attract enough investment in capacity reserves due to market and institutional failures. We first study several measures that have been proposed internationally to cope with this problem including strategic reserves, capacity payments, capacity requirements, and call options. The analytical and practical strengths and weaknesses of each approach are discussed.

We then construct a model to analyze the structure of incentives for the expansion of electricity supply in the spot market and capacity in the long run electricity reserve market. Through a game-theory model, we analyze price convergence in three markets: the peak spot market, the non-peak spot market, and the long-run capacity reserve market.

We finally carry out a simulation for Mexican power generation, which has been subject to a dim reform process and is characterized by large efficiency differences among plants. First, we construct a benchmark, by using the merit order model for dispatch, for the generation cost given the technologies that are actually employed in Mexico. Then we compare that with the real cost for electricity generation.

1.- Resource Adequacy: Several Measures

Several measures have been proposed to ensure an enough amount of generation capacity reserves. Such measures might be analyzed in terms of their degree of centralization or decentralization with regards to the *amount* of capacity and the *price* of capacity (see Knopff, 2002). In this section we carry out an analysis of each one of these measures both studying their theoretical fundaments as well as their international application and assessment.

1.1.- Totally Centralized Resource Adequacy

We start analyzing two extreme approaches to resource adequacy and investment in capacity reserves. One extreme is a fully centralized solution where a vertically integrated utility centrally deals with imbalances and manages congestion and ancillary services using its own generation resources. This is the "wheeling" model that is used in the United States in areas that have not gone into a competitive structure and that have no spot market (Hunt, 2002). The Mexican model is currently another example of centralized supply adequacy where private independent power producers sell energy to the state monopsony CFE under long-term power purchase agreements that are supported by government funds.[3]

Another centralized alternative is the creation of a "moth ball" (or strategic) reserve with government subsidy and centralized decisions

[33] See Carreón-Rodríguez, *et al.,* (2003), and Madrigal and de Rosenzweig (2003).

regarding both amount and price of capacity. The moth ball reserve would imply a strategic reserve of generation capacity,[4] with an operation centrally controlled by the government and that would only be used during emergencies. There is of course, a social cost to this procedure since subsidies would be financed through public funds at large. Supply of capacity reserves would then be categorized as a public service obligation (Knopff, 2002).

1.2.- Totally Desentralized Resource Adequacy

An opposite extreme approach to resource adequacy is a fully decentralized solution where the market determines the amount and price of capacity resource that will grant resource adequacy. Under such a solution, the different energy markets would be separated and a sequential equilibrium would be reached in the spot market, the forward energy market, the market for capacity reserves, and the forward transmission market through the voluntary participation of agents, and a minimal supervision of an ISO (Wilson, 2002).

Different decentralized models have been tried internationally as in Texas, California, Australian Victoria pool, and NETA in the United Kingdom.[5] The aim has been in some cases (NETA) to get the System Operator out of the spot markets, so that traders manage the spot market as well as manage congestion, and separate arrangements are set up for ancillary services. Typically, the primary income for recovery of capacity costs is the difference between the market clearing price and the generators' marginal cost (scarcity payments).

Hunt (2002) argues that the basic problem of a decentralized model is precisely that it ends up creating private markets not only for spot energy, but also markets for congestion energy, markets for imbalance energy, and markets for ancillary services. She states that all these markets deal with the same energy product, and in an efficient market all these products would end up being traded at the same price.[6] In reality, these prices do not converge, and alternatively higher prices, shortages, bureaucracy and new transaction costs are created.

This view is endorsed by Joskow (2003) who shows that wholesale market designs that separate energy and individual ancillary service markets have performed poorly and have made electricity markets subject to unilateral behavior that leads to price increases. California did an actual separation of

[4] In Norway and Sweden there is direct ownership of some peaking plants (Güllen, 2000).

[5] In England and Wales the existing integrated system was substituted with an extreme version of a decentralized model that discourages the use of imbalances and trading in markets remote from the system operator. According to Hunt (2002) this implies a reduction in the transparency of energy markets because imbalance prices do not reflect efficient contract prices.

[6] We will later show that prices in the capacity reserves, peak capacity and non-peak capacity markets converge to the same price in a model that separates these three markets.

five electricity markets (Hunt, 2002). Some theoretical studies try to find the optimality conditions for such an approach (e.g., Wilson, 2002, and Chao and Wilson, 2002). However, elegant in theory,[7] the electricity industry practice has clearly shown the inconvenience of separating the different markets.

Borenstein (2002) also agrees that electricity markets do not fulfill the conditions for full competition to work, so that decentralized sequential and efficient equilibrium of the different electricity markets is impossible. Market power and volatility are really inherent to electricity markets since demand is inelastic and difficult to forecast. Likewise, supply faces binding constraints at peak times, and it is very costly to store. This implies that short-term prices are extremely volatile so that small changes in demand or supply conditions lead to price bursts, and even small-share generators can exercise market power. Borenstein then claims that the best way that regulators can handle market power is through long-term forward contracts between power buyers and sellers together with *real-time* pricing. Forward contracts help to lower the average price paid in both spot and forward markets, while real-time pricing also makes the demand curve flatter.[8]

In the context of an integrated ISO that reaches a centralized equilibrium in all the electricity markets, De Vries and Neuhoff (2003) analyze the "energy-only" market solution. Such a solution relies on the spot market run by the ISO to take care of resource adequacy so that price spikes signal the need of investment in generation capacity. De Vries and Neuhoff argue that there are not enough incentives for generators in an energy-only market to invest in capacity whenever there exists economic uncertainty, or fluctuations in demand. Moreover, they show that when generators and consumers are risk averse, the optimal level of investment from the perspective of generators is below the level consumers wish to finance with long-term contracts. The main reason is that market designs do not have the *institutions* that permit long-term contracts to develop sufficiently, and generators are restricted in the amount of risk that they can transfer to consumers. Likewise, complete reliance on price spikes is not advisable because they are usually not politically acceptable,[9] and they can also be manipulated by the generation companies. Even more, electricity markets that rely on short-term energy

[7] For example, Chao and Wilson (2002) analyze the two-part Californian procurement auction for the market of spinning reserves. One part of the auction was designed for making capacity available, while the other part was for supplying incremental energy. A scoring rule is meant for comparing bids, while a settlement rule for paying accepted bids. The revelation principle applied to this model makes that each suppliers' optimal energy bid reveals their true marginal cost. Additionally, the ISO and the generators are not required to agree on the probability distribution of dispatched energy

[8] Most of the recent electricity reform proposals also promote the use of demand side bidding measures (see for example Commonwealth of Australia, 2002)

[9] Gülen (2002) shows that if the probability of lost load in the PJM market is 1 day in 10 years, price spikes in the range of $12,000-$30,000 per Mwh are needed in an energy-only market. Energy-only markets work however in Australia and New Zealand with maximum prices between $2,500 and $5,000.

revenues might lead to shortfalls in capacity over time that might originate investment cycles where investment lags the demand in the market.

Regulators worldwide are then very concerned that energy prices are not enough to cover generators' capacity costs. Most markets have implemented some type of resource adequacy measure. Texas has recently changed to generation adequacy assurances, and FERC's Standard Market Design (SMD) also recognized the adequate contracted provision of capacity reserves (FERC, 2002).[10] California in 2001 also changed its market approach to capacity supply and prompted a proposal for an available capacity requirement (ACAP) to be imposed on load serving entities (LSEs).

It is therefore not surprising that several methods have been formally studied in the literature on incentives for investment in reserve capacity such as capacity payments, capacity requirements, and capacity options. The literature on resource adequacy analyzes these mechanisms in the context of an integrated ISO. We next study such mechanisms.

1.3.- Capacity of Payment

Capacity payments provide remuneration to generators for making available their generation capacity (whether they get dispatched or not). The price of capacity is set while the market determines the amount of capacity available. That is, prices are centrally determined while capacity decisions are decentralized. Capacity payments are collected from consumers through an uplift charge and determine the cost behavior of the firm but leave the amount of reserves uncertain. Oren (2003) explains that capacity payments are rooted in the theory of peak-load pricing so that energy is priced at marginal cost and a capacity payment is used to recover the fixed capacity cost imposed on peak-period energy users. The optimality condition is such that the shadow price of the capacity constraint is equal to the incremental cost of capacity.

Capacity payments have been used in Argentina,[11] Chile, Colombia, Peru, Spain (together with bilateral capacity contracts), and the United Kingdom.[12] Two different kinds of capacity payments have been applied in the international practice: fixed payments and fluctuating payments. Fixed payments per MW have been implemented in Spain, where the compensation depends on the availability and the technology of the power plant, and in Argentina, where the *Secretaría de Energía* set a \$10 MWH (\$5 for base capacity and \$5 for reliability) payment paid during peak demand blocks (6 am-11 pm during workdays).

[10] However FERC has recently backed of and recognized the State's jurisdiction over resource adequacy measures.
[11] Argentina changed to a capacity market in 2000.
[12] With the adoption of "NETA" in October 2000, the UK abandoned capacity payments based on the loss of load probability (LOLP) method along with the pool system.

Fluctuating payments vary with the need for reserve capacity. Although later rescinded under NETA, they were implemented in the early UK (England and Wales) electricity market. The market merit-order pricing rule is modified during periods of high demand when reserve capacity margins are low. In such circumstance, the market price is defined as the weighted average of two factors: the price of the last accepted offer to generate (*LAO*) and the value of lost load (*VOLL*). The weight is the loss of load probability (*LOLP*). The formula for the market price is then *market price = LAO * (1-LOLP) + VOLL * LOLP*, where: $0 \leq LOLP \leq 1$. The greater (lower) the surplus reserve capacity the smaller (higher) is *LOLP*. Generators would ideally add capacity when the expected sum of all these payments over all hours of the year is greater than the cost of installing new capacity. This formula also implies a price cap for VOLL when the system is short of power.

Joskow and Tirole (2004) analyze the effects of an uplift charge of an ISO to recover the costs of resources. They do so in the context of a general model that studies the effects on the theorems of welfare economics of market failures as those existing in electricity markets. They find that capacity payments grant inefficient results:

- When the uplift charge is applied both to peak and non-peak periods, large ISO purchases discourage the build up of base load capacity and push down the peak price.

- For small purchases, off-peak capacity decreases when the uplift is applied in both peak and off peak periods, and the peak capacity decreases when the uplift is only applied during the peak period.

In a model of imperfect information, Oren and Sioshansi (2003) analyze payments for reserve capacity in a joint day-ahead energy and reserves auction. Reserves are procured through the energy market using energy only bids, and capacity payments are made based on the generator's opportunity cost. Oren applies the revelation principle to show that generators have an incentive to understate their costs so as to capture higher capacity rents.[13]

Such theoretical assessments are confirmed in practice by the case of Argentina that substituted its fixed capacity payment mechanism for a hybrid system of payments and contracts because fixed payments were found to distort the merit order dispatch and negatively affected the long-term financial situation of thermal generators. In the UK, the LOLP system was manipulated by large players at the end of the pre-NETA period.[14] In several other countries, capacity payments have also led to construction of inefficient peaking units, promote the use of one fuel over others, and eliminated the incentive for availability during crisis of deficit supply.

[13] See also Newbery (1995).
[14] See Green (2004).

Likewise, Singh (2002) asserts that, as in any price-cap procedure, setting the optimal level of capacity payments is very difficult, and Knopff (2002) points out that a practical problem of fluctuating capacity payments is that variations in such mechanism happen in the short run, whereas the relevant time for investment in capacity reserves is the long term. Additionally, Gülen considers that the LOLP method is not adequate for largely hydro-based systems (as Brazil) as the LOLP would be very small during wet seasons, which would lead to disproportionate low revenues for thermal generators. Hunt (2002) then claims that any capacity adder should be designed to reflect the value of the plant to the system, which is in turn affected by the technology plant composition in such a system.

Capacity payments might be combined with price caps to protect consumers (International Energy Agency, 2002) because when capacity is paid separately there is no need that price spikes remunerate reserve capacity. Hobbs, B. F., Iñón, J. and S. E. Stoft (2002) show that the result of such combination could be a reduction in price volatility without affecting average prices and reserves. However, price caps can also have a locational influence on generators that would seek high price-cap areas.

1.4.- Capacity Requirements

Capacity requirements are set as an obligation to maintain a certain amount of reserve capacity. Such an amount is centrally determined through an administratively forecast of demand, and is usually imposed by the ISO (or the regulator) to LSEs. Conversely to capacity payments, the price is decentrally determined by the market once the amount of reserve capacity is set. LSEs must buy enough "capacity tickets" to meet the expected peak load of their customers multiplied by (1+X), where X is the expected reserve margin that will cover an estimated level of reliability to cope with random outages. The tickets are sold by generators who are usually allowed to export their reserve capacity to other markets. With a capacity requirement, the regulator is able to control the reserve level but the cost remains uncertain (IEA, 2002).

Capacity requirements are used in the Pennsylvania, New Jersey, Maryland (PJM), New York and New England markets where an obligation is imposed on LSEs to arrange for Installed Capacity (ICAP). In particular, PJM put into practice a bid-based, day-ahead and month-ahead ICAP markets.[15] LSEs are required to buy ICAP in order to be able to serve loads, and they can trade their ICAP with other LSEs. The ICAP requirements can be met by LSEs through self supply, bilateral transactions with suppliers, capability period auctions (several–month strip), monthly auctions, deficiency-spot market auctions, and so forth. Capacity resources can be exported from (or imported

[15] On October 1st, 1998, PJM initiated monthly and multi-monthly capacity markets, while daily capacity markets initiated their operation in 1999.

to) the PJM area. Generators sell a recall right that enables PJM to recall energy exports from capacity resources when required. When capacity is recalled, the supplier is paid the market price for energy. The System Operator determines demand through the choice of obligations of LSEs, which must own or purchase capacity resources greater than or equal to their expected peak-load plus a reserve margin. If an LSE is short of capacity, it pays a penalty that equals the daily amount of deficiency in capacity times the number of days. When the system itself is short of capacity, the deficiency charge is the double of the capacity deficiency rate (equal in 2003 to USD 174.73 per MW-day).[16]

Oren (2003) then proposes to view long-term reserves as price insurance and be treated as a private good but within the framework of a centralized provision of the ISO that imposes mandatory levels of such insurance on LSEs. These mandatory rules would compensate for several obstacles that consumers face when choosing an adequate level of protection, such as technological barriers on metering control, politically barriers to set electricity tariffs efficiently, and so forth.

For a market based on operating reserves backed by high prices Stoft (2002) shows that optimal investment in generation capacity depends on the inverse relationship between capacity requirements and the purchase price limit on the System Operator: the higher the reserve requirement, the lower the optimal price limit.[17]

Creti and Fabra (2004) make a theoretical analysis of the PJM ICAP market. They build a two-stage game theory model. In the first stage, prior to the realization of demand, generators compete in the capacity market and receive their payments for the capacity amounts they commit. In the second stage, once demand is realized, generators compete in the domestic and foreign markets. When there is an excess of demand, the regulator recalls the suppliers' committed capacity resources, which are paid at market prices. Finally, suppliers get their payments for the energy sold. Creti and Fabra analyze this game for the monopoly and the perfect competition cases, and also study the role of the regulator in choosing the capacity requirement as well as in setting a capacity price cap.

Creti and Fabra derive several results from their model on:

1. The opportunity costs of committing capacity resources.

2. The firm's optimal behavior in the capacity market.

[16] The capacity deficiency rate indicates the annual fixed cost of a combustion turbine in PJM plus transmission costs (PJM, 2003).
[17] Stoft (2002) also shows that in a perfectly competitive market a price cap equal to the average value of lost load results in an optimal level of investment in generation capacity. Ford, 1999; Hobbs *et al.*, 2001 also discuss the need for price caps when markets do not clear.

3. The regulator optimal decisions regarding capacity price caps and the optimal reserve requirement.

In their first result, Creti and Fabra show the trade-off that a generator faces between committing more resources to the capacity market against the foregone revenues from exports (in the case of being recalled). The difference between the foreign and domestic prices then determines the opportunity cost of committing capacity resources.[18] The second result shows that two types of equilibria are possible for the firm's optimal behavior given the value of the capacity price cap, and the reserve requirement set by the regulator. When the price cap is too "low", the generator's opportunity costs will not be covered and a capacity deficit would arise (*capacity deficit equilibrium*). When the price cap is "high", enough capacity resources are able to cover the needed capacity requirement (*market clearing equilibrium*).[19] Finally, Creti and Fabra show that the regulator should always set the capacity requirement equal to peak demand so as to fully avoid the risk of shortage, and to set the capacity price cap equal to the firm's opportunity costs of providing full capacity commitment.

Creti and Fabra's results show the fragility of the ICAP system, which crucially depends on the capacity price cap and the capacity requirement. The administrative calculation of the latter variable is a subjective one,[20] while the optimality of the former variable depends on the market structure of financial transmission rights (FTRs) since the opportunity cost of the generator is given by the price difference between the domestic and foreign markets: if the FTR is subject to market power that will be reflected in the ICAP market.

In practice, ICAP mechanisms have failed to provide investment signals when they are most needed. ICAP markets were subject to market manipulation[21] that caused price spikes in 2000 in PJM. The pool was deficient

[18] More specifically, the opportunity cost is also a function of the probability of recall, the amount of resources needed by the system to assure resource adequacy, and the intensity of price competition in the energy market.

[19] Joskow and Tirole (2004) also build a model that shows how a combination of capacity requirements with capacity price caps might potentially restore investment incentives. Even in the presence of market power, a (Ramsey) optimum can be achieved when: (i) LSE capacity requirements can be met both by peak and base load generators, (ii) capacity requirements are determined using the demand from *all* consumers, and the capacity prices reflect the prices paid by *all* retail consumers, and (iii) the market for peaking capacity is contestable. However, this result is not true when there are more than three states of nature (where two state of nature are "off-peak" and "peak"). In such a case strict price-cap regulation might be used to alleviate market power off-peak and allow peakers to recover their investment (Joskow and Tirole, 2004, pp. 45-46).

[20] There have been efforts to improve the calculation of the capacity requirement. For example, in the New York ISO a demand curve is proposed to be constructed as an alternative to an ICAP market that intends to increase resource reliability by valuing additional ICAP above the fixed capacity requirement (Harvard Electricity Policy Group, 2003).

[21] ICAP gives incentives in the short run for manipulating the availability of plants to increase revenue. Anticompetitive behavior is potentially higher when capacity and system constraints are binding. Another practical

some days in June, July and August 2000 since owners of capacity increased their exports for periods when external prices surpassed the PJM market price. In January 2001, there were price spikes of more than $300 MW-day with a deficiency in system capacity. Furthermore, high market concentration in capacity ownership has also been observed.

In New England, Joskow (2003) has showed that the scarcity rents generated are far below from what would be necessary to attract reserve "peaking" capacity to invest (or continue operation) so as to supply the needed operating reserves and energy during scarcity conditions.[22] This means that the combination of an ISO spot market with ICAP markets has not been capable to provide enough incentives to attract generating capacity to maintain adequate reliability levels. Similar results have been obtained for the New York ISO (Patton, 2002).

The ICAP system is flawed in part because it derives from short-term adequacy concerns rather than long-term, and since it depends on a subjective estimation of a "right" capacity level which depends on generation stocks, fuel prices, load shapes, and elasticity of demand for reserves. Also, since ICAP is combined with the possibility of exportation of capacity, the value of the ICAP depends on the price differences across the adjacent markets. Furthermore, ICAPs have not provided incentives to build new generation facilities and, conversely, have contributed to keep old inefficient plants in place (Harvard Electricity Policy Group, 2003).[23]

PJM has then been looking to modify its ICAP system by developing a new methodology for peak load obligation, and by changing the month-ahead and day-ahead markets to a price-taker auction while retaining mandatory participation in the day-ahead market. Likewise, the ISO New England proposed a new locational installed capacity (LICAP) market since the capacity markets in New England were registering at certain times prices of zero while generation in constrained areas needed to be valued more highly (Davis, 2004).[24] The LICAP proposal includes basing prices in demand curves for Maine, Connecticut, metropolitan Boston, and the rest of New England. New prices are to be phased-in through capped increments in a five-year period. These proposals have been widely opposed by LSEs and other consumers since –in their opinion-- it will only produce huge transfers from

problem of ICAP is the interaction among systems with and without capacity requirements, which might lead to inefficient distortions. (IEA, 2002).

[22] The average scarcity rents in New England of $10,000 Mw-year are very low compared to the fixed cost of a new combustion turbine built to provide reserve capacity estimated in between $60,000-$80,000 Mw-year (Joskow, 2003).

[23] Joskow and Tirole (2004) theoretically show that the inefficient dispatch of resources procured by the ISO in order to be used during reserve scarcity conditions will lead in the long run to substitution of base load units by peak units.

[24] Creti and Fabra (2004) deduce from their theoretical model the possibility that capacity markets clear at zero prices if there is no spread between national and foreign prices.

LSEs to generators, without providing long-term incentives to increase new generation (Davis, 2004).

FERC's original SMD (FERC, 2002) also criticized ICAP requirements and proposed instead the use of resource adequacy requirements with targeted curtailments, penalties for undercontracting, and long-term contracting mandatory measures (FERC, 2002). Chandley and Hogan (2002) argue that this is a further flawed policy because there is no objective way to solve the resource-adequacy problem in accordance with SMD without incurring the many difficult issues faced in ICAP design. They think that the best solution would be to allow prices to clear the energy and reserve markets (so that scarcity costs are properly signaled)[25] while allowing financial hedging contracts and demand-side measures. According to Chandley and Hogan, FERC should not mandate the replacement of ICAP mechanism while totally discouraging a market-clearing alternative for reserve capacity markets.

1.5.- Call Options

As seen in the previous section, capacity requirements have the problem of artificially setting a capacity requirement and the value of maintaining such a capacity. Call options are proposed as an alternative system that would represent a more real value of capacity (Vázquez *et al.*, 2001), and that bundles generation adequacy with price insurance. The desired capacity is centrally determined, while price is decentrally determined but consumers are hedged against huge price spikes. Typically, the System Operator would purchase call options from the generators in a competitive bidding process that would cover the desired capacity.[26] The buyer exercises the option if the spot price is greater than the strike price (and receives a premium equal to the difference between the spot price and the strike price).[27] The strike price of options is used as a price-cap in case of emergencies, and high penalties are imposed for failure to deliver when the option is called. This assures that the promised capacity is really made available, especially during peak periods.

The price cap of a call options system works as a protection to consumers, which will assure that prices stay within a socially acceptable range so that the regulatory intervention becomes a form of insurance against price volatility. Compared to the ICAP system, the risk is now changed to the System Operator that now bears the uncertainty of whether the options are

[25] This is of course confronted with the political motivation to keep prices low. However, from a strictly economic point of view, the experience in industries different from the electricity industry is that "the best cure for high prices is high prices", Harvard Electricity Policy Group, 2003, pp.18.

[26] Alternatively, LSEs could be the buyers of options through self-provision from their own controlled resources or bilateral contracts with generators.

[27] The buyers of the call option may choose the strike price that suits their risk aversion: high (low) strike prices have small (high) premiums. Option premiums also work as substitute efficient signals compared to price signals generated by ICAPs (Singh, 2002).

used or not. Risk is removed from generators that now face a more stable revenue horizon compared to an uncertain and volatile income for peak generation. The expected generators income for prices above the strike price equals the price of the call options, and the generators now receive a fixed payment for the option. Prices and corresponding capacity payments are then derived as market based premia from the market players' strategies for risk management.

Oren (2003) claims that the provision of supply adequacy through LSE's hedging obligations captures several important features. If the LSE obligations are adjusted (say) monthly to reflect fluctuations in forecasted peak demand, a secondary market for call options should emerge that would permit the trading of call options among LSEs. However, while secondary markets permit the LSEs to adjust their positions each month, price volatility in such markets increases the LSEs risk. Oren proposes then to treat hedging as another ancillary service, allowing LSEs self provision through bilateral contracts with the ISO acting as a provider of last resort. The danger is of course that this may interfere with incentives in the contract market, and be perceived by LSEs as an alternative to prudent risk management.

Oren (2003) further alerts that in countries where there is not a well-developed infrastructure of financial markets, LSEs or generators may assume more risk than they might reliably handle.[28] In particular, LSEs might not be able to manage risk in a socially optimal way, so that the regulator should need to set a minimum contracting or hedging level on LSEs. Then again, this would lead to non-market arbitrariness.

Vázquez *et al.* (2002) analyze a call-option mechanism for the electricity market in Colombia. The regulator requires the System Operator to buy a prescribed volume of reliability contracts that allow consumers to get a market compatible price cap in exchange for a fixed capacity remuneration for generators. This entitles consumers to enough available generation capacity. Reliability contracts then consist of a combination of a financial call option with a high strike price, and an explicit penalty for generators in case of non-delivery.[29] The regulator carries out a yearly auction of option contracts and sets the strike price (at least 25% above the variable cost of the most expensive generator) and the volume of capacity to be auctioned (in terms of the expected peak demand and the available installed capacity). However, generators decide how to divide their total capacity into different blocks (firm, less-firm, new entrants, and least-firm) and how to price each

[28] Likewise, the capital market might not be able to provide the long term financing for generation investments commensurate to the associated risk. This combined with inexperience with commodity trading in the electricity industry and the perceived regulatory risk, might raise the cost of capital so much that the investment level will be far below than the needed for an efficient resource adequacy level (Oren, 2003).

[29] When the market price p is greater than the strike price s, and the generator is unable to honor its obligation to produce, the generator will have to pay an additional penalty *pen* (apart from the difference $p–s$). The additional penalty is intended to discourage even more bids not backed by reliable capacity.

block, so that capacity assigned to each generator is a market result and not the outcome of an administrative process. The Vázquez *et al*. proposal is very sensitive to market power. Therefore, they propose for implementation in the Colombian electricity market that: a) the maximum amount that a generator can bid is limited to its nominal capacity; b) portfolio bidding is not allowed; and c) the winning bids cannot transfer their obligations of physical delivery to other generators.

2.- The Model

Now, we propose a simple model to try to analyze some of the facts covered in the above discussion. We want to see if the market gives the right signals to the generation plants to install enough capacity to satisfy current demand, in the spot market, and expected demand, in the long-run reserve market. We assume there are no regulations at all. The most important actor is the System Operator, who is in charge of the dispatch to satisfy demand at all times.

2.1.- Merit Order Model

In this Section, we analyze the profit-maximizing behavior of a generation plant in the electricity market. Any generation plant must choose one of three possibilities: (a) supplying for the short-run spot market, (b) supplying for the long-run reserve market, and (c) not supplying at all. Plants are allowed to supply for only one submarket. Firstly, the generator might supply energy in the short-run market (or pool) or capacity in a long-run market for capacity reserves or not supplying at all. Secondly, if the generator decides to supply for the spot market, it might choose to sell energy in non-peak or peak periods.

The conditions that characterize the optimal behavior of the generators under these scenarios should hopefully provide the ISO with key clues to evaluate the impacts of different pricing rules that seek to enhance supply of energy and capacity reserves.

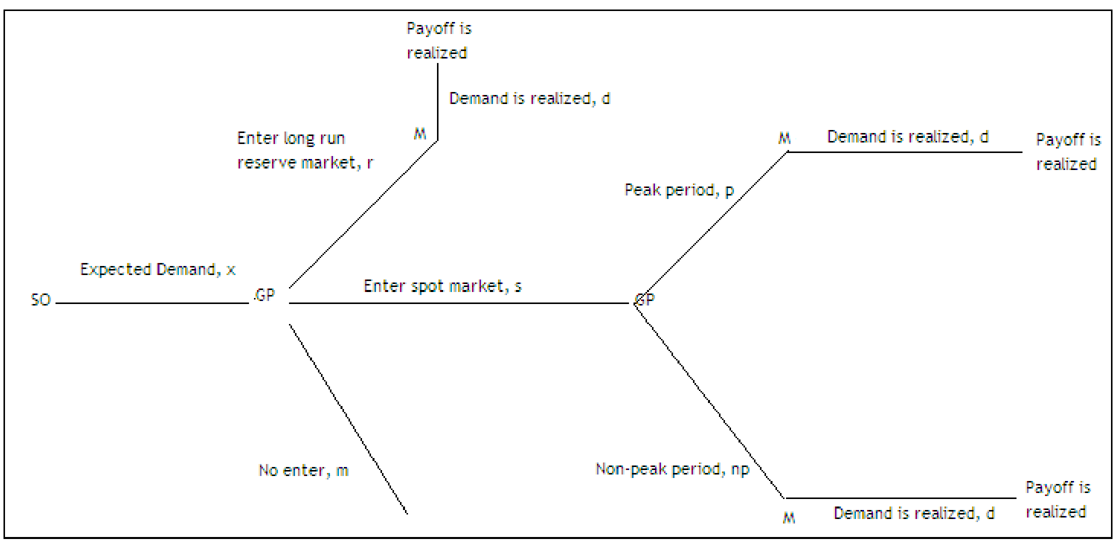
We think in the following mechanism (See Graph 1). There exists a sequence of decisions that any generation plant (GP) must take. First, after the System Operator announces the expected demand for the following day, the GP must decide whether to enter the spot market, to enter the long-run reserve market, or not participate at all in the electricity market. Second, if it decides to participate in the spot market, it must decide to supply for the non-peak or for the peak period. Once any GP has made its decision, the market plays and decides the size of demand in any of the three markets: non-peak and peak periods in the spot market and the long-run reserve market. After this, any GP gets its payoff by computing its expected profits in these three markets. Finally, based on this information, each plant must decide in which market it will participate. Therefore, we will think of this set up as a sequential game.

Definition 1:

We define a sequential game as $\Gamma = \left\{ N, (A_n)_{n=1}^{N}, (u_n)_{n=1}^{N}, P, Z \right\}$ where

(i) N is the set of players
(ii) A_n is the set of actions available for player $n = 1,2,\ldots,N$
(iii) u_n is the payoff function for player $n = 1,2,\ldots,N$
(iv) P is the player function
(v) Z is the set of histories in Γ

Graph 1. Sequence of Decisions



Therefore, in this mechanism we have the following. The set of players is $N = \{SO, GP, M\}$, where SO is the System Operator, GP is the generation plant and M is the market. The SO´s actions are $A_{so} = [0, \infty)$ denoting the expected demand for the following day. The PG´s actions are $A_{gp} = \{n, s, r, np, p\}$. The

Market´s actions are $A_m = [0,\infty)$ denoting the realized demand. The set of histories is $Z = \{\phi, x, xn, xs, xrd, xsnp, xsp, xsnpd, xespd\}$. For example, *xsnpd* means that the SO expects a demand *x*, the PG decides to enter the sport market and supply energy for the non-peak period, and the market chooses a demand *d*. The terminal histories are $T = \{xn, xrd, xsnpd, xspd\}$. The non-terminal histories are $NT = \{\phi, x, xs, xsnp, xsp\}$. The player function is defined as $P: NT \rightarrow N$, where $P(\phi) = SO$, $P(x) = GP$, $P(xs) = GP$, $P(xsnp) = M$, and $P(xsp) = M$. Finally the payoff function for any generation plant is given by its producer surplus, which will be defined below.

Therefore, we will be looking for equilibrium in this game. In this case, the concept of solution will be the Perfect Subgame Nash Equilibrium.

Definition 2:
A Nash Equilibrium is a Subgame Perfect Nash Equilibrium if the combination of strategies is a Nash Equilibrium in each subgame.

Following definitions 1 and 2, we are looking for a configuration of plants in which no plant has incentives to move from one market to another. Based on this model, we will analyze if this mechanism has the right structure to give incentives to expand the generation capacity in the spot market and in the long-run reserve market.

2.2.- Incentives for expasion of capacity

For now, we will analyze the strategic behavior of the generation plants in the short-run spot market. The only choice for them is to choose to generate electricity for the non-peak period or for the peak period, once they decided to enter the spot market. After solving for this model, we will allow plants to decide whether to generate for this market or to offer capacity for the long-run reserve market or staying out of the generation market. In this context, all generators will make their decisions depending on the expected profits they would get in each market

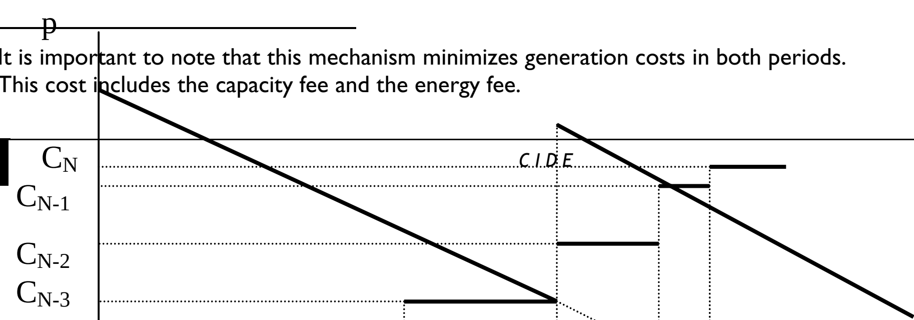
2.3.- The spot market

The spot market works as follows. Each generator decides voluntarily whether or not to participate in the market. Once it decides to participate, it chooses to supply for the non-peak or for the peak period. The System Operator coordinates the market with operations in real-time and forecasting for a day in advance from an engineering technical scope as well as from an economic perspective. Based on the expected demand for the non-peak period, each participating generator makes a merit order bid based on its capacity and costs for the next day. Then, in the real-time market, the System Operator ranks the bids and offers economic dispatch service based on marginal-cost power pricing. That is, generators are dispatched, according to their price bids, from the lowest to the highest one until demand is satisfied. After that, the market price in the non-peak period is the price bid of the last dispatched generator. For the peak period, the System Operator and the participating generators follow the same rules.[30]

Let us consider the following set up. There are N potential generators. Each generator $n = 1,2,\cdots,N$ has capacity of Q_n and cost of $C_n(Q_n)$.[31] Each generator makes a merit order bid based on Q_n y C_n. Suppose that each generator makes a bid of c_n for each unit of capacity that it is willing to supply. That is, generator n offers q_n units of capacity at cost c_n for each unit. Without loss of generality we suppose that $c_1 < c_2 < \cdots < c_N$. So, we have ordered plants according to their bids and name them accordingly. The generation capacities for these plants are $Q_1, Q_2,..., Q_N$ and they offer $q_1, q_2, \cdots, q_N$ to the spot market.

We now make the following assumptions. If the generator decides to participate in the spot market, it offers all capacity in the non-peak or in the peak period; that is $q_n = Q_n$. We do not allow plants to participate in both periods. Let $P = P(q)$ be the inverse demand function, which includes the peak load. We assume that this function is linear in both, the peak and non-peak periods. This inverse demand function has the shape shown in Graph 2. In this Graph we have ranked all generators according to their bids. The quantity supplied in the market is the sum of all the quantities supplied by each one of these plants. That is, the supply curve is the upward sloping curve shown in this same Graph. Then, price and quantity are defined according to this graph following the rules described above. For example, in this case, the price in the non-peak period will be $p = c_{N-3}$ and the quantity supplied will be $q = q_1 + q_2 + \cdots + q_{N-4} + q_{N-3}$.

Graph 2



30 It is important to note that this mechanism minimizes generation costs in both periods.
31 This cost includes the capacity fee and the energy fee.

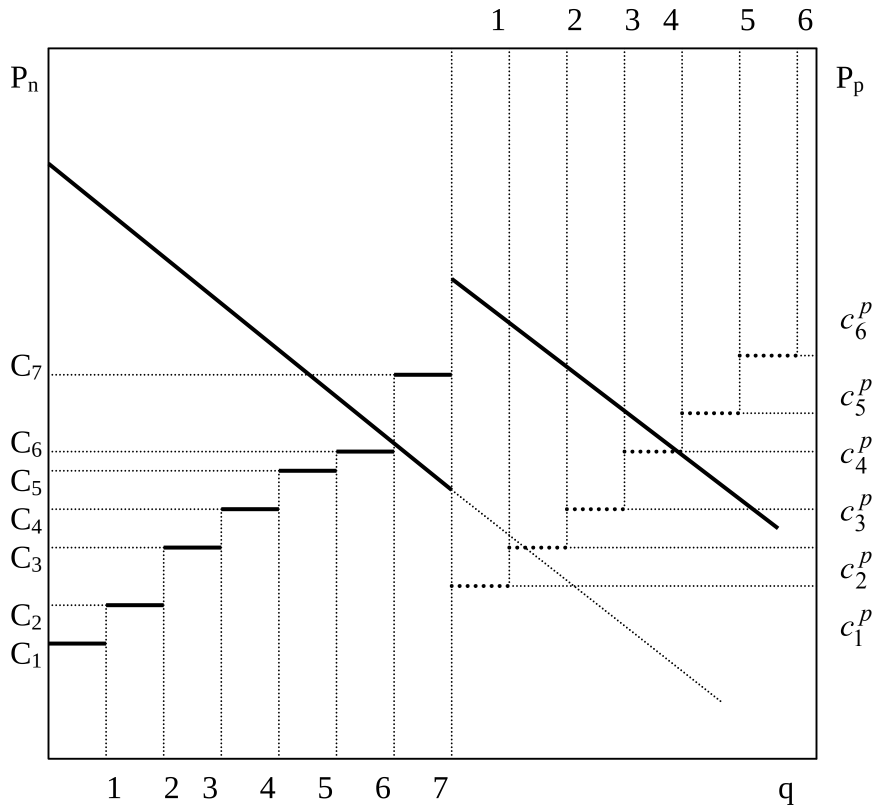
From now on, we simplify this model furthermore. We assume that each plant has only one unit of capacity. This makes computation easier. In this context, we compute the market price for generation, the quantity supplied by the generators, the producer surplus and the consumer surplus. Based on this information, each firm will decide to supply for the non-peak period or for the peak period.

Thus, we have that $q_1 = q_2 = \cdots = q_N = 1$. From the total number of generators, there are N_{np} supplying in the non-peak period, N_p plants supplying for the peak period, and N_r supplying energy for the long run reserve market. This configuration satisfies $N_{np} + N_p + N_r = N$. This situation is depicted in Graph 3. In this Graph we do not show the offers made by the N_r participating in the long run reserve market. Given the demand function and the bids made by these generators, we get the following results.

For the non-peak period: price is $p_{np} = c_6$; quantity is $q_{np} = \sum_{n=1}^{6} q_n = 6$; producer surplus is $PS_n^{np} = p - c_n = c_6 - c_n$ for $n = 1,2,\cdots,6$; and consumer surplus is $CS_{np} = \frac{1}{2}(6)[P(0) - c_6] = 3[P(0) - c_6]$.

For the peak period: price is $p_p = c_4^p$; quantity is $q_p = q_{np} + \sum_{n=1}^{4} q_n = 4$; producer surplus is $PS_n^p = p - c_n = c_4^p - c_n$ for $n = 1,2,\cdots,4$; and consumer surplus: $CS_p = \frac{1}{2}(4)[\overline{P}(7) - c_4^p] = 2[\overline{P}(7) - c_4^p]$

Graph 3



Let us analyze the behavior of plant 1 in this market. Its decision of offering for the non-peak or the peak period depends on the last plants dispatched in each period. In this particular case, plant 1 will get lower producer surplus by offering for the non-peak period since $PS_1^{np} = c_6 - c_1 < c_4^p - c_1 = PS_1^p$ because $c_6 < c_4^p$ and, as a consequence, we have that $p_{np} < p_p$. For the same arguments, all plants offering for the non-peak period have incentives to move to the peak period. However, it could also be that $p_{np} > p_p$. In this case, the incentives would be to move from the peak to the non-peak period. Thus, all plants will decide depending on the cost of the last plant being dispatched in each period. Therefore, the actual prices for the non-peak and the peak period depend on the configuration of plants choosing to serve each period.

Let c_m^{np} and c_m^p be the bids of the last plants dispatched in the non-peak and peak periods, respectively. Then, we have that $\frac{\partial PS_i^k}{\partial c_m^k} \geq 0$ for $k = np, p$. That is, the higher (lower) the bids of these plants, the higher (lower) the producer surplus of all plants in that period. Therefore, the incentives to move from one period to the other will depend on the configuration of each set of generators. No plant will move if $c_m^{np} = c_m^p$.

Finally, an equilibrium in this market is a configuration of plants $\{N_{np}^*, N_p^*\}$ such that $c_m^{np} = c_m^p$. This gives a Nash Equilibrium in this market since no plant has incentives to move from one period to the other.

Now we proceed to analyze the long-run reserve market. In this case, we compute the expected profits of a generator that decides to offer capacity in this market. We then compare these profits with profits it would get in the short run spot market. Based on this, the generator will decide its strategy that maximizes its profits.

2.4.- The long run reserve market

In this section we model the behavior of the generation plants that choose to supply electricity for the reserve market in the long run. This is an uncertain market, since it is unknown the size of demand at that particular point in time. All plants deciding to participate in this market have a probability of being dispatched. The bigger the capacity demanded in this market at that time, the higher the probability of being dispatched. Clearly, given the merit order mechanism, the generator with the lowest bid will be dispatched for sure. For the other plants, it will depend on the size of the actual demand at that moment in time.

In order to analyze this market, we construct a simple model that gives us some hints of what could happen. We assume that demand might be $d = 1,2,3,\cdots,M$ units of electricity. There exists a probability distribution over this demand. The lower the quantity demanded the higher the probability. Let P be a probability distribution over d given by $\{p_m\}_{m=1}^M$. So, p_m is the probability of having a demand of $d = d_m$ for $m = 1,2,3,\cdots,M$ where $p_1 > p_2 > \cdots > p_M$, $p_m > 0$ for $m = 1,2,3,\cdots,M$ and $\sum_{m=1}^{M} p_m = 1$,

Suppose that each plant, n, entering this market makes a bid. It will offer one unit of electricity at cost of c_n. Once all plants willing to supply for the long run reserve market make their offers, they are ranked according to their bids. Say we have N_r plants in the market. Then the ordering will be $c_1^r < c_2^r < \cdots < c_{N_r}^r$. Given this ordering, we compute the expected profits for entering this market.

Plant 1 will get c_1^r with probability p_1, $c_2^r - c_1^r$ with probability p_2, $c_3^r - c_1^r$ with probability p_3, and so on. That is, it will get an expected profit of

$\Pi_1^{re} = \sum_{j=1}^{M} p_m \left(c_m^r - c_1^r \right) = \sum_{j=1}^{M} p_m c_m^r - c_1^r$.[32] Thus, plant n will get an expected profit of $\Pi_m^{re} = \sum_{m=i}^{M} p_m c_m^r - c_m^r$.

Therefore, according to this set up, we have the following results. First, we have $\frac{\partial \Pi_i^e}{\partial c_i^r} < 0$; that is, the lower the costs of plant n the higher the expected profits. Then, the less costly plants are the ones that are more likely to enter this market. Second, if we have a probability distribution $\overline{P}$ given by $\overline{p_m}$ where $\overline{p_1} > \overline{p_2} > \cdots > \overline{p_N}$, $\overline{p_m} > 0$ for all $m = 1,2,3,\cdots,M$ and $\sum_{m=1}^{M} \overline{p_m} = 1$ that is stochastically dominated by the probability distribution P, then the expected profits for all generators will be higher under $\overline{P}$ than under P. Therefore, the higher the expected demand, the higher the expected profits in the long-run reserve market. In this case, more generators will be willing to supply capacity for this market. Third, $\frac{\partial \Pi_i^{re}}{\partial \left(c_j^r - c_i^r \right)} > 0$. That is, the bigger the difference between the n´s bid and the bids of the other firms (which are more costly), the higher the expected profits of generator n. Therefore, the less costly generators with respect to all generators in the market, are the ones that are more likely to enter this market.

Finally, we compare these expected profits with the profits in the short-run spot market. Think of generator 1, the less costly one, for the case depicted in Graph 3 above.

In the non-peak period, it would get $c_6 - c_1$. In the peak period, it would get $c_4^p - c_1$. In the long run reserve market, it would get $\Pi_1^{re} = \sum_{m=1}^{M} p_m c_m^r - c_1$. Given that $c_6 < c_4^p$, it prefers the peak period than the non-peak period. However, if $\sum_{m=1}^{M} p_m c_m^r > c_4^p$, this generator will prefer to supply capacity for the long-run reserve market. That is, if the expected costs of all plants in the long run reserve market is bigger than the price bid of the last dispatched generator in the peak period, then generator 1 would get higher profits in the long-run reserve market. Therefore, we have the following result.

Proposition 1:

[32] We should say at this point that we are assuming that all plants are risk neutral

The Subgame Perfect Nash Equilibrium in this game is a configuration of plants $\{N_n^*, N_s^*, N_r^*\}$ such that $N_n^* + N_s^* + N_r^* = N$ and $N_{np}^* + N_p^* = N_s^*$ where no plant has incentives to move.

Finally, based on these profits, we see that there are incentives for building more capacity for two reasons. First, the new potential generators would use better technologies, which imply lower costs and higher expected profits for them. Second, given that demand is growing over time, the more costly plants will likely be dispatched even though more capacity is installed. The only case when these more costly plants are displaced from the market is when the growth rate of demand is lower than the growth rate of new capacity. In this case, there would be gains in consumer surplus, since the new generation is entering at lower cost and, therefore, there would be lower generation prices.

Moreover, this new capacity would enter the non-peak period, the peak and the long run reserve market depending on the configuration of plants that are generating electricity at that moment in time. These new plants will get producer surplus that is strictly positive. It would be a matter of choice whether they enter the non-peak or the peak or the long run reserve. This decision would depend on the market prices that are expected to prevail in each period. However, it is important to note that generation prices could not decrease over time if the expansion in capacity grows at the same or lower rate than demand.

IV. Simulation

In this section we make a simulation to compute the generation cost in Mexico for 2004. We want to find the minimum generation cost, according to the merit model discussed above, and compare it with the actual cost for the Mexican electricity sector. All data used in this exercise come from CFE. The year used for comparison is 2004 only for the interconnected system.[33] Total installed capacity (in MW) is given for each technology. The load factor is the weighted average for each technology depending on the capacity of each plant. Load factors for 2003 are used for the computation in 2004. We think there is no problem since there is almost no change from year to year. One important caveat applies for hydroelectricity since the load factor depends on the previous raining season and could imply some biased in the computation of the weighted average.

On the other hand, demand is needed to compute the amount of power that is required for the peak and non-peak period. It is classified in three

[33] Baja California is not connected to the rest of the country. For this reason we eliminate both, supply and demand, in the simulation. We also eliminate other small subsystems.

groups. (a) Base demand, which is required 24 hours a day. (b) Intermediate demand, which is required during some hours almost all days a year. Finally, (c) peak demand, which is required for some hours some days a year. For 2004 we have the following estimation. Total demand is 212,480.47 GWh. Base demand is 187,837.06 GWh, intermediate demand is 12,179.11 GWh and peak demand is 12,464.3 GWh. We classify the first two as non-peak and the last one as peak demand for our simulation. Therefore, the non-peak demand is 200,016.17 GWh and the peak demand is 12,464.3 GWh. Finally, since 10% of total demand is satisfied by self supply, cogeneration, etc., we used only 90% of the total demand reported for 2004 for our estimation.

Finally, data about costs, load factors, and capacity for each technology is presented in Table 1. Also, following the merit order model, this table shows the dispatch order for each one.

Table 1: Data for Technology for the Mexican Electricity System

Technology	*Total Capacity*	*Load Factor*	*Cost (Kwh)*	*Dispatch order*	*Total Generation*	*Cumulative Generation*
Carbon	2600	0.73	0.33	1	1417	1417
Dual	2100	0.75	0.33	2	1176	2593
Geothermic	230	0.57	0.43	3	97	2690
Hydro	9900	0.22	0.51	4	163	4321
Combined Cycle	9865	0.70	0.78	5	5169	9490
Steam	14179	0.58	0.93	6	6155	15645
Nuclear	1365	0.88	0.95	7	892	16537
Wind	2	0.40	1.23	8	0.65	16537
Turbogas	2890	0.52	1.32	9	1115	17653

Table 2: Results

Month	*Nonpeak demand*	*Last plant to be dispatched*	*Cost (Kwh)*	*Peak demand*	*Last plant to be dispatched*	*Cost (Kwh)*
January	14110	Steam	0.93	879	Steam	0.93
February	13328	Steam	0.93	831	Steam	0.93
March	15072	Steam	0.93	939	Nuclear	0.95
April	14339	Steam	0.93	894	Steam	0.93
May	14598	Steam	0.93	910	Steam	0.93
June	15638	Nuclear	0.95	974	Turbogas	1.32
July	16327	Nuclear	0.95	1017	Turbogas	1.32
August	16719	Nuclear	0.95	1042	Turbogas	1.32
September	15571	Nuclear	0.95	970	Turbogas	1.32
October	15890	Nuclear	0.95	990	Turbogas	1.32
November	14121	Steam	0.93	880	Steam	0.93
December	14301	Steam	0.93	891	Steam	0.93

We use monthly data for 2004 as our simulation exercise by using a seasonality factor computed from the total generation in that month. Table 2 shows the results of this simulation.

Following the merit order dispatch, the cost of generation in the peak period is higher than the cost for the non-peak period. Once the generation plants observe these prices, they will have incentives to move from the nonpeak to the peak period. In this way the price will equalize after some lower cost plants go to the peak period.

On the one hand there is some monthly seasonality in the consumption of electricity. The highest consumption is during July and August. The lowest consumption months are from January to May. In January, February, April and May, the last plants to be dispatched in both periods are the steam ones. Therefore, the generation cost is the same for peak and nonpeak periods. In March, the last one is the nuclear. For the other months, there are differences in the plants that are dispatched during the peak or nonpeak periods. In these cases the cost is higher for the peak one. We could expect some movement of plants from the nonpeak to the peak period, as stated in Proposition 1.

On the other hand, during July and August the reserve margin is only 1%, while for June, July and September the reserve margin is 5%. In these for months we do not have a reliable system because the reserve margin is below international standards, given by 6% to 9%. So, the Mexican Electricity System has been lucky by not having any disturbance in the system.

Finally, we compare our simulation with data reported for CFE about total generation in 2004. This allows us to get some conclusion about the performance of the system. That is, to know if CFE is following the merit order model for dispatch. Table 3 shows actual total generation and simulated generation by technology.

Table 3: Actual and simulated generation by technology (2004)

Technology	*Simulated generation (Gwh)*	*Actual Generation (Gwh)*
Carbon	16725	17883
Dual	13890	7915
Geothermic	1142	2922
Hydro	19827	25076
Combined cycle	63498	63696
Steam	69170	65783
Nuclear	4697	9194
Wind	3	6
Turbogas	2281	2772
Total	191232	195247

There are some differences between our simulation and the actual generation. The most important ones are the hydro, dual and nuclear technologies. For the hydroelectricity the possible explanation is that the load factor we used in our simulation is smaller than the actual one. This is a possibility if we think that year 2003 was a wet year. In that situation the load factor for these plants can be bigger. However, this could be a special year. It is not guarantee that this will happen all years. The differences for the dual technologies could reflect congestion problems because in general these are located in the most congested zones. Therefore, CFE could decide not to dispatch some of them and replace by more costly ones. Finally nuclear generation is smaller in our simulation because the starting cost of this technology is so high that the best strategy is to dispatch this plant all the time. Moreover, since its capacity is needed during six months, the best strategy is to put in all the time. Finally, there exist a small difference in total generation, which is not so relevant.

Conclusions

In assessing the different alternatives across the literature on supply adequacy, the trend is to look for some kind of transitory regulatory intervention that grants resource adequacy. However, Hunt (2002) claims that capacity obligations or capacity payments can only be useful if hourly metering, hourly pricing, and demand bidding are “woefully inadequate” and cannot be implemented expeditiously. Otherwise, the energy and the reserve markets should not be separated. The ideal would be an ISO that runs day-ahead markets and spot markets that takes care of imbalances and reaches equilibrium of all electricity markets in an integrated way. Market players would meet their long run expectations for the demand-supply balance in well-developed forward and futures markets. Energy and reserve pricing would take care of supply adequacy.

However, in practice electricity markets are usually implemented together with transitory resource-adequacy measures. Capacity payments and requirements alone have been found to be inadequate both in theory and practice. The most advanced developments in the literature point to the use of some type of hedging instruments such as call options. Oren (2003) even argues that capacity payments or requirements might work efficiently if combined with risk management approaches and hedging instruments that promote demand side participation. Regulatory intervention would then be focused on promoting rules that facilitate liquid markets for energy futures and risk management.

Following this discussion in the literature, we proposed a simple model to explain the strategies of the generation plants in the spot market together with the long run reserve market to satisfy the expected demand. We find that the expected cost of generation will be the same in the spot market, for the nonpeak and the peak periods, and in the long-run reserve market. To see the accuracy of this model, we compared our simulated dispatch with the actual dispatch for the Mexican Electricity System for 2004. There are some differences that can arise because of differences in the load factor, congestion costs, or entry costs, which are omitted from our simulation. However, the total generation is very similar. This tells us that this simple model is a good approximation of the dispatch in the Mexican system to satisfy the requirements of capacity to satisfy demand.

Does this mean that the “market” alone is enough to assure resource adequacy in the Mexican electricity market at all times, and that no additional measures should be taken in the future? In answering this question, there are important caveats. Firstly, Mexico does not currently have an open market. As previously discussed, private independent power producers sell energy to the state monopsony under long-term power purchase agreements which are publicly funded. Only in recent years, CFE argues that a mock (or shadow) market has been implemented inside the vertical integrated state monopoly. This virtual market seeks to emulate a competitive market. It uses an optimization model that least-cost dispatch based on actual generation costs (*merit order rule*) in one-day-ahead and real-time markets. The one-day-ahead market establishes production, consumption and price schedules for each of the hours of the following day. The differences between forecasted and actual schedules are cleared at real-time prices. Bids are actually submitted to the System Operator (CENACE) by the different "programmable" thermal CFE's generation plants, which are administratively separated so that they function as different power producers.[34]

Secondly, in this virtual market payments to generators include a "capacity" payment intended to foster the development of generation capacity reserves. It then seems that the combination of this virtual market (with still some elements of central control and subsidy scheme of the state-owned holding company) together with capacity payments has eventually resulted in capacity generation expansion similar to what would be attained in an open electricity market as the one modeled in our study. But this by no means proves that the Mexican electricity industry will not need in the future some of the additional capacity expanding mechanisms discussed in this paper either if it stays under the current monopsonistic schemes, or even if it is the subject of a more aggressive structural reform.

Finally, our model, that simulates a "market" solution to assure resource adequacy, has also some simplifying assumptions that should be relaxed in a more general setting. First, plants have only one unit of capacity. This assumption is avoiding problems about capacity payments and strategic behavior trying to push up price in one of the three submarkets. Second, plants are risk neutral and there is not discount. This assumption assures enough capacity for the long-run reserve market. Third, we ignore the rest of the electricity system, avoiding possible congestion in the transmission lines, among others. Finally, this is a static setting. Plants have no chance to move from one submarket to other over time. Therefore, we can not conclude that any electricity system does not need an additional mechanism to assure resource adequacy, or that the "market" is the right mechanism.

[34] "Non-programmable" generators are small producers that only supply power according to a previously set energy delivery schedule. Hydro generators also make available all their generation capacity, and face production constraints in the one-day-ahead market. Both types of generators then have zero variable costs.

References

Borenstein, S. (2002), "The Trouble With Electricity Markets: Understanding California's Restructuring Disaster," *Journal of Economic Perspectives*, 16: pp.191-211.

Bouttes, J.P. (2004), Roundtable "Market Design and Competition in Electricity," presented at IDEI-CEPR conference "Competition and Coordination in the Electricity Industry", Jan. 16-17.

Carreón-Rodríguez, V.G. and J. Rosellón (2002), "La Reforma del Sector Eléctrico Mexicano: Recomendaciones de Política Pública", *Gestión y Política Pública*, Vol. XI, No. 2.

_______ (2001), "Incentives for the Expansion of Electricity Supply and Capacity Reserves in the Mexican Electricity Sector," Working Paper, CIDE, E219.

Carreón-Rodríguez, V.G., Jiménez San Vicente, Armando, and J. Rosellón (2003), "The Mexican Electricity Sector: Economic, Legal and Political Issues," Working Paper at the Program on Energy and Sustainable Development of the Center for Environmental Science and Policy of Stanford University (available at: http://iis-db.stanford.edu/viewpub.lhtml?pid=20311&cntr=cesp).

Chandley, J. D., and W. W. Hogan (2002), "Initial Comments of John. Chandley and William Hogan on the Standard Market Design NOPR," November 11, available at: http://ksghome.harvard.edu/~.whogan.cbg.Ksg/.

Chao, H.P. and R. Wilson (2002), "Multi-Dimensional Procurement Auctions for Power Reserves: Robust Incentive-Compatible Scoring and Settlement Rules," *Journal of Regulatory Economics*, 22:2, 161-183.

Commonwealth of Australia (2002), "Towards a Truly National and Efficient Energy Market," (available at http://www.ksg.harvard.edu/hepg/Papers/Australian.govt_National.energy.report_02.pdf)

Creti, A. and N. Fabra (2004), "Capacity Markets for Electricity," Working Paper, UCEI, CSEM WP-124.

Davis, T. (2004), "New England Officials, Utilities Yowl Over Installed Capacity Plan," *The Energy Daily*, Wednesday, March 24.

De Vries, L. and K. Neuhoff (2003), "Insufficient Investment in Generating Capacity in Energy-Only Electricity Markets," paper presented at 2nd Workshop on Applied Infrastructure Research (Regulation and Investment in Infrastructure Provision-Theory and Policy), TU Berlin WIP, DIW Berlin, October 11.

FERC (2002), "Notice of Proposed Rulemaking. Remedying Undue Discrimination through Open Access Transmission Service and Standard Market Design", Docket No. RM01-12-000, July 31.(FERC: Federal Energy Regulatory Commission)

Ford, A. (1999), "Cycles in Competitive Electricity Markets: A simulation Study of the Western United States," *Energy Policy* (27): 637-658.

Green, R. (2004), "Did English Generators Play Cournot?", Mimeo, University of Hull Business School.

Gülen, G. (2002) "Capacity Payments" presentation, 22nd. Annual IAEE/USAEE North American Conference, Vancouver 6-8 October.

Harvard Electricity Policy Group, (2003), "Rapporteur's Summaries of HEPG Thirty-First Plenary Session," May 21-22.

Hobbs, B. F., Iñón, J. and S. E. Stoft (2002), "Installed Capacity Requirements and Price Caps: Oil on the Water or Fuel on the Fire," *The Electricity Journal*, July, pp. 23-34.

Hunt, S. (2002), *Making Competition Work in Electricity*, New York, John Wiley & Sons.

IEA (2002), "Security of Supply in Electricity Markets: Evidence and Policy Issues", OECD/IEA. (IEA: International Energy Agency).

Joskow, P. (2003), "The difficult transition to competitive electricity markets in the U.S.," Working Paper, The Cambridge-MIT Institute Electricity Project, CMI 28,

Joskow, P. and J. Tirole (2004), "Reliability and Competitive Electricity Markets," paper presented at IDEI-CEPR conference "Competition and Coordination in the Electricity Industry,", Jan. 16, 17.

Knops, H. (2002), "Electricity Supply: Secure under Competition Law?," proceedings, 22nd Annual IAEE/USAEE North American Conference, Vancouver 6-8 October.

Madrigal, M, and F. de Rosenzweig (2003), "Present and Future Approaches to Ensure Supply Adequacy in the Mexican Electricity Industry," paper presented at the Power Engineering Society Meeting, IEEE, vol. 1, pp. 540-543, 13-17 July, Toronto,.

Murphy, F. H., and Y. Smeers (2002), "Generation Capacity Expansion in Imperfectly Competitive Restructured Electricity Markets," CORE Discussion Paper, 69.

Newbery, D.M., (1995), "Power Markets and Market Power", *The Energy Journal*, Vol. 16, No. 3, pp. 39-66.

Oren, S. (2003), "Ensuring Generation Adequacy in Competitive Electricity Markets," Energy Policy and Economics Working Paper, UCEI, EPE 007.

Oren, S. and R. Sioshansi (2003), "Joint Energy and Reserves Auction with Opportunity Cost Payment for Reserves," paper presented at IDEI-CEPR conference "Competition and Coordination in the Electricity Industry,", Jan. 16-17.

Patton, D. S. (2002) "Review of New York Electricity Markets," *New York Independent System Operator*, October 15, Summer.

PJM, Monitoring Market unit (2003), "State of the Market Annual Report," (http://www.pjm.com) PJM: Pennsylvania, Maryland, and New Jersey ISO.

Singh, H. (2002), "Alternatives for Capacity Payments: Assuring Supply Adequacy in Electricity Markets," presentation, 22nd Annual IAEE/USAEE North American Conference, Vancouver 6-8 October.

Vázquez, C., M. Rivier, and I. Pérez Arriaga (2002), "A Market Approach to Long-Term Security of Supply," *IEEE Transactions on Power Systems*, Vol. 17, pp. 349-357, May 2002.

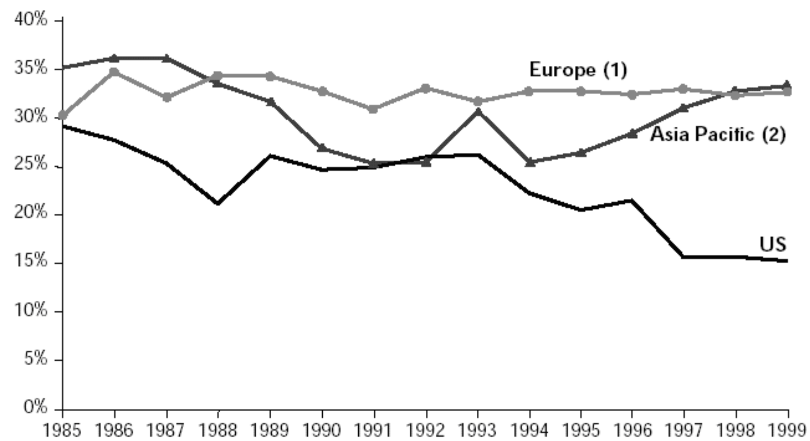
Wilson, R. (2002) "Architecture of Power Markets," *Econometrica*, 70: 1299-*1340*.

Annex

Reserve Margins in IEA Countries:

Figure 1

Reserve Margins in IEA Countries, 1985-1999



(1) Portugal, Italy, Denmark and the Netherlands not included.
(2) Australia not included.
Source: IEA.

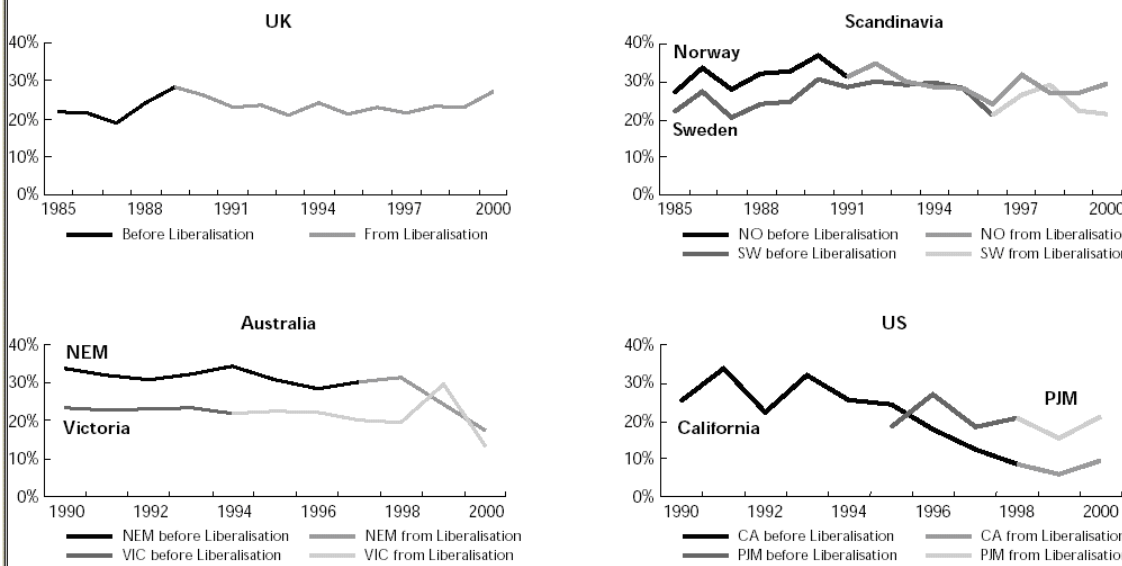
Reserve Margins in IEA Countries (%)

	1985	1990	1995	1999
Australia	36	28	-	21*
Austria	-	61 (2)	60	54 (3)
Belgium	38	26	21	18
Canada	26	19	24	-
Denmark	36	36	46	49
Finland	22	23	22	23
France	31	39	38	37
Germany	27	25	28	29
Greece	42	42	32	31
Hungary	6	9	23	26
Ireland	34	32	24	14
Italy	45 (1)	36	40	42
Japan	35	27	26	33
Luxembourg	54	-	-	-
Netherlands	43	39	41	26 (3)
New Zealand	37	29	34	29
Norway	27	37	28	27
Portugal	-	-	52	57
Spain	46	39	44	39
Sweden	27	36	27	23
Switzerland	47	42	42	33
Turkey	40	46	36	34
United Kingdom	21	26	21	23
United States	30	26	20	16

(1)1986 data. (2)1991 data. (3)1998 data. (-) Missing data.

Source: IEA Database except * taken from ESAA(2001).

Reserve Margins in Selected Power Markets



Notes:

- *In PJM, there is approximately 5 per cent interruptible demand which has been included in the reserve margin calculation, for comparative purposes.*
- *VIC means Victoria, Australia.*

Source: IEA.

Change in Reserve Margins in the Reformed Markets

	UK	Norway	Sweden	Australia Victoria	Australia N.S. Wales	US: California	US: PJM
Change in reserve margin since year of liberalisation until year 2000	0	-2	0	-24	-13	1	0
Change in average reserve margin(1)	5	-3	-5	-16(2)	-7	-7.5	-3(3)
Year of liberalisation	1990	1991	1996	1994	1997	1998	1998

(1) *Difference between average reserves in the five years before liberalisation and average reserves from year of liberalisation to year 2000.*
(2) *Average four years before liberalization in 1994.*
(3) *Average three years before liberalization in 1998.*

Source: IEA.