Testing the Performance of a Forward Capacity Market with Barriers to Entry

by

Timothy D. Mount and Surin Maneevitjit

Department of Applied Economics and Management Cornell University, Ithaca, NY 14853, USA

Abstract

In the northeastern electricity markets, regulators generally favor keeping spot prices close to short-run competitive levels, and as a result, some generating units with low capacity factors that are essential for reliability may not earn enough to be financially viable. This is the "missing money" problem, and regulators in New England, New York and PJM have introduced some form of capacity market to supplement the earnings of all generators. The rationale for adopting this strategy is based on the implicit assumptions that 1) the mix of different types of generating capacity is economically optimal, and 2) the spot prices really are equal to the true marginal operating cost of generation. The question addressed in this paper is whether the economic incentives provided in a capacity market will give the right incentives to get the right type of new capacity built to make the mix of generating capacity more economically efficient.

1. Summary

Maintaining system reliability comes with extra costs because some generating units that are needed to maintain reliability are dispatched at minimum levels most of the time. Given current conditions in the northeastern electricity markets, regulators generally favor keeping wholesale prices close to short-run competitive levels, and as a result, some generating units with low capacity factors that are essential for reliability may not earn enough to be financially viable. This is the "missing money" problem, and regulators in New England, New York and PJM have introduced some form of capacity market to supplement the earnings of all generators. The rationale for adopting this strategy is based on the implicit assumptions that 1) the mix of different types of generating capacity is economically optimal, and 2) the wholesale prices really are equal to the true marginal operating cost of generation. Since neither of these assumptions is realized in practice, there are situations in which some generators (e.g. peaking units) are not earning enough in the wholesale market to cover their capital costs while other units (e.g. baseload units) are earning substantial profits above their capital costs.

The main question addressed in this paper is whether the economic incentives provided in a capacity market will give the right incentives to get the right type of new capacity built to move the mix of generating capacity towards the economically efficient mix. The capacity market run by the New York Independent System Operator (NYISO) has some obvious deficiencies (prices can be manipulated by the incumbent firms and the market clears too close to real time to make it feasible for potential new firms to participate). As a result, the paper focuses the Forward Capacity Market (FCM) proposed by regulators in New England. As the name suggests, the FCM purchases capacity three years ahead of real time, and as a result, potential new entrants can participate and build a new unit only if their offer price is accepted in the auction. In addition, the offer prices allowed for incumbent firms are severely restricted. If an offer for a new unit sets a high price, all capacity is paid this high price but installed capacity can not set a high price.

A series of economic experiments were conducted to test the performance of the FCM using graduate students at Cornell University to represent incumbent firms and software agents to represent potential new entrants. In the first test, there was only one type of generating capacity in the market. The incumbent firms were successful in 1) maintaining market share and keeping out new entrants by undercutting the cost of building a new unit (gains in wholesale market earnings were more than enough to offset the loss in building a new unit), and 2) creating artificial scarcity using legal ways to withhold capacity and therefore allow a new unit to set a high price (by repowering existing units, for example).

In the second test, there were two types of generating capacity, peaking and baseload. The earnings of baseload units in the wholesale market depended on the amount of time that peaking units set the price. Consequently, the earnings of an installed baseload unit will increase if higher loads are met by building new

Corresponding author: Timothy Mount, Tel: (607) 255-4512; Fax (617) 255 9984; E-mail: tdm2@cornell.edu. Timothy Mount is a professor and Surin Maneevitjit is a graduate research assistant in the Department of Applied Economics and Management at Cornell University. This research was supported by the US Department of Energy through the Consortium for Electric Reliability Technology Solutions (CERTS) and by the Power Systems Engineering Research Center (PSERC) through a collaborative project with ISO New England. The authors are responsible for all conclusions presented in the paper, and the views expressed have not been endorsed by the sponsoring agencies.

peaking units. Even though the profits are very high for an installed baseload unit, the results show that the incumbent firms have no incentive to build new baseload units if new entrants can only build new peaking units. The incumbent firms will not reinvest their profits in new capacity unless potential new firms can build new baseload units, and therefore, make low offers to build new baseload units that take into account the earnings in the wholesale market. Hence, institutional barriers to entry associated with the safety of nuclear plants and environmental restrictions on emissions from coal plants may undermine the performance of the FCM given the current high prices of natural gas.

2. Background

The evidence to date about the performance of deregulated electricity markets is not encouraging for the advocates of deregulation. The energy crisis in California in 2000/01 is the most obvious example of a market design that failed and ended up increasing the cost of electricity for many customers in the western states. Although the California crisis was limited to the western states, there is more recent evidence that all deregulated regions in the nation are having trouble getting investors to commit to building new generating capacity when it is needed. A recent report by the North American Electric Reliability Council (NERC) summarizes the current outlook for maintaining reliability in terms of the capacity needed for both generation and transmission in different regions ("2006 Long-Term Reliability Assessment", NERC, October 2006). Four regions have adopted some form of deregulation (ERCOT (Texas), MRO (Midwest), NPCC (Northeast), and RFC (PJM)), three regions are still governed by traditional regulation or public power like the TVA (FRCC (Florida), SERC (Southeast), and SPP (South)), and the Western Inter-Connection (WECC) includes a combination of deregulation (California) and public power in the Northwest (Bonneville).

A comparison of the projections of the margins for generating capacity for the regulated and deregulated regions shows a remarkable difference. In the four deregulated regions, the 2006 projections of the capacity margin fall from the current level of about 15% to below 5% by 2015 in three regions and to less than 10% in the Northeast (NPCC). In contrast, the 2006 projections in the three regulated regions are relatively level at about 15% in two of the regions, and fall to less than 10% by 2015 in one region, the South (SPP). The projection for the West (WECC) is very similar to the situation in the deregulated regions and it falls below 5% by 2015. Comparing the projections made in 2003 and 2006 shows that the recent 2006 projections in the deregulated regions are substantially lower than they were in 2003 in two regions and about the same in the third region (there is no 2003 projection for PJM (RFC)). In contrast, the 2006 projections in the regulated regions are roughly the same in two regions and substantially higher in the third region (Southeast (SERC)). The projections for the WECC are like the deregulated regions, and the projection is much lower in 2006 than it was in 2003.

An important difference between regulated and deregulated markets is that the revenues received by generators in a regulated market are tied to actual costs. In a deregulated market, a large part of the net-revenue earned above the operating costs is fungible and does necessarily go towards the capital costs of generating capacity in a particular region. This problem is exacerbated by the fact that generators receive revenue from more than one market. In New York, for example, generators participate in markets for electricity, ancillary services and capacity. The capacity market was designed by the state regulators specifically for the purpose of encouraging investors to build new generating capacity when it is needed.

There is no general agreement among regulators on whether the earnings in a deregulated wholesale market for electricity should be sufficient to cover both operating and capital costs. Regulators in Australia, Alberta and Texas, for example, support "energy only" markets that cover all production costs and provide the financial incentives needed to get new generating units built when they are needed. In contrast, the deregulated markets in the northeastern and mid-Atlantic states provide generators with supplementary payments above their earnings in the wholesale market. These supplementary payments are designed to correspond to the shortfall anticipated by regulators in the net-revenue needed to cover capital costs.[1] The Independent System Operators (ISO) in New England, New York and PJM advocate using a capacity market to provide this supplementary revenue. However, there is still no general agreement among regulators about the best design for a capacity market.

The challenge for regulators in an energy-only market is to make sure that high prices above the true marginal operating cost occur infrequently and to avoid the type of market "meltdown" experienced during the energy crisis in California in 2000/01. A proposal for the new market design in Texas is to monitor the cumulative net-revenue earned by a proxy peaking unit in the wholesale market during a year. If this net-revenue gets above a specified level, related to the amount needed to cover the annualized capital cost of a Peaking unit, the market rules are changed for the rest of the year and a relatively low price cap is imposed on the market. The most important implication is that the regulators are using

[1] Generally referred to as the "missing money" by regulators and system operators.

Long-Run Marginal Cost (LRMC) pricing to judge the market's performance. In other words, the wholesale market is considered to be competitive by regulators if the annual net-revenue earned in this market is sufficient to cover all of the production costs of a Peaking unit.

Regulators in the northeastern states and PJM have not supported the rationale for LRMC pricing in an energy-only market and have established a more traditional approach using Short-Run Marginal Cost (SRMC) pricing to judge the performance of a wholesale market. Given this criterion, wholesale prices should be equal to the true marginal operating costs at all times unless there is a genuine lack of available generating capacity to meet the system load. However, this is a difficult policy to implement because the structure of electricity markets makes it almost inevitable that some suppliers will speculate by submitting offers to sell that are well above the SRMC. As a result, regulators have implemented additional restrictions on the behavior of suppliers by using, for example, Automatic Mitigation Procedures (AMP) to discourage speculation. To a large extent, they have been quite successful in their efforts to reduce the number of price spikes in these wholesale markets compared to the period immediately after the markets were first deregulated. Figure 1 shows how the behavior of wholesale prices has changed in New York City.

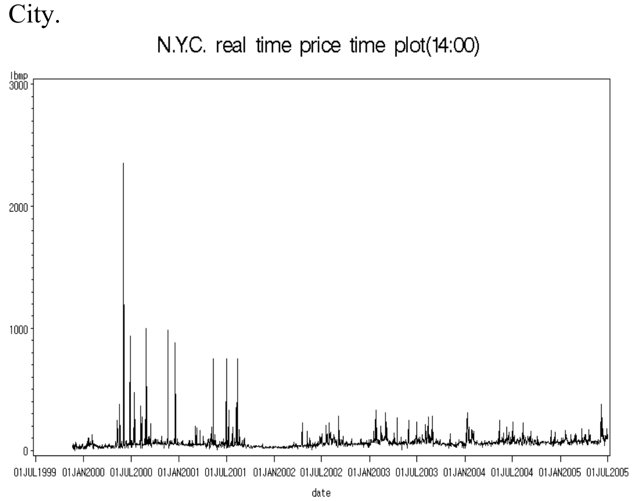

Figure 1: Daily Zonal Wholesale Prices ($/MWh) for NYC in the Balancing Market at 2pm

An important consequence of reducing the number of price spikes in the wholesale market is to introduce the problem of missing money for generators, particularly the owners of Peaking units. This is true even though the evidence to date shows that wholesale prices are still above SRMC. The approach favored in New England, New York and PJM is to modify the structure of an Installed Capacity (ICAP) market to ensure that generators receive enough additional income to cover the missing money. Although most system operators recognize that some supplementary income for generators is needed with SRMC pricing, there is no general agreement about how much income is needed and how this extra income should be provided. The capacity market adopted in 2003 by the New York Independent System Operator (NYISO) provides the most extensive source of evidence to date about how well a specific form of capacity market works. The performance of this market has been disappointing in terms of getting new generating capacity built when it is needed. This poor performance is likely to be a major reason why the new market designs proposed by the Independent System Operators in New England (ISO-NE) and PJM are substantially different from the design of the ICAP market operated by the NYISO.

Minimum amounts of installed generating capacity are determined by the NYISO for three different regions in the Locational Installed CAPacity (LICAP) market.[2] For each region, the regulators specify an explicit "demand curve" for purchases in the capacity market three years in advance. However, this demand curve is only implemented one month ahead of the time when the capacity is needed. The economic rationale for the demand curve is to ensure that the price paid to generators for making their generating units available to meet load is enough to cover the prorated annual capital cost of a Peaking unit. If the amount of capacity purchased is less (more) than the amount required for reliability, the price paid for capacity will be higher (lower). Consequently, when there is not enough generating capacity to meet reliability standards, the price of capacity will be high. Regulators assumed that expectations about future outcomes in the capacity market would provide sufficient incentives for investors to build new generating capacity. In reality, if there really is insufficient capacity offered into the capacity market to meet reliability standards, it is much too late to build new capacity only one month ahead of the time when it is needed.

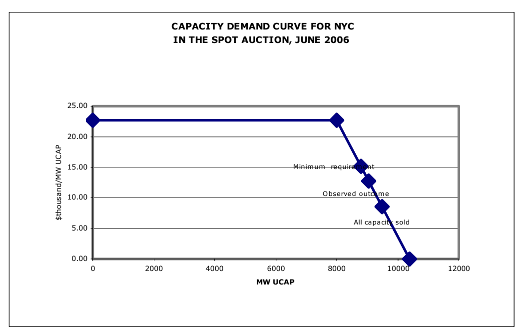

Figure 2: The Capacity Demand Curve for New York City set by Regulators for June 2006

---

2 Demand response is also included if the corresponding reductions of load can be authenticated by the regulators.

An example of the Demand Curve for generating capacity in NYC is shown in Figure 2 for June 2006. It is calibrated to pay the prorated capital cost of a Peaking unit at the "Minimum Requirement" needed to meet the reliability standard (8798 MW UCAP[3]). If all of the installed capacity (9843MW UCAP) had been sold, the market price would have been $8,600/MW (All Capacity Sold in Figure 2). However, the actual market price was $12,712/MW because only 9054MW UCAP were sold (Observed Outcome in Figure 2). Why did this happen? The answer is simple. Some firms get paid more money for selling less. The income from selling all capacity (9843MW UCAP at $8,600/MW) is $82million/month, compared to the actual outcome (9054MW UCAP at $12,712/MW) of $115million/month. The difference of $33million/month is substantial and corresponds to a 40% increase of the total cost of purchasing the capacity.

Even if many of the firms in NYC submit all of their capacity into the auction, it is still perfectly rational for the largest firms to withhold some of their installed capacity from the auction. In fact, the market price and the total cost would be even higher if the regulators had not introduced additional restrictions on how high the market price could be set by the largest firms. These firms are able to manipulate the market price to get exactly the amount that they are allowed. By setting a price cap on the incumbent firms, regulators have set an arbitrary limit on how much market power is allowed in the capacity market. In other words, the regulators consider that a payment of almost $700million over the summer (assuming the price cap is paid to all capacity sold) is an acceptable amount to pay to incumbent firms for being available to generate electricity. If regulators had wanted the generators to offer more capacity into the auction and lower the price of capacity, they could have implemented different rules on behavior by, for example, requiring firms owning more than 1000MW UCAP to submit all of this capacity at a minimal price and be price-takers in the market.

Figure 3 shows the estimated earnings for combined-cycle and combustion turbines in NYC, LI and the upper Hudson valley for 2002-04. These estimates show that total earnings in NYC and LI are well over $100,000/MW/year for combustion turbines and well over $200,000/MW/year for combined- cycle turbines. Furthermore, a large part of the earnings in NYC and LI comes from the capacity auction, and for combustion turbines, the payments for capacity are the dominant source.

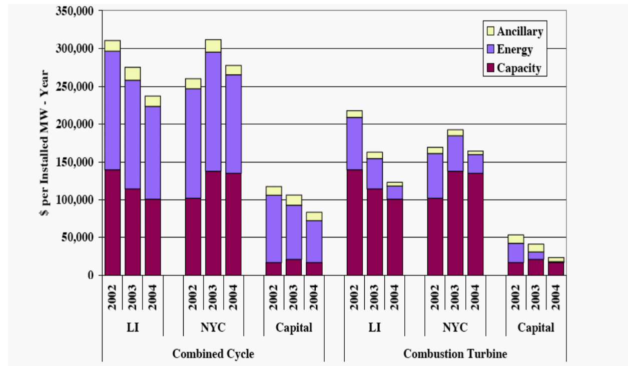

Figure 3: Estimated Earnings (Net-Revenue) of Combined Cycle and Combustion Turbines in Different Locations in the NYCA ("Capital" is the upper Hudson valley).
Source: [13] Figure 16 on p. 23 of the "NYISO 2004 State of the Market Report" <www.nyiso.com>

The important question is why have investors delayed the construction of new generating capacity in NYC given the high level of payments being made in the capacity auction. A plausible explanation is that the earnings from the capacity auction are risky and do not provide the financial security needed to build a new project. In other words, high average payments in a capacity auction are not equivalent to making the same payments through a multi-year Power Purchase Agreement (PPA).

In New York City, over a billion dollars has been paid each year through the capacity market to the owners of existing generating capacity. In spite of this major expenditure, the financial incentives have not been high enough to get investors to commit to building new generating capacity. The main accomplishment of these extra payments has been to increase the market value of the existing capacity. There is no obligation placed on generators in the NYISO capacity market to build new generating capacity when and where it is needed. The basic mistake made by regulators in New York State was to use one policy instrument to treat two very different policy objectives, namely 1) meeting the short-run objective of ensuring there is enough installed capacity available to maintain operating reliability, and 2) meeting the long-run objective of ensuring there is enough new investment to maintain generation adequacy. The financial needs of generators are very different for 1) installed Peaking units, 2) installed Baseload units, and 3) new generating capacity. For the first two types of capacity, financial arrangements already exist for the capital costs of past investments. For new capacity, new

3 The Demand Curve is specified in terms of "Unforced Capacity" (UCAP) to account for different levels of operating reliability for different types of generating unit. The UCAP is equal on average to 94.58% of the "Installed Capacity" (ICAP).

financial arrangements must be established, and typically, potential investors need to have a credible source of future earnings to secure financing.

3. The Forward Capacity Market

Recently, a new form of Forward Capacity Market (FCM) has been proposed for New England.[4] The design of the FCM addresses two of the major problems with the capacity market in New York State. First, the ISO determines how much generating capacity will be purchased three years ahead to maintain generation adequacy. This makes it feasible for investors to participate in the FCM before new capacity is built. Second, restrictions are placed on incumbent firms in the FCM that limit their ability to withhold capacity and to submit high offer prices. By discriminating between new and existing capacity in this way, the type of exploitation of market power by large firms in the capacity market in NYC is likely to be severely limited in the FCM.

The current situation in PJM is similar to the situation in New England, regulators in PJM are not satisfied with the performance of the existing capacity market and have proposed an alternative design that has not yet been implemented. This new design is called the Reliability Pricing Model (RPM)[5]. The RPM is based on a demand curve similar to the one in Figure 2, but, like the design of the FCM, the capacity is purchased four years in advance. However, the responses to a shortfall of capacity are quite different in the RPM and the FCM. When insufficient capacity is offered into the FCM, the ISO cancels the auction and establishes bilateral contracts to build the additional capacity. In contrast, if insufficient capacity is offered into the RPM, the market price of capacity will be higher and less capacity will be purchased than the "right" amount needed to maintain reliability.[6] This higher price may provide the financial incentive needed to build new capacity in the future, but there is still no guarantee that this will actually happen. The strategy followed in the FCM is more direct and is designed to get the right amount of capacity installed in time to maintain reliability standards.

There are other objectives that underlie the design of the RPM. Market prices in the existing capacity market in PJM have been very volatile and have exhibited a boom-and-bust characteristic. This has made earnings from the capacity market very risky for generators. An important objective of the RPM is to stabilize earnings, and this may well occur. However, a key feature of the FCM is missing from the RPM, and this is the ability for investors in new capacity to lock-in a price for up to five years in the FCM. Having a firm contract for five years is a major step towards having the type of financial security needed to raise capital for building new capacity. Holding a standard form of Power Purchase Agreement (PPA) would provide even more financial security.

The basic structure of the FCM is that the ISO determines how much generating capacity will be needed three years ahead to maintain generation adequacy. The ISO purchases this amount of capacity for one year using a descending-clock auction[7] with the same market price paid to all installed and new capacity. This price is a commitment, and generating units accepted in the auction are paid this price three years later if specified standards of performance are met (e.g. being available when needed). The costs of these purchases are covered by the ISO and then allocated to Load Serving Entities (LSEs) in proportion to their actual loads. There are a number of potential advantages of the FCM design compared to the LICAP design. The most important of these are as follows:

1) The ISO determines the amount of capacity purchased and backs this purchase. This responsibility is not given to LSEs, and therefore, the problems of uncertainty about how much capacity will be purchased and the limited credit-worthiness of LSEs are eliminated.

2) The capacity purchased is for availability three-years-ahead, and furthermore, investors in new capacity can lock-in this price for up to five years. Installed generating units can only sell capacity for one year at a time. This rule effectively discriminates between installed capacity and new capacity. Potential investors can establish a secure price for capacity up to eight years ahead. If a similar forward contract for fuel has been secured, an investor would hold contracts that are similar to a PPA for up to five years. However, these contracts do not

4 Affidavit of Peter Cramton, Appendix to ER03-563-000, 030, 055, <http://www.iso-ne.com/markets/othrmkts_data/fcm/filings/index.html>.

5 Statements of Audrey A. Zibelman and Andrew Ott for a Technical Conference on the Reliability Pricing Model, Filed by PJM Interconnection, LLC for FERC Docket Nos. ER05-1410-000 and, EL05-148-000 on February 3, 2006.

6 The minimum reserve margin for generating capacity needed to meet a given reliability standard.

7 This type of auction starts at a high price specified by the ISO. All suppliers owning installed generating units and investors considering the construction of new generating units register the amount of capacity that they would supply at the initial price. Assuming this amount is greater than the amount specified by the ISO, the price (clock) is lowered until some generating units are withdrawn from the auction. The final market price is set when the amount of capacity remaining is equal to the amount needed.

cover the remaining uncertainty about earnings in the spot market for electricity[8].

3) Using a descending-clock auction that starts at a high price implies that the ISO knows in advance whether or not there is sufficient capacity offered into the auction to cover the capacity requirement for generation adequacy. The FCM has explicit rules about starting the auction, and if there is not enough capacity offered to meet this minimum requirement, the auction is cancelled. In this situation, there is a formula for paying all of the capacity that was offered into the auction, and more importantly, the ISO can issue a Request For Proposal (RFP) to build any additional capacity that is needed. The FCM design puts the main emphasis on ensuring that the physical quantity of capacity is sufficient to meet generation adequacy three years ahead. In contrast, the LICAP design only determines the actual amount of capacity purchased one month ahead.

4) Restrictions are placed on incumbent firms in the FCM that limit their ability to withhold capacity and submit high offer prices. The allowed ranges of offers are defined in terms of a specified Cost Of New Entry (CONE), and the offers for new capacity can be up to 2xCONE but the offers for existing capacity must be below 0.8xCONE[9]. By discriminating between new and existing capacity in this way, the type of exploitation of market power by large incumbent firms in the LICAP market in NYC is likely to be severely limited in the FCM. New generating units can submit higher offer prices for capacity, and if this new capacity is needed, the higher price is paid for all capacity. However, the existing generating units cannot set such a high price. There are modifications to this restriction on offers for units that are going to be retired (de-listed), but the general objective of limiting the market power of incumbent firms is an explicit feature of the FCM.

5) Although the market price of capacity is set three years ahead in the FCM, the actual payment to generators occurs in the actual year of delivery. This payment is only made if a generating unit meets explicit standards of availability. In this way, the FCM addresses both the long-run criterion of generating adequacy three years ahead, and the short-run criterion of operating reliability for that delivery year. Payments are reduced by poor performance, and this type of regulatory mechanism has a lot in common to the rationale for using Performance Based Regulation (PBR).

6) A final and unusual feature of the FCM is that generating units accepted in the auction held three years earlier are not allowed to get "excess" earnings from high prices in the spot market for electricity. The combined earnings from the FCM and the spot market are limited in the following way. The ISO sets a price cap of $150/MWh, for example, in the spot market, and all revenues paid to suppliers above this cap are offset by equivalent reductions of the capacity payments. In this respect, the FCM provides a minimum level of earnings, and excess earnings above the price cap are, in effect, returned to the ISO. This mechanism is also similar to the rationale for using a PBR[10]. An important implication of this price cap in the spot market is that it will reduce the incentives for speculating, and therefore, will make the spot prices more competitive. By making the spot prices more competitive, the level of income obtained from the FCM will be more critical for securing the financial viability of new generating units.

The overall conclusion is that the design of the FCM has addressed most of the obvious deficiencies of the LICAP market in New York State. Although there is no evidence available at the time to determine how well the FCM will perform in practice, it seems likely that the FCM will be able to identify possible shortfalls of generating capacity far enough in advance to get new peaking capacity built in time. This is not possible for the LICAP market because clearing only one month ahead effectively limits the range of options for meeting a shortfall to shedding load. Given the importance of maintaining generation adequacy for reliability, it is

[8] This remaining financial uncertainty is likely to be more of an issue for a new baseload unit than a peaking unit because the FCM is designed to cover the full capital cost of a peaking unit but this will be only part of the capital cost of a baseload unit. A baseload unit is expected to earn enough net-revenue in the spot market to cover the rest of its capital cost. In addition, anyone building a new baseload unit would probably want to have a PPA for at least ten years. Hence, the FCM is more likely to accommodate the construction of new peaking units than baseload units, and there is still a substantial amount of financial risk associated with building a new baseload unit without some form of PPA.

[9] Withholding existing capacity from the FCM is also discouraged by requiring generators to present a justification for withholding a generating unit that must be authenticated by the ISO.

[10] A typical example of PBR would provide a floor on the earnings of suppliers in return for some form of profit sharing between suppliers and the public.

unrealistic to rely on the LICAP market. The responsibility for this important task must be taken by some other regulatory mechanism. The LICAP market is effectively a way to provide additional income for incumbent firms. Decisions to build new generating capacity may be influenced by expectations of future earnings in the LICAP market, but the existence of this market does not represent a reliable way to maintain generation adequacy. Most of the money spent in the LICAP market does nothing more than inflate the market value of existing generating units.

4. Results

A series of economic experiments were conducted in spring 2007 to test the performance of the FCM using graduate students at Cornell University. The students represented three incumbent firms and software agents represented potential new entrants. The experiments consisted of two tests, Test 1 and Test 2. In Test 1, there was only one type of generating capacity. Incumbent firms owned some installed capacity and could build new capacity, and new entrants (firms) could also build new capacity. The economic challenge for the incumbent firms (students) was 1) could they maintain market share by keeping new entrants out of the market, and 2) could they build new capacity, and get a high price for all capacity sold, even though new capacity was not really needed to maintain generation adequacy.

In Test 2, there were two types of generating capacity, peaking and baseload. The basic economic conditions corresponded to a situation with high prices for natural gas, and therefore, high costs for peaking units. Under these conditions, the economically efficient choice for building new capacity when it was needed was to build baseload units, because the higher earnings in the spot market were more than enough to cover the higher capital costs. However, the mix of installed generating capacity in the spot market affected the earnings of baseload units, and adding new baseload units lowered the earnings of the installed baseload capacity in the spot market.

Test 2 consisted of two sub tests, Test 2-A and Test 2-B. In Test 2-A, there were barriers to entry and new entrants could only build peaking units. The incumbents could build either new peaking units or new baseload units. In Test 2-B, new entrants and incumbents could build baseload and/or peaking units. The additional economic challenge for the incumbents was to decide what type of capacity to build. In Test 2-A, the incumbents would earn higher total profits by ensuring that only new peaking units were built even though the profits for a new unit would be higher for a baseload unit. In Test 2-B, although the basic economic logic was the same as Test 2-A for the incumbents, new entrants were more aggressive and were willing to build new baseload units at a low price if the combined profits for baseload units from the spot and capacity markets were high.

In both Test 1 and Test 2, the capacity market was run for sessions consisting of ten trading periods. The demand (load) was constant for the first three periods and then grew at a rate of 10 units per period from period four to period eight. Demand was constant for the last two periods. The initial amount of installed capacity was sufficient to cover the demand for the first four periods, and therefore, the amount of new capacity needed to maintain generation adequacy over the ten periods was only 40 units.

In Test 1 when there was only one type of generating capacity, the incumbent firms were successful in maintaining market share and keeping out new entrants. This was accomplished by submitting offers in the capacity market that were lower than the true cost of building a new unit and offsetting this loss with the corresponding higher earnings in the spot market from installed capacity. In addition, the incumbent firms created artificial scarcity during flat demand periods by using legal ways to withhold capacity (i.e. exporting and repowering existing units), and therefore, made it possible for a new unit to set a high price.

In Test 2-A when there were two types of generating capacity but new entrants could not build baseload units, the incumbent firms had no incentive to build new baseload units even though they were more profitable than a new peaking unit. Instead, the incumbent firms used their profits to project earnings from their installed baseload units by building new peaking units. In Test 2-B when new entrants could build both baseload and peaking units, it was more difficult for the incumbent firms to protect their high profits from installed baseload units in the spot market, and as a result, they built some new baseload units as well as new peaking units to prevent new entrants making very low offers to build new baseload units and bringing down the price of all capacity.

Figure 4 illustrates the results for Tests 1, 2-A and 2-B for selected groups of students who were able to manipulate the market successfully. In all three cases, the solid black line represents the economically efficient cumulative additions to generating capacity needed to maintain generation adequacy, and in all three cases, the actual additions were higher than the efficient amounts. In particular, the students were able to withhold existing capacity using legal means to create artificial scarcity in the early periods so that new capacity had to be purchased. By doing this, the price paid for all capacity was higher than it would be if installed capacity had set the price.

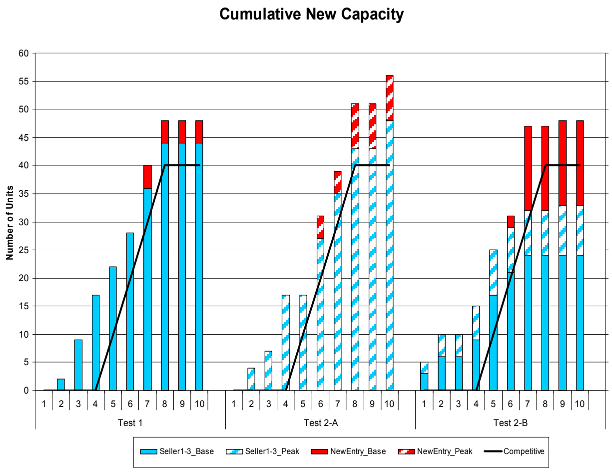

Figure 4: The Cumulative New Capacity plots for FCM Test 1, Test 2-A and Test 2-B (selected groups)

Figure 4 also shows how many new units were built by incumbents (blue) and by new entrants (red). For Test 1, incumbents built more than the 40 units needed in an efficient market and new entrants built less than 5 units. For Test 2, the results were similar. However, the important additional result is that only new peaking units (striped) were built even though a new baseload unit (solid) would be more profitable. These students realized that building new baseload units would reduce earnings from their installed baselaod units in the spot market. In Test 2-B, both incumbents and new entrants built new baseload units. The incumbents also built some peaking units, but it was much harder for them to protect the profitability of their installed baseload units in the way that they had in Test 2-A.

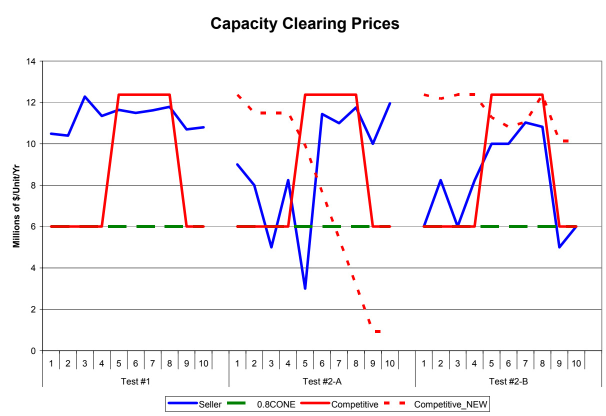

Figure 5: An Example of the Prices for Capacity in Test 1, Test 2-A and Test 2-B (selected Group)

Figure 5 shows the capacity clearing prices for the three cases. The green dotted line is the maximum allowed offer for installed capacity. The competitive price (red) is equal to this maximum offer when no new capacity is needed, and is equal to the actual cost of new entry when new capacity is needed. In Test 1, the incumbents firms were able to maintain the price well above the maximum offer for installed capacity (by legal withholding so that a new unit was able to set the price), and in addition, they undercut the cost of a new entrant when new units were needed to meet growth in demand.

In Tests 2-A and 2-B, the solid red line is the competitive price assuming that only new peaking capacity can be built. The red dotted line is the entry price for a new baseload that includes earnings in the spot market. In Test 2-A when new entrants could not build baseload units, the incumbents built less profitable peaking units to increase their earnings from their installed baseload units. As a result, the true entry price for a new baseload unit fell almost to zero. Under these circumstances with barriers on the construction of baseload capacity by new entrants, the market for the incumbents is not incentive compatible. It does not pay the incumbents to use their high profits from installed baseload capacity to build new baseload capacity even though a new baseload unit would be profitable on its own. As a result, the incumbents built less profitable peaking capacity.

The results were different in Test 2-B when new entrants could build new baseload units. The incumbents had to accept modest profits from their installed baseload units to avoid having new entrants bring the market price down by making very low offers to build new baseload units. Consequently, the entry price for a baseload unit (red dotted line) did not drop the same way that it did in Test 2-A.

	Test1#1		Test1#2		Test1#3		Test#4		Test1#5		Test1(Av)	
	Se	Ne	Se	Ne	Se	Ne	Se	Ne	Se	Ne	Se	Ne
G1	35	10	24	24	37	10	27	14	32	10	31	14
G2	28	20	35	10	47	0	44	4	42	2	39	7
G3	30	16	44	4	41	6	48	0	51	4	43	6
G4	12	38	26	20	34	14	34	6	41	2	29	16
G5	50	6	48	0	41	6	32	18	44	4	43	7
AV	31	18	35	12	40	7	37	8	42	4	37	10

Table 1: New Units Built by Incumbents (Se) and New Entrants (Ne) in Test 1 (Five 10 period Sessions conducted by five different Groups of students, G1-5)

Table 1 summarizes the results of Test 1 for all five groups of students and all five sessions in terms of the number of new units built by incumbents (Se) and by new entrants (Ne). In 24 of the 25 sessions, the incumbents built more units that the new entrants, and in 3 sessions, new entrants did not build any new units.

	Test2-A_#1			Test2-A_#2			Test2-A (AV)		
	Seller1-3		NeEn	Seller1-3		NeEn	Seller1-3		NeEn
	Ba	Pe	Pe	Ba	Pe	Pe	Ba	Pe	Pe
G1	32	18	0	29	19	0	31	19	0
G2	36	19	0	16	27	0	26	23	0
G3	20	9	22	34	8	10	27	9	16
G4	0	48	8	0	40	14	0	44	0
AV	22	24	8	20	24	6	21	24	4

Table 2: New Units Built by Incumbents (Sel1-3) and New Entrants (NeEn) in Test 2-A (Baseload, Ba and Peaking, Pe. Two 10 period Sessions conducted by four different Groups of students, G1-4)

	Test2-B_#3				Test2-B_#4				Test2-B (AV)			
	Seller1-3		NeEn		Seller1-3		NeEn		Seller1-3		NeEn	
	Ba	Pe	Ba	Pe	Ba	Pe	Ba	Pe	Ba	Pe	Ba	Pe
G1	36	12	2	0	32	19	2	0	34	16	2	0
G2	30	16	4	0	5	26	18	0	18	21	11	0
G3	16	18	18	0	24	9	15	0	20	14	17	0
G4	22	34	19	0	29	23	6	0	26	29	13	0
AV	26	20	11	0	23	19	10	0	24	20	11	0

Table 3: New Units Built by Incumbents (Sel1-3) and New Entrants (NeEn) in Test 2-B (Baseload, Ba and Peaking, Pe. Two 10 period Sessions conducted by four different Groups of students, G1-4)

Tables 2 and 3 summarize the results for Tests 2-A and 2-B. In Test 2-A, only group G4 truly understood that the most profitable strategy was to add new peaking capacity to increase earnings from their installed baseload units. However, this group lost some money by undercutting the cost of building new peaking units. They would have earned even higher profits by allowing new entrants to build the new peaking capacity.

In Test 2-B, the incumbent firms had to deal with two important issues at the same time, namely, 1) maintaining market share by not allowing new entrants to build profitable baseload capacity, and 2) trying to keep the earnings of their installed baseload units from the spot market as high as possible. All four groups realized that they should build some new baseload units as well as peaking units. However, the incumbents found it much harder to keep new entrants out of the market and new entrants were able to build some new baseload units in all 8 sessions. On average, the total number of new units built by all participants was higher in Test 2-B (55) than in Test 2-A (49) and Test 1 (47). In all three cases, only 40 units would be needed in a competitive market.

5. Conclusions

The overall conclusion about the performance of deregulated markets for electricity in New York, New England and PJM is that regulators have adopted procedures that make the financial incentives in the wholesale market insufficient to get investors to build new Peaking units when they are needed. In response to this problem, additional sources of income for all generators have been established and the primary source of income is to use some form of capacity market. Most of the evidence about the performance of a capacity market comes from New York State because this market has been operating since 2003. The current performance of this market has been disappointing. First, it has still not overcome the problem of delays in the construction of new generating units in NYC. This is true in spite of making payments of over $1billion/year to incumbent firms in NYC. Second, the largest firms have been able to increase the market price of capacity and increase their earnings by exploiting market power in this capacity market.

Partly in response to the ongoing problems with the performance of the NYISO market, regulators have proposed new market designs for New England and PJM that have not yet been implemented. Even though these two designs are quite different from each other, they share a common feature of purchasing capacity three and four years ahead, instead of just one month ahead as is done in the existing capacity market in New York. This change in design makes it feasible for potential investors in a new generating unit to participate in the capacity market before the unit is built, and to determine the capacity price for this unit when they commit to bringing it on-line. Hence, it is likely that these new capacity markets will be more effective than the NYISO market in getting new generating units committed in time to maintain generation adequacy. It remains to be seen how expensive these new markets will be for customers. The current situation in New York is that the capacity market is very expensive and there is still a lot of uncertainty about whether new capacity will be built when it is needed.

A series of economic experiments were conducted to test the performance of a Forward Capacity Market (FCM) using graduate students at Cornell University to represent incumbent firms and software agents to represent potential new entrants. In the first test, there was only one type of generating capacity in the market. The incumbent firms were successful in 1) maintaining market share and keeping out new entrants by undercutting the cost of building a new unit (gains in wholesale market earnings were more than enough to offset the loss in building a new unit), and 2) creating artificial scarcity using legal ways to withhold capacity and therefore allow a new unit to set a high price (by repowering existing units, for example).

In the second test, there were two types of generating capacity, peaking and baseload. The earnings of baseload units in the wholesale market depended on the amount of time that peaking units set the price. Consequently, the earnings of an installed baseload unit increase when higher loads are met by building new peaking units. Even though the profits are very high for an installed baseload unit, the results show that the incumbent firms have no incentive to build new baseload units if new entrants can only build new peaking units. The incumbent firms will not reinvest their profits in new capacity unless potential new firms can build new baseload units, and therefore, make low offers to build new baseload units that take into account the high earnings in the spot market. Hence, institutional barriers to entry associated with the safety of nuclear plants and environmental restrictions on emissions from coal plants, for example, may undermine the performance of the FCM. Under these circumstances, there is no economic incentive for owners to use profits from installed baseload units to expand baseload capacity even though it would be the socially efficient to do so.

6. References

[1] J. Chen, T. D. Mount, J. S. Thorp and R. J. Thomas, *Location-based scheduling and pricing for energy and reserves: a responsive reserve market proposal*, Decision Support Systems, 40 (October 2005), pp. 563-577.

[2] J. Chen, J. S. Thorp and R. J. Thomas, *Time–space methods for determining locational reserves: a framework for location-based pricing and scheduling for reserve markets*, CERTS Report on Reliability Adequacy Tools (2002 (Dec.)).

[3] J. Chen, J. S. Thorp, R. J. Thomas and T. D. Mount, *Locational Pricing and Scheduling for an Intergrated Energy-Reserve Market*, *Proceedings of the 36th Annual Hawaii international Conference on System Sciences (Hicss'03) - Track 2 - Volume 2 (January 06 - 09, 2003)*, IEEE Computer Society, 2003, pp. 54.1.

[4] Federal Energy Regulatory Commission, *Energy Policy Act of 2005 (EPACT05)*, Federal Energy Regulatory Commission (FERC), 2005.

[5] K. H. LaCommare and J. H. Eto, *Understanding the cost of power interruptions to U.S. electricity consumers*, Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA (US), 2004 Sep 01.

[6] New York Independent System Operator, *2004 Load and Capacity Data*, New York Independent System Operator (NYISO), 2004.

[7] New York Independent System Operator, *Comprehensive Reliability Planning Process Draft Reliability Needs Assessment*, New York Independent System Operator (NYISO), September 2005.

[8] New York Independent System Operator, *NYISO Installed Capacity Manual Revision 5.1.1*, New York Independent System Operator (NYISO), September 2004.

[9] New York Independent System Operator, *NYISO Power Trends 2004*, New York Independent System Operator (NYISO), May 2004, pp. 23.

[10] New York Independent System Operator, *NYISO Power Trends 2005*, April 2005, pp. 16.

[11] New York State Reliability, *New York Control Area (NYCA) Installed Capacity Requirements for the Period May 2006 through April 2007*, New York State Reliability Council (NYSRC), 2006

[12] North American Electric Reliability Council, *Long-Term Reliability Assessment : The Reliability of Bulk Electric Systems in North America*, North American Electric Reliability Council (NERC), 2005.

[13] D. B. Patton, *2004 State of The Market Report New York ISO*, Potomac Economics, Ltd - Independent Market Advisor to the NYISO, July 2005.

[14] M. Reeder, *Government Intervention into Wholesale Electric Markets to Assure Generation Adequacy*, New York State Department of Public Service November 2002.