The Role of Demand Underscheduling in the California Energy Crisis

Ezra D. Hausman, Senior Associate
Richard D. Tabors, President
Tabors Caramanis & Associates
50 Church Street
Cambridge, MA 02138
ehausman@tca-us.com
rtabors@tca-us.com

Abstract

The lack of demand response to rapidly increasing prices in the California electricity market in 2000 and early 2001 has been identified as one significant factor in the descent of that market into dysfunction. Only consumers in San Diego saw their bills increase with the increase in spot market costs in the state, and they saw it after the fact. Less noticed and now hotly denied was the effect of the day-ahead demand bidding behavior of Pacific Gas & Electric (PG&E), specifically, on the disequilibrium and ultimate collapse of the California electricity market.

The objective of this paper is to evaluate the impact of PG&E's bidding behavior during the crisis. The paper will examine PG&E's underscheduling of load through the creation and bidding of hourly demand curves that bore no relation either to the realities of inelastic demand within the service territory or to the spatial distribution of that demand within the state. The paper concludes that the silent initiator of the California crisis was the economically rational, but structurally destabilizing, bidding behavior of PG&E in their effort to shift purchased energy from the day-ahead to the real-time market. This shift was neither anticipated by the designers of the market nor effectively responded to by California or Federal regulators when its impact was identified.

1. Introduction

The design flaws and dysfunctions of the California electricity market leading up to the crisis of 2000 and 2001 were plentiful enough that the convergence of events leading to the collapse of the market, and the rapid losses of billions of dollars, has been described as a "perfect storm." There has been no shortage of analyses of the underlying factors that led to spiraling prices and spectacular bankruptcies. Often identified as a primary cause of the problem has been the lack of opportunity for demand response by customers in California. The political compromise that underlay restructuring, California Law AB 1890, provided for a guaranteed reduction in customer cost by instituting fixed retail rates for the three Investor Owned Utilities (IOUs) who remained with an obligation to serve load. (California customers had the right to switch to suppliers other than the three IOUs; however, the rate ceiling made competition by other suppliers on the basis of price impractical.)

The retail rates for the initial phase of deregulation were designed to include what was referred to as "head room," which was the difference between the IOUs' allowed charge to customers and their costs of procuring energy in the combination of the Power Exchange (CalPX) and Independent System Operator (CAISO) administered spot markets. This head room was to be used to pay down the stranded costs of the IOUs. After these costs were recovered, the IOUs would be allowed to enter the competitive market released from the rate freeze.

The IOUs had a maximum of five years in which to recover these costs, after which time they would be required to compete for customers on the basis of price. The incentive of the IOUs was to minimize their cost of energy (a positive effect) in order to maximize their paydown on stranded costs and be freed from the rate freeze more quickly. In fact, only San Diego Gas and Electric (SDG&E) was ever able to achieve this goal and charge market based rates to their customers. This occurred in the early spring of 2000, allowing rates in the SDG&E region to skyrocket and provoking a public backlash which contributed to the political refusal, later on, to allow rate increases as a solution to the power crisis [1]. Ironically, this price signal may

0-7695-2056-1/04 $17.00 (C) 2004 IEEE

well have provided the demand response that worked to keep blackouts to a minimum.

This stranded-cost paydown approach was based on the premise that wholesale energy prices would inevitably decrease under deregulation, and that the IOUs would thus be able to procure energy at a cost lower than the amount that they were allowed to charge their customers. This assumption did not turn out to be justified. As soon as the cost of supply began to approach and then exceed the price cap, the IOUs were in a position of losing money on every kWh that they sold.

The separation of the CalPX and the CAISO markets provided a strategic option for the IOUs. The CalPX market was predominantly a day-ahead market from which, by design, the IOUs were to bid to purchase all of their forecast requirements for the next day on an hour by hour basis. It was also the market into which the generators and other potential suppliers into the California market were to offer their energy day-ahead on an hour by hour basis. The result was an hourly intersection of the supply and demand curves which produced the market clearing price for each hourly market. This was to be, in essence, the single state-wide hourly price for wholesale energy, subject only to congestion management and load balancing in real time.

Putting aside the many fine points of the California markets, the CAISO was to complement the CalPX-administered day-ahead market by providing real time operations, as well as a supplemental energy market for load balancing and congestion management functions. This required that the CAISO receive supply and demand bids in the real time market, which could also serve as an energy market of last resort. While the day-ahead and the real-time clearing prices were expected generally to converge, events in the market and changes in the availability of suppliers that occurred between day-ahead and real-time might reasonably cause these prices to differ for any particular hour or day.

The intent of those defining the market rules was that the IOUs would bid all of their forecast demand into the CalPX market, but this was not required by the rules. In fact, there was sufficient room for interpretation to allow any of the IOUs to bid less than their anticipated demand into the day-ahead market in an attempt to reduce the day-ahead market clearing price. The residual demand would be floated forward into the real-time market, and as a result would inflate the real-time clearing price. The key to this strategy was to move relatively small amounts of energy between the markets, playing off the clearing prices so as to minimize the overall cost of wholesale energy procurement. It is difficult to pinpoint when this strategy was begun, but it was clearly in full swing by the beginning of 2000 on the part of PG&E.

The Federal Energy Regulatory Commission (FERC) identified this behavior pattern in November 2000 [2], finding that the CPUC-mandated exclusive reliance on the CapPX and CAISO markets produced a chronic underscheduling of load by PG&E. This led to the real-time market not being a market of last resort for balancing load, as intended, but becoming instead a market in which up to 30 percent or more of the energy for the CAISO control area was procured.

The objective of this paper is to analyze the effects of PG&E's load underscheduling on spot market prices and, ultimately, the role of this behavior in the failure of the California electricity market. It is acknowledged that more work remains to be done before the full impact of strategic underscheduling on the functioning of electricity markets is well understood.

2. Incentives for Underscheduling Load

Most competitive electricity markets are designed with multiple settlement periods, generally including at least a day ahead market and a real-time load balancing and congestion management market [3]. This makes sense from an operational perspective: most power is intended to be scheduled based on day-ahead purchases or long-term contracts, so scheduling coordinators (SCs) can submit balanced transaction schedules in advance, subject to modest balancing and congestion management adjustments through the real-time market when actual load and system conditions are known.

In a well-functioning electricity market, the forward and day-ahead clearing prices are based on an expectation of real-time prices, so the clearing prices in the two markets will tend to converge. Market participants might be indifferent to when they purchase power if price were the only consideration. However, managing uncertainty and risk for both buyers and sellers provides a significant incentive for transacting in the day-ahead market, and for forward contracting in general. If a persistent price difference were to exist between the day-ahead and real-time markets, opportunistic market participants would respond in such a way as to eliminate it – to arbitrage it away, until the only difference is a small risk premium. Suppliers who foresaw a higher real-time price for power, for example, would have an incentive to withhold supplies from the day-ahead market, bringing the two markets back into line.

The California wholesale market was designed along these lines. Purchases and sales of wholesale power were to be cleared day-ahead through the CalPX and submitted as balanced schedules to the system operator (the CAISO), which would accept schedules and resolve any day-ahead transmission congestion. (Further refinements were to be made through the CalPX-administered "day-of" market, although this market never had enough liquidity to have a significant role in the market.) In real-time, the CAISO would run a balancing energy market, and use incremental and decremental bids and offers in the CAISO market to resolve any congestion which appeared in real time. It was generally anticipated that well over 95% of the energy would be scheduled in the forward markets, with the balance (positive or negative) transacted in real time.

The market structure and the resulting economic incentives turned out to be inconsistent with this expectation. If most of the energy is scheduled in the day-ahead market, and in the absence of any apparent mechanism for arbitrage, it is not in the interest of the market participants to have the prices in the two markets converge. Since both markets operate on the principle of the single clearing price auction, sellers would like to see a higher price in the more robust day-ahead market if that is where most of the power is going to clear, even at the expense of a lower price in real time. Buyers have the opposite incentive. Which direction the market will take, and the ultimate effects on the functioning of the system, depend on the relative market power held by market participants on either side and on the details of the market rules.

In California's "deregulated" electricity market, the purchasers of electricity (the IOUs, specifically PG&E, SCE and SDGE) were in a curious position: while they had a high degree of market concentration on the demand side of the market, they had very little flexibility in managing costs. The three IOUs were only permitted to purchase power in spot markets, either in the day-ahead market through the CalPX, or from the CAISO in real time. Even power from their own generation assets had to be offered through these spot markets, and there was no provision for hedging price risk by negotiating long-term power purchase contracts. Further, their obligation to serve load was inflexible and price elasticity was nonexistent due to fixed retail rates. As a result, there was no direct way to control or recoup costs associated with increases in the price of wholesale power.

Given their dominant position on the demand side of the market, however, the IOUs were in a position to know the aggregate supply curve from which their load would be purchased with reasonable accuracy, and to exert some control over the demand curve. By under-bidding load in the day-ahead market, they had the ability to make that market clear at a lower price and for a lower quantity, effectively transferring demand into the real-time market; this is shown schematically in Figure 1. The diagonally shaded region illustrates the value of such a strategy for a purchaser in a two-settlement market, assuming that all supply not sold day-ahead is offered into the real-time market at the original price.

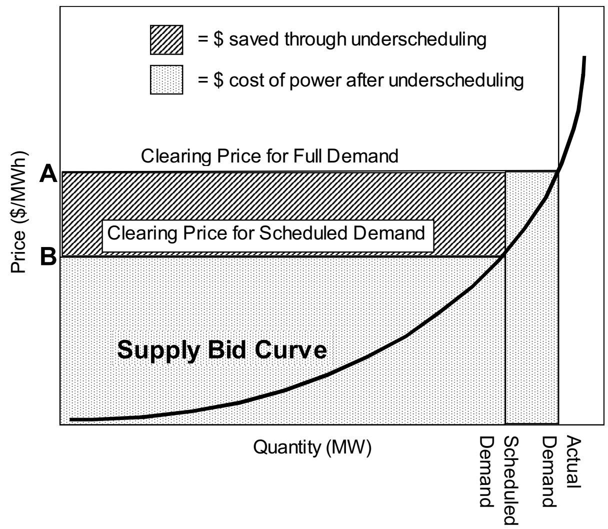

Figure 1. Benefit to LSE of underscheduling load

Had the IOU represented in Figure 1 bid full demand requirements in the day-ahead auction at the true value of lost load, the full requirement would have been procured at the clearing price labeled **A**, where the supply curve intersects with actual demand. On the other hand, by effectively withholding (or low-ball bidding) a portion of demand in the day ahead auction, the IOUs can try to obtain a portion of their demand at a lower price, labeled **B** the Figure. If the supply curve in the real-time market is composed of the unsold energy from the day ahead market offered at the same price, the IOU can purchase the remaining requirements at the original clearing price, for the total savings represented by the diagonally-shaded region.

Figure 2 shows a sample PG&E bid curve for September 26, 2000, hour 17. Although PG&E's load for this hour was over 11,600 MW, the utility bid a curve which declined to unrealistically low prices well below this quantity. They would have obtained their full requirements for this hour in the day-ahead market only if the clearing price had been about half of what it was. As a result of this bidding strategy, PG&E served 70 percent of its load requirement at a price which, presumably, was significantly lower than it would have been had the demand bids been more representative of actual load; the remaining 30% was obtained in the real time market.

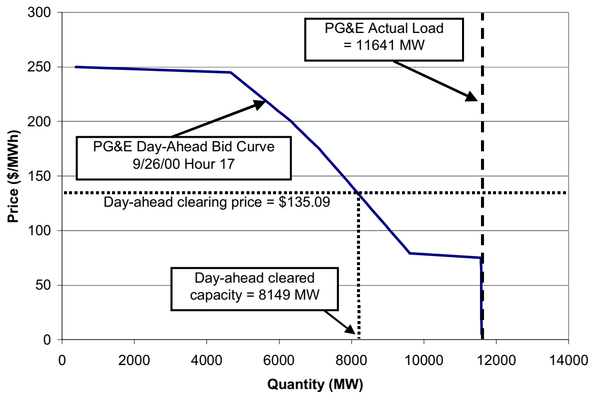

Figure 2. Example of PG&E underscheduling, September 26, 2000 hour 17.

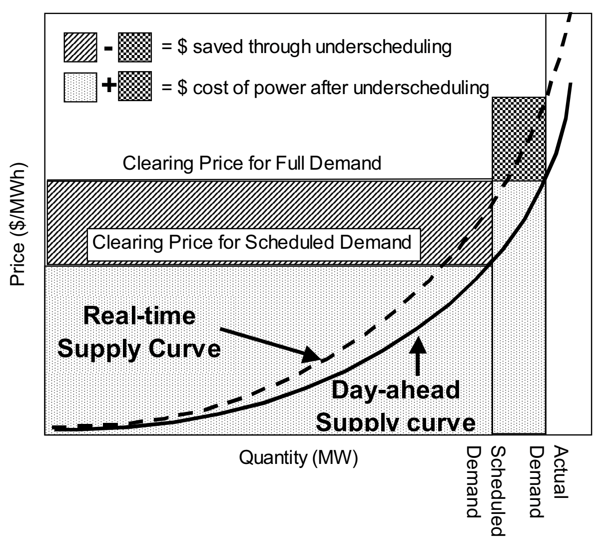

Figure 3. Impact of price divergence between day-ahead and real-time markets on effectiveness of load underscheduling strategy.

In fact, the real-time spot price is likely to be higher than the day-ahead price due to the shorter lead time for fuel contracting and start-up operations, among other factors. This situation is more likely to resemble that in Figure 3, which shows a divergence of the day-ahead and real-time supply curves. This makes underscheduling in the day-ahead market somewhat less attractive, and probably keeps it in check to some degree in many markets.

In California, however, the incentive for load Underscheduling remained strong for two reasons. The first was the effective requirement that all power be offered and purchased through the CalPX, so the volume in the day ahead market was much higher than it would have been had there been a robust bilateral market. The second, paradoxically, was the price cap in the real-time market. Because the IOUs knew that their exposure was limited to a price which was not far in excess of the "true" clearing price, there was little disincentive to aggressively push power into the real-time market, to the detriment of the smooth functioning of the market and operational efficiency.

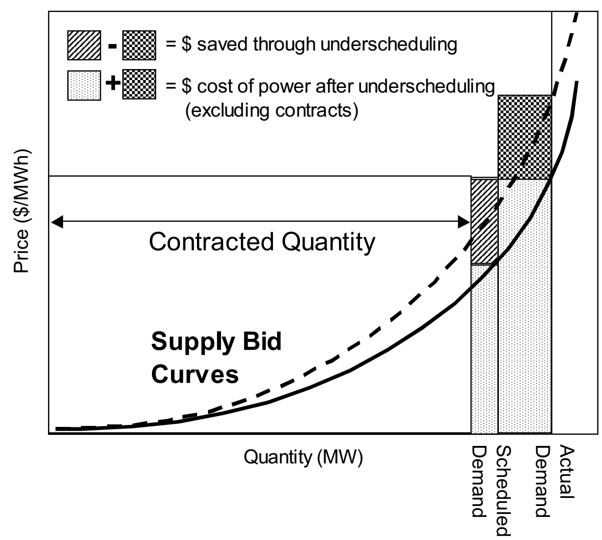

Figure 4. Impact of long-term power purchase contracts on effectiveness of load underscheduling strategy.

3. Absence of Long-Term Contracts

Much has been written about the legislatively-mandated lack of long-term forward contracts for power in the California energy market, leaving the three IOUs completely unhedged in the face of soaring energy costs [*e.g.,* 4]. A less noted result of this market design flaw, however, is its role in promoting underscheduling of the form described here. Figure 4 displays the case in which most (perhaps 80%) of an IOU's requirements are met through preexisting contracts. In this case there is little or no incentive to underschedule load, for two reasons. First, the region of the supply curve over which underscheduling can have an effect is much smaller, so there is a much smaller potential impact on the clearing price. Second, the volume of power for which this price effect would provide an advantage is much smaller. As there is uncertainty in the shape of both the day-ahead supply curve and the real-time supply curve, it is unclear that there would be any advantage to underscheduling load if such contracts were in place. During the restructuring process in California it was generally assumed that power prices would decline under deregulation and that long-term power supply contracts would become new stranded liabilities. As a result, no such bilateral arrangements were permitted in the California wholesale electricity market to mitigate the incentive for load underscheduling.

Given the lack of long-term contracts and the price cap in the real time market, underscheduling in the California day-ahead market was an economically rational strategy for PG&E and other IOUs. There was nothing in the market rules to prevent this behavior, and the ISO was well aware of this incentive structure and practice [5]. As a result, PG&E underscheduled load increasingly throughout 2000, up to the point where the collapse of the market and the decline and ultimate bankruptcy of the CalPX led to the state taking over power purchasing functions in January 2001.

4. Extent and Method of underscheduling

Figure 5 shows the weekly average on-peak (weekday hours ending 8 through 23) load for PG&E for May through August 2000, along with the amount of capacity that cleared in the day-ahead market during this period. Figure 6 shows the percent of PG&E's requirements that were actually scheduled day-ahead, again as a weekly average of on-peak periods. From May through July, PG&E transacted and scheduled an average of 81% of on-peak load on a day-ahead basis. In August and through the end of the study period, that percentage dropped to closer to 62% of on-peak load. At the same time as this shift in underscheduling, the price cap in the CAISO-administered real-time market dropped from $750 per MWh to $500 (July 23) and finally to $250 (August 6). This convergence of events supports the idea that a lower price cap in the real-time market diminishes the risk and cost of underscheduling load by IOUs, making the practice more attractive as a cost management measure.

PG&E underscheduling was accomplished by submitting a demand curve to the CalPX market in which much of the anticipated demand was bid at a price well below that at which the market was expected to clear. Figure 7, based on PG&E's bid curves, shows the amount of power that would have been purchased during on-peak periods at price levels from $25 to $500 per MWh. For comparison, Figure 8 shows the on-peak average day-ahead clearing price. From mid-June forward, the clearing price in the day-ahead market rarely reached below the $100/MWh level, and frequently exceeded $200/MWh. Despite this fact, PG&E consistently bid between 10% and 40% of their load at a price below $100/MWh, in full awareness that it was unlikely to clear the day-ahead market.

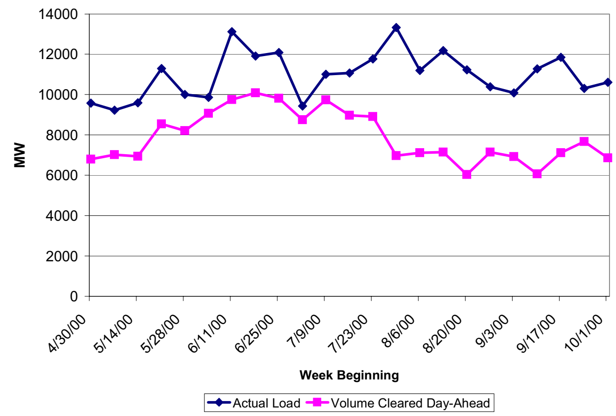

Figure 5. PG&E weekly average on-peak load vs. capacity cleared day-ahead, May through August 2000.

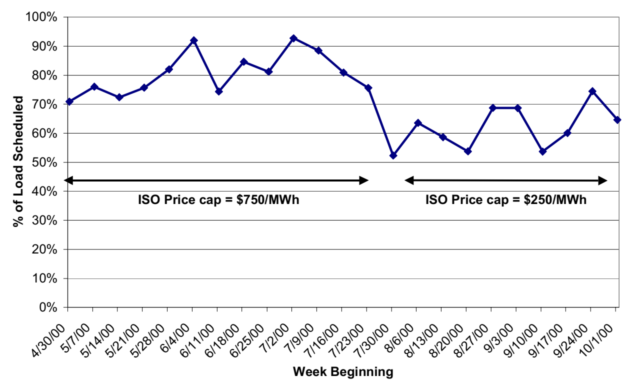

Figure 6. Weekly Average Percent of PG&E On-Peak Load Scheduled with CAISO Day-Ahead, May through August 2000.

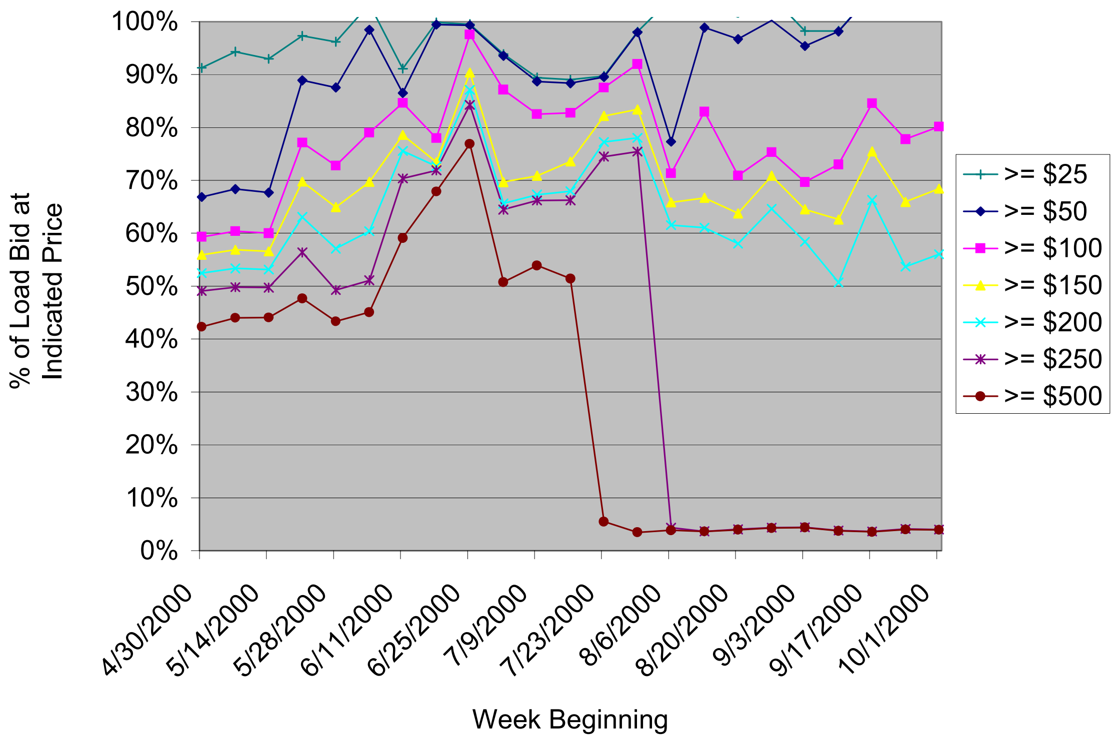

Figure 7. Weekly Average On-Peak Percent of PG&E Load Bid by Price Level

5. Supplier Response

If IOUs yield to their economic incentive to underschedule load, the economically rational response from generators is either to offer power at what they anticipate the "true" clearing price to be, or to withdraw supplies from the day-ahead market and offer them only at real-time. They would adjust their offers in this manner whether or not they were aware of the sort of bidding behavior underlying the price signals. As prices show a persistent low bias in the day-ahead market relative to real time, it is rational for any supplier to find ways to move power into the higher priced market until the two markets come back into line.

In general, electricity market rules are not designed to facilitate suppliers selling large amounts of power in the real-time market. For out of state suppliers in California, for example, it was impossible to secure firm transmission in advance for power which was overtly intended for sale in the real-time market. One solution to this dilemma is to offer the power in the day-ahead market, but also to bid load into that market which is not expected to materialize in real time—that is, to overschedule load in compensation for the underscheduling on the part of the IOUs. When the overbidding entity found itself with excess power on its hands, as expected, this power could be released to the real-time market for whatever turned out to be the real-time clearing price.

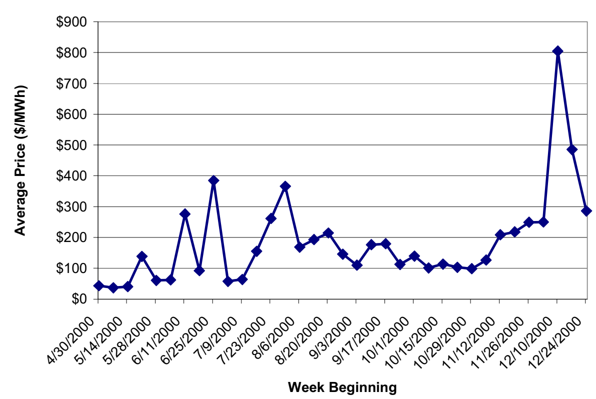

Figure 8. Weekly average on-peak Day Ahead Clearing Price in the California Power Exchange

This kind of load overscheduling can take many forms. In New York [6], energy providers can engage in “virtual bidding”, by bidding for power day-ahead even if they have no obligation to serve load in the New York ISO region. In real time they make this power available to the ISO as balancing energy, which in turn uses it to compensate for underscheduling by IOUs. The suppliers run the risk of being unable to sell the power in real-time, as it was not scheduled against any real load. In return, they capture the price difference between the artificially depressed day-ahead market and the real time cost of power, and thus apply pressure on these two markets towards price convergence.

In California there was no specific mechanism for virtual bidding, and only entities who had an in-state obligation to serve load were permitted to submit schedules to the ISO. However, there were no rules about how well the load schedule of any particular entity submitted to the ISO had to match actual delivery. If a generator scheduled more power than was consumed by their load, the power was simply released to the system at the real time price—exactly the same effect as seen with virtual bids in New York.

This was the only way for out-of-state generators to reliably sell power into the real-time market to compensate for the persistent underscheduling of load by PG&E and other IOUs. Because transmission capacity could not be secured without a day-ahead schedule, there was no other way to guarantee delivery in the real-time market. The only option for suppliers was to schedule power imports day ahead as if they had a contractual obligation to serve load, and then to release the energy to the real-time market, at the real-time price, to meet the “unanticipated” demand of the IOUs.

Many out-of-state generators engaged in this practice, but Enron became most famous by giving it a nickname (“fat boy”) and by nominally scheduling large amounts of power over small transmission links that could not possibly handle the flow. This has been cited as evidence of market manipulation, and it was indeed a fast and loose application of the market rules in order to get firm transmission access for delivery to the real time market. However, the net result of this behavior, on the part of Enron and others, was to mitigate the distortions caused by underscheduling on the part of the IOUs. This was explicitly recognized by FERC in their March 26, 2003 Show Cause order [7], in which they concluded that overscheduling endeavors, and specifically “fat boy”, were a legitimate response to underscheduling by PG&E and did not represent market manipulation.

Not only is overscheduling to load a reasonable economical response to underscheduling by IOUs, it is essential to the functioning of the market. Without virtual bidding of some sort, the ISO’s unit commitment and congestion management systems would have been operating on the basis of grossly distorted information, or at best in complete disregard of the schedules submitted by the utilities. This was recognized as a problem in the New York market, which began operation without any mechanism for virtual bidding, as day-ahead and real-time prices began to diverge early and due to load underscheduling. In New York the problem was successfully addressed by the introduction of virtual bidding for both supply offers and demand bids, which, if cleared, become financially binding day-ahead schedules [6]. Mismatches between day-ahead and real-time prices were effectively mitigated by market forces, as market participants exploited the opportunity to extract profit from these mismatches through virtual bidding.

6. Equilibrium or divergence to market failure?

As the IOUs underscheduled load to depress prices in the day-ahead market, power suppliers responded by effectively withdrawing power from that market and offering it at real time, or by offering it at the price they anticipated would be the true clearing price. Now the IOUs had run out of room to maneuver by shifting load to real time, but the effectiveness of their cost containment efforts had turned out to be short lived and ultimately counterproductive. The market operator had insufficient and inaccurate information for matching resources to load, and was thus hobbled for unit commitment and congestion management operations. The IOU-submitted schedules had almost no value as load forecasts, and apparent congestion in the day-ahead market had no predictive value for congestion in real time. Efficient unit commitment and congestion management are essential to maintaining reliability in electricity markets, but both were compromised as a result of the corruption of information which should have been conveyed through balanced, day-ahead schedules.

In terms of the auction structure for electricity, the day-ahead market now more closely resembled a pay-as-bid market than a clearing price auction. Suppliers had a strong disincentive from offering power at close to marginal cost, and would more reasonably offer power day-ahead at what they expected the clearing price should be. The real-time market remained glued to the bid cap, providing no meaningful market signal whatsoever. This resulted in an overall loss of efficiency, as units were no longer necessarily bid in merit order (lower cost, more efficient units first) and to generally higher prices more of the time.

This dysfunctional state of affairs may represent a highly inefficient equilibrium, as neither the producers nor the consumers could improve their position be unilaterally rectifying their bidding strategy. Worse, it may be characterized by continuing efforts by either side to control costs by playing the markets off each other, causing a positive feedback loop that eventually leads to a complete market failure. It is hard to conclude from theory which result would occur, but history suggests that the resulting market conditions were not sustainable for long.

7. Discussion

While it is difficult to identify the primary cause of the market failure in California at the end of 2000, it is clear how a dysfunctional market was amplified into into a train wreck, bringing the major utilities up to or over the brink of bankruptcy. This factor was the requirement that all power be purchased in the day-ahead or real-time spot market, coupled with a complete lack of demand response. When the spot markets experienced supply shortages, the resulting price spikes impacted not the fraction of energy requirements that would have been affected had long-term contracts been in place, but the entire energy requirement for the CAISO-controlled region.

This paper has focused on underscheduling behavior by PG&E in relation to both the market dysfunction and the magnification of the associated cost. Underscheduling by IOUs, and specifically PG&E, contributed to the market dysfunction and supply shortages by sending meaningless signals to the ISO through the day-ahead market, inhibiting the ISO's ability to perform both unit commitment and congestion management functions. The overscheduling by some suppliers in response helped to mitigate these problems, but also contributed to a breakdown in the efficient dispatch of power to meet California's energy needs. Underscheduling by IOUs was rational behavior given the constraints written into the California market rules, and was made perhaps irresistible because of the large volume of power transacted in the day-ahead and real-time spot markets. Had the utilities been allowed to hedge their requirements through long-term contracts, the load underscheduling would never have reached a level at which it caused reliability problems on the grid. Any market failures which did occur would have been far less costly to the utilities, to consumers and to the state, and could probably have been remedied with far less draconian and expensive measures than those being proposed and implemented today.

[1] Van Vactor, S. and F. Pickel, "Money, Power and Trade", *Public Utilities Fortnightly,* Vol. 139, No. 11, May 15, 2001.

[2] San Diego Gas & Electric v. Sellers of Energy and Ancillary Services Into Markets Operated by the California Independent System Operator Corp. and the California Power Exchange Corp., 93 FERC ¶ 61, 121 at 61, 359-62 (2000).

[3] Stoft, Steven, *Power System Economics: Designing Markets for Electricity*. Wiley-Interscience, 2002.

[4] Jurewitz, John, "California's Electricity Debacle: A Guided Tour", *The Electricity Journal,* May, 2002, pp. 10-29.

[5] California ISO Department of Market Analysis, "The Firm Transmission Rights Market: Review of the First Nine Months of Operation, February 1 – October 31, 2000", November 30, 2000.

[6] New York Independent System Operator, Inc., "Evaluation of the Impact of Virtual Trading on the Summer 2002 New York Electricity Markets", filed as FERC Docket Nos. ER01-3009-000, ER01-3153-000 and EL00-90-000, 12/16/02.

[7] FERC Docket No. PA02-2-000, *Final Report on Price Manipulation in Western Markets,* 3/26/03, Chapter VI.