

COMPETITIVE ELECTRICITY MARKETS AND THE CASE OF CALIFORNIA

by

CLAY D. DAVIS

B.S., Kansas State University, 2005

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF ARTS

Department of Economics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2008

Approved by:

Major Professor
Dennis L. Weisman

Abstract

The primary purpose of this report is to address the potential benefits and drawbacks of competitive electricity generation. A number of countries have introduced various forms of competition into the electric utility industry. The most notable attempt in the United States and the focus of this report took place in California in the late 1990s. This report is divided into two parts. The first covers the history of the electricity industry by reviewing influential policies, cost of service regulation, and concluding with incentive regulation. The second part discusses the potential benefits and drawbacks of a competitive generation sector, through the lens of the California experiment.

Government policies have dramatically changed in the last twenty years. Many of these changes were aimed towards increasing competition within the generation sector and have made a competitive generation sector possible. Whether these policies are correct has been the focus of much debate. A competitive generation sector could potentially operate more efficiently than under traditional regulatory regimes. Whether this potential will be realized is in question. This report assesses this question by looking at the events that took place in California.

The competitive wholesale markets in California functioned properly for nearly two years before the events of the “California Electricity Crisis” took place. This showed that a competitive wholesale market is possible once certain criteria are met; most importantly adequate competition to reduce a firm’s potential to exercise market power. The “Crisis” in California showed what can happen if these criteria are not satisfied. Nevertheless, there is still much optimism about the potential benefits of competitive electricity markets.

Table of Contents

List of Figures	iv
Acknowledgements	v
CHAPTER 1 - Introduction	1
CHAPTER 2 - History of the Electricity Industry	2
Regulatory Policies and How They Shaped the Industry	2
Industry Structure Pre-Reregulation	4
Rationale for Reregulation.....	8
CHAPTER 3 - The California Electricity Crisis.....	10
Market Structure Pre-Restructuring.....	10
Reasons for Restructuring in California	11
Transition Towards and Setup of the Restructured System.....	13
Potential Problems with the New Market Structure	15
The California Electricity Crisis Spring 2000	20
Solutions That Could Have Reduced the Effects of the Crisis	31
CHAPTER 4 - Conclusions	34
References.....	36

List of Figures

Figure 3.1 Wholesale Spot Prices, California PX (Day-Ahead), CAISO Real-Time. <i>Source: Sweeney 2002.</i>	21
Figure 3.2 PX Day-Ahead Demand and Price, July 1999. <i>Source: Sweeney 2002.</i>	22
Figure 3.3 Natural Gas Spot Market Prices: PG&E Citygate and Southern California Border. <i>Source: Sweeney 2002.</i>	24
Figure 3.4 Natural Gas Spot Prices in California: PG&E Citygate and Southern California Border. <i>Source: Sweeney 2002.</i>	24
Figure 3.5 Average Generating Capacity Off-Line. <i>Source: Sweeney 2002.</i>	28
Figure 3.6 Price Versus Quantity for California PX, November 2000. <i>Source: Sweeney 2002.</i>	29
Figure 3.7 A shift in variable demand in an electricity market. <i>Source: Borenstein, 2002.</i>	30

Acknowledgements

I would like to thank my Major Professor, Dr. Dennis L. Weisman, for valuable comments, continuous help, and overall guidance with my work on this report.

CHAPTER 1 - Introduction

The electric power industry has sparked much debate in the past 30 years, leading to many changes in beliefs about the industry's structure. What was once thought of as an industry best served by vertically integrated monopolies, now has many people believing certain aspects of the industry are better served by competitive markets. Many changes are taking place based on these new beliefs; whether they are correct will play out in the coming years.

The transmission and distribution networks were and still are thought of as being best served by one firm. This naturally led to the idea that the other aspects of electricity generation (e.g. generation and retail) were also best served by natural monopolies. This is now not necessarily believed to be the case. The various sectors are more than just linked, but they rely on each other for proper functioning of the entire grid. Because electricity cannot be easily stored, supply (generation) and demand (load) must equal each other in real-time. This can make coordination challenging, leaving many industry observers to continue to believe the monopoly structure is still the best design.

Countries have recently experimented with creating competitive markets for the generation sector, and some in the retail sector. One of the most notable attempts, and the focus of this report, was in California in the late 1990s. The California experiment has left many industry observers leery at the possibility of a competitive electricity industry.

This report will start out by covering the history of the electricity industry and how we arrived at where we are now, covering the policies that have shaped the industry along the way. It will then focus on current regulation and the specific problems that competition might solve. Then it will cover the case of California including the events leading up to, through, and past what is known as the "California Electricity Crisis". Finally, it will cover possible solutions to the market problems, concluding with some of my own thoughts on the events that transpired and the future of the industry.

CHAPTER 2 - History of the Electricity Industry

Regulatory Policies and How They Shaped the Industry

Regulatory policies have greatly affected the electric utility industry up to this point and are driving many of the changes currently taking place. The first major act to affect the industry was the Public Utility Holding Company Act of 1935 (PUHCA) (Warkentin-Glenn 2006). This act was enacted shortly after the Great Depression, in response to many public utilities going bankrupt. The act prohibited holding companies from owning both utility and nonutility assets. Utility companies were generally seen as safe investments, so holding companies would use utility assets to reduce the risk of owning nonutility assets. This in turn exposed the utility to the risks of these other assets and around the time of the Great Depression forced many of them into bankruptcy. Since the enactment of PUHCA no utility has filed for bankruptcy. This act was repealed on August 8, 2005 leaving much uncertainty as to the future of the utility industry (Warkentin-Glenn 2006).

In 1978 another piece of legislation was enacted that greatly affected the utility industry. As discussed by Warkentin-Glenn (2006), the Public Utility Regulatory Policies Act (PURPA) came about in response to the energy crisis of the 1970s. Its intent was to promote diversity and efficiency with respect to the generation of electricity. This policy emphasized the use of cogeneration facilities, which allow the excess heat from the generation of electricity to be collected and used for heating and other industrial processes. It also put in place policies to allow facilities meeting certain efficiency criteria to be classified as Qualifying Facilities (QFs). Utilities were required to purchase electricity from these facilities at a price no greater than the utilities' avoided costs from not generating the electricity themselves. This, according to Borenstein (2002), was one of the major reasons that California faced high retail electricity prices. These QFs do not face any sort of cost or price regulation (Warkentin-Glenn 2006). According to Joskow (1989), the Federal Energy Regulatory Commission (FERC) left most of the implementation of PURPA policies to the states. This would lead to large discrepancies in contract prices with QFs between these states. If prices were set too high it would encourage unnecessary QF capacity and the opposite if prices were set too low. With

bundled prices it is difficult to determine the utilities' "avoided costs" of generation. Under a vertically integrated structure, firms have an incentive to favor their own generation over that from QFs. This could cause them to misrepresent their avoided costs, in order to avoid purchasing from a QF. Nevertheless, PURPA marked a big step towards promoting independent generators and a competitive generation sector.

In 1992 the Energy Policy Act (EPAAct of 1992) was passed with the intent of promoting competitive electricity generation (Warkentin-Glenn 2006). While PURPA promoted independent generators, the EPAAct of 1992 gave them the power to sell directly to final consumers. This act created a new class of generators known as Exempt Wholesale Generators (EWGs). These generators could be owned by utilities, holding companies, or unaffiliated investors. The EPAAct of 1992 also led to what is known as "wholesale wheeling," by forcing utilities to allow EWGs access to their transmission lines. This meant EWGs could sell directly to the final customer (Warkentin-Glenn 2006).

FERC Orders 888 and 889 were enacted in 1996. As discussed by Warkentin-Glenn (2006), Order 888 reduced the incentive for utilities to discriminate between their own generation and that of the EWG, by allowing the utility to recover the stranded costs associated with unused generating capacity. The utility would recover these costs in the form of a transmission access charge, charged to EWGs. Another product of this legislation was the formation of Regional Transmission Organizations (RTOs) each of which would have an Independent System Operator (ISO) control the transmission wires over a given region. The hope was that these RTOs would lead to cheaper retail prices by reducing the multiple rates charged from electricity being transmitted over multiple utilities' transmission lines. These RTOs would run a regional spot market for electricity. They would not answer to any one utility, but to all of them collectively. This would eliminate any incentive they would have to favor one party over another. As noted by Joskow (1996), there is still the problem of the access charge being invariant to the congestion on the network. Order 889 forced the utilities to make available network congestion information to everyone on a real-time basis and led to the development of open access same-time information systems (OASIS) (Warkentin-Glenn 2006).

The final major piece of federal legislation that takes us up to the present was the implementation of the Energy Policy Act of 2005 (EPAAct of 2005). According to Warkentin-Glenn (2006) this policy created an organization overseen by the FERC to promote system reliability. It also allowed for faster depreciation and tax incentives for investments in transmission infrastructure, so as to promote infrastructure investment. The most controversial implication of the EPAAct of 2005 was the repeal of PUHCA to promote investment in infrastructure, by allowing more parties to own utility assets. This may cause firms to consolidate leading to larger corporations that are more difficult for state and federal regulators to oversee. Also, state regulation may become more challenging with more utilities operating over state lines. The EPAAct of 2005 also reformed PURPA, no longer forcing utilities to purchase electricity from QFs. Lastly, customers in some areas can now choose between different generators.

Federal policy has done much to change the outlook of the electricity industry in a short amount of time. In thirty years the electric utility industry went from being dominated by natural monopolies to being regarded as a competitive industry, at least in some of its sectors.

Industry Structure Pre-Reregulation

As mentioned above, electricity generation was traditionally thought of as a “natural” monopoly industry. Joskow (1997, p. 119) states,

For nearly a century, the electricity sector in all countries has been thought of as a “natural” monopoly industry, where efficient production of electricity required reliance on public or private monopoly suppliers subject to government regulation of prices, entry, investment, service quality and other aspects of firm behavior.

Retail customers had to buy their electricity from the regulated monopoly supplier who had the legal right to distribute electricity in their region (Joskow, 1997). The regulated monopoly is legally required to provide all electricity demanded in their region, as well as planning for growth in demand and making necessary improvements to its infrastructure. The utility and the state regulatory commission determine the retail prices. These retail prices are “bundled” prices in that they combine the prices for generation, transmission, distribution, and retail into a single price. A great deal of coordination is

required between these four functions to maintain system stability. Supply and demand not only must equal each other in real-time, but frequency, voltage, and other physical characteristics of the system are required to stay within tight constraints. Even though generators might not be affiliated with each other, they all affect the stability of the entire system. To improve coordination and stability, the United States is divided into three interconnection regions, the Eastern, Western, and Texas Interconnects. These interconnects have developed operating protocols that all participating parties abide by to keep the system functioning smoothly.

Utilities have been regulated primarily by cost of service (COS) regulation, also known as rate of return regulation (RORR), where the utility earns a return based on their costs of providing electricity (Comnes et al., 1995). Joskow (1972) determined what criteria actually determine the return a firm is allowed to earn. These include capital costs, reasonableness of the firm's request, whether interveners are involved, the type of firm, and judgment of the commission. Using the model Joskow developed and the given criteria, one can consistently determine the return the firm will be allowed to earn. A few possible drawbacks to RORR discussed by Cave, Majumdar and Vogelsang (2002), include limited incentives for innovation and cost reduction, over-capitalization, the high cost of regulation, and excessive risk imposed on consumers. The direct tie between earned revenue and costs can reduce the incentive of a firm to minimize costs because no return will be earned on foregone costs. RORR may actually cause overinvestment in capital, what is known as the "Averch-Johnson Affect." In order to ensure the firm acquires the necessary capital it may give a higher than fair return. Here a fair return is the lowest return on capital necessary to acquire sufficient quantities of electricity. This may cause the firm to acquire more than the necessary level of capital. Regulation can be costly depending on the extent of firm oversight, the level of scrutiny a firm receives from regulators, and the frequency of rate hearings. Consumers can face the risk of changes in a firm's operating costs. If a firm's costs are decreasing then it would not push for a rate hearing to reduce prices, but would likely want to schedule a rate hearing if costs were increasing, in order to push for a price increase. This puts the consumer in the position of being faced with the downside risk, while the firm would gladly take any upside gain.

As discussed by Comnes et al. (1995), there are two main types of COS/RORR regulation: 1) Pure and 2) Regulatory Lag. Pure COS regulation continually updates rates based on costs. Continually updating rates based on costs will reduce any incentive the firm has to reduce costs because any decreased costs will be passed on as lower rates to the consumer, instead of increased profits to the firm. That is to say, no good deed (cost reduction) goes unpunished. To increase the incentive of the firm to decrease costs COS with regulatory lag can be used. Regulatory lag is where the time between rate cases is set and the longer the time between cases the greater the firm's incentive to reduce costs, by allowing it more time to realize increased profits. The utility does not always know when the next rate case will be or how much of their cost reductions/increases will be passed on in the form of lower/higher retail prices. These uncertainties can also reduce the incentive to decrease costs.

These potential downfalls discussed by Cave, Majumdar and Vogelsang (2002) and Comnes et al.(1995) led to the development of performance-based regulation (PBR). The main intent of PBR is to reduce or remove the link between a firm's costs and its retail rates. PBR can lead to increased incentives to lower rates, costs, or improve non-price performance. Some of the main types of PBR include revenue caps, price caps, and sliding scale regulation. The discussion here will focus on price cap regulation (PCR), as this is the method used to regulate retail electricity prices in California. As discussed by Cave, Majumdar and Vogelsang (2002), PCR can provide greater incentives for technological innovation and cost reduction over RORR. Through PCR, the firm will be rewarded for cost reductions through the realization of increased profits. PCR forces the firm to behave more like a firm in a competitive market. By removing the ties between costs and revenue, the firm will no longer have the incentive to accrue unnecessary expenses. Doing so would cause the firm to earn less profit. The firm assumes the risk/reward of increasing/decreasing costs. Under RORR, consumers faced the risk of increasing costs leading to a rate hearing, while decreasing costs would be realized by the firm in increased profits, going unchallenged until the next rate hearing. See Cave, Majumdar and Vogelsang (2002) for further discussion on the benefits and drawbacks of the various regulatory regimes.

The argument has been made that under COS/RORR the firm is obligated to perform efficiently, making incentive regulation unnecessary. Weisman and Pfeifenberger (2003) show why incentive regulation is generally superior to traditional regulation methods. They show that incentive regulation is superior for two reasons. The first is that incentives motivate increased performance better than performance mandates. They generalize this argument to the historical superiority of free market economies versus centrally-planned economies. Efficient outcomes are difficult to mandate because there is uncertainty over future market events. Furthermore, decentralizing control to the regulated firm can provide incentives for the firm to use its private information about costs and demand to make improvements in efficiency. Secondly, if a firm made efficiency improvements once it switched to some form of incentive regulation, it does not mean the firm was operating inefficiently under prior regulation. Efficiency improvements are learned over time as a “discovery process.” So the argument should be which method creates a greater incentive to “discover” these efficiency gains. Much empirical research has shown that incentives better promote these performance gains.

Weisman (2002), showed that the performance of PCR can be undermined by regulatory opportunism. Here regulatory opportunism refers to actions taken by the regulator in the short run that reduce the profitability of the regulated firm and undermine the performance of PCR in the long run without technically breaching the price cap commitment. For example, regulators could artificially reduce barriers to entry in a way that greatly liberalizes competitive entry. This could have the effect of making the effective (market) price lower than the capped price, causing a reduction in the firm’s profit without breaking the price cap contract. The regulator could also undermine the incentives created by PCR by not honoring the regulatory commitment. If firm profits became greater than projected, then the regulator could be under pressure to change the capped rate. Doing so would reduce the incentive of the firm to increase profits through efficiency gains because future profits would be decreased by lowering rates. The firm could also push for price adjustments if its return became sufficiently low. Whether rate adjustments should be granted depends on the reasons for the low return. Rate increases should not be granted if low returns are the result of poor performance. If the reason for the low return is due to some criteria outside of the firm’s control, then it may have an

argument for a rate adjustment. These examples show that the performance of PCR can be greatly reduced through both regulator and firm actions.

Rationale for Reregulation

A number of factors are motivating the push towards a more competitive electric utility sector. Some of the potential improvements are lower prices, increased investment in innovative technologies and infrastructure, improved electricity quality and system reliability, increased labor productivity, more efficient use of capital, and new value-added services. Most of the changes are aimed at the investor-owned segment, which accounts for over 75 percent of U.S. retail electricity sales (Joskow, 1997).

One of the major contentions is that the move towards a competitive market in electricity generation will lead to lower prices. In 1997, electricity prices in the United States averaged 8.4cents/kWh for residential customers and 4.7cents/kWh for industrial customers, which are near the low end when compared to other Organization for Economic Cooperation and Development (OECD) countries (Joskow, 1997). Roughly half of the retail cost of electricity is due to generation and this is the area where most of the cost reductions will likely take place (Joskow, 1996). While U.S. prices are low compared to other OECD countries, there is considerable variability between U.S. states. In the late 1990s most of the states pushing for a competitive generation sector were states with higher than average retail prices. The two main reasons, according to Borenstein and Bushnell (2000), for these price differences were largely due to investments in nuclear facilities and purchase contracts with QFs. These two reasons for higher retail prices would not be removed by transitioning to a competitive market for electricity generation. During the 1970s and 1980s many states were pushed into building costly nuclear facilities, which was thought to be the clean generation source of the future. These facilities turned out to be very costly to build and greatly exceeded their cost of construction estimates. Making the generation sector more competitive would not eliminate these sunk costs. The other main reason for these price differences was due to utilities signing long term contracts with QFs at higher than avoided cost prices, largely at the urging of the states. These contract costs would not go away by changing the market structure.

Price decreases could also come through more efficient use of capital and increased labor productivity. Joskow (1997, p. 124) notes,

The operating performance of both fossil and nuclear units varies widely even after controlling for age, size and fuel attributes, and some utilities have performance that lags behind industry norms (Joskow and Schmalensee, 1987).

Also according to Joskow (1997, p. 125),

Regulatory cost recovery rules may encourage utilities to continue to operate generating plants even though it would be economical to close them.

Until about 1980, utilities realized economies of scale by building larger generating plants, when new technologies (e.g. natural gas turbines) led to smaller efficient scales (Joskow, 1997). These smaller generating facilities are able to be built in a manufacturing setting, reducing the build cost and resulting in more uniform costs across generating facility projects. Also, shorter lead times are required to build these smaller units, leading to generating capacity responding faster to market signals.

As discussed by Joskow (1997), costs could possibly be reduced by increasing labor productivity, which has happened in the restructuring of electricity markets in other countries. The gains from labor productivity might not be as significant in the United States as compared to these countries, since pre-restructuring productivity in the U.S. is significantly higher.

Other improvements that might not necessarily lead to lower prices are improvements in innovation and value-added services. Technological innovation in the electric utility industry has not kept up with other industries (Munson 2005). Many of the technological innovations were developed by other industries and modified to work for the electric industry. For example, natural gas turbine technology was originally developed by the aerospace industry and later modified for the electricity industry. It is believed that moving towards a competitive market structure will lead to increased investment in technology, specifically for the electric industry. Most of these technology improvements will likely be aimed at improving value-added services for retail customers. Customers may have more options regarding bill structure, quality of

electricity, and reliability of service. These possible benefits, in addition to lower prices, have led to increased optimism about competitive generation and retail sectors.

CHAPTER 3 - The California Electricity Crisis

The California Electricity Crisis began in the spring of 2000 and ended by the spring of 2001. The Crisis was characterized by high and very volatile electricity prices, caused by a number of factors. These factors will be covered in this chapter along with events leading up to, during, and after the Crisis.

Market Structure Pre-Restructuring

The three main investor-owned utilities (IOUs) in California were Southern California Edison (SCE), Pacific Gas and Electric (PG&E), and San Diego Gas and Electric (SDG&E) (Sweeney 2002). These utilities provided electricity to 78% of California's customers and accounted for 72% of the generating capacity. The market was vertically integrated in that generation, transmission, distribution, and retail services were all provided by a single utility operating over a specified franchise service area. Municipalities served the remaining 22% of customers and owned the remaining 24% of generating capacity. These municipalities had local governing boards and were exempt from state regulation, which in California was the California Public Utility Commission (CPUC). The fact that these municipalities were exempt from state regulation will be an important distinguishing factor when the events of the Crisis take place, leaving the municipalities largely unaffected.

As discussed by Sweeney (2002), the CPUC oversees all four sectors of the electricity market. The CPUC determines the rate of return that the utilities are allowed to earn on their investment. In order for a utility to earn a return on an asset the CPUC has to deem the asset to be "used and useful." Once the asset is determined to be used and useful, the utility can earn a return on it, even if the asset turns out to be a bad investment. This transferred the risk from the utility to the retail consumer, who would pay for the asset and the return through retail prices. This means that retail prices are set mainly by cost of service and secondly by market conditions. Retail prices would be

affected by market conditions usually by being indexed against historical prices. Setting retail prices until the next rate case gave the utility an incentive to delay rate hearings when costs were falling and push for a rate hearing during times of increasing costs. The regulators have to balance the interests of the utilities who want to hold off on a rate hearing during times of decreasing costs and the consumer who would want the hearing in order to decrease retail prices.

Reasons for Restructuring in California

According to Borenstein (2002, p. 192),

In 1993, California's average retail electricity price was 9.7 cents per kilowatt-hour, compared to the national average of 6.9 cents.

High retail prices were the major reason for promoting restructuring in California (Sweeney 2002). While most people believe restructuring could lead to lower retail prices, the causes of higher retail prices in California were independent of the market structure. The two main causes of high retail prices were investments in nuclear generation that turned out to be poor investments, and expensive contracts with Qualifying Facilities. California invested heavily in nuclear generating facilities, which turned out to be more costly to build than estimates had forecasted. Under rate of return regulation if these costs were deemed used and useful, then they would be passed on to consumers through higher retail prices. Also, under PURPA, the CPUC pushed the utilities to enter into high priced contracts with QFs. These contracts were at prices much greater than the avoided costs of the utilities. By 1994, approximately twenty percent of the generating capacity in California was from QFs, and these two causes of higher retail prices were considered sunk costs and would not be reduced through restructuring (Borenstein, 2002). Some industry observers believed that these costs should be transferred, through restructuring, from the consumer to the utility shareholders because the utilities were responsible for making these poor investments. This transfer of costs would not take place.

Generating capacity would become an unavoidable problem in the coming years, as electricity demand outpaced growth in capacity. It was thought that a competitive generation sector would allow capacity to better keep up with electricity demand

(Sweeney 2002). Due to assets having to be deemed used and useful by the CPUC, the process for increasing capacity can be inordinately slow. This is necessary because the CPUC is making decisions on behalf of the consumers who are bearing the risks of the investment. In a competitive generation sector the firm would determine whether the risks warranted the rewards, making it unnecessary to go through the bureaucratic process of determining whether the asset is useful.

A competitive generation sector would likely lead to more distributed generation, which for a few reasons is cheaper in California (Sweeney 2002). These would generally be smaller natural gas facilities, due to the decreasing economies of scale exhibited by smaller, more efficient turbines. Since California does not have a large supply of coal, natural gas will likely be the fuel chosen for new generating capacity. Natural gas is a cleaner burning fuel, as compared to coal, allowing the facilities to be located closer to the point of electricity consumption. Allowing the facilities to be closer to the point of final consumption will result in two attributes that allow for numerous potential benefits. First, less investment in transmission capacity will be required, which accounts for a significant portion of the overall investment in new generating capacity. Fewer line losses will occur as a result of the electricity traversing shorter distances to the customer, which in addition to using a cleaner fuel will also make the facility more environmentally friendly. Additionally, with shorter transmission distances less land use will occur. Secondly, allowing the facility to be located closer to end use will make the steam part of the cogeneration process, which is used for heating and industrial processes, more useful.

Wolfram (2003) conducted a study that showed significant efficiency improvements had been realized by states going through electricity restructuring, as compared to states that have not. She separates the firms into two groups based on whether they had passed restructuring legislation as of April 2001. Her results show that employees per megawatt fell by 8% and non-fuel operating expenses by 14%. With 800,000 MWs of capacity in the U.S., this would equate to industry costs falling by a total of approximately \$3-\$4 billion. The study also showed that heat rates, a measurement of the amount of energy needed to generate a unit of electricity, improved by 2-2.5% after restructuring. These are short-run effects, so in the longer term industry costs could move in either direction. The gains to the post-restructured states could

decrease costs in the immediate term, but may increase over the longer term. Also, these results may be understated if firms in non-restructured states are already making changes in anticipation of future restructuring. If nothing more, it does show significant differences between the two groups.

Transition Towards and Setup of the Restructured System

In February 1993 the CPUC published a staff report, known as the “Yellow Book,” which outlined various strategies for market restructuring (Sweeney 2002). Out of the “Yellow Book” came the “Blue Book,” which proposed a plan to restructure the electricity industry in California.

Sweeney (2002) examines the issue of “stranded costs” associated with the utilities generation assets and the parties that should be responsible for covering these costs. These “stranded costs” included investments by utilities in excess generation capacity that might not be needed in a competitive wholesale market. If the costs were passed on through higher retail rates, then the customers purchasing from the utilities’ generators would pay for these costs and ultimately be driven towards purchasing from other generators. On the other hand, some people believed the utilities should pay for these costs since they were responsible for making these investments in the first place. Under the regulated structure, the utility was required to make the necessary investments to accommodate future demand growth. The solution proposed in the “Blue Book” was a “Competition Transition Charge.” This solution added a charge onto all retail purchases whether or not the electricity was purchased from a utility generator. In other words, this was a non-bypassable charge. The revenue from this charge would go to the utility covering any stranded costs.

Included in this plan was the requirement that utilities divest at least 50% of their generating assets, but most ended up divesting nearly all of their generating assets. These assets on average sold for 70% over book value, which greatly reduced the time required for the competition transition charge to be in effect. This divestiture was necessary to eliminate any incentive the utility would have to favor its own generation over that of a competing generator. The utilities would still have monopoly rights over the distribution of electricity, but consumers could choose their retail provider (Sweeney 2002).

According to Sweeney (2002), in 1995 the CPUC proposed a restructuring order, which was their plan to setup the new markets. Out of this order came Assembly Bill 1890 that enacted the policy into legislation. In addition to the CTC already discussed, this plan created an Independent System Operator (ISO) to manage the grid and a Power Exchange (PX) to manage the wholesale market. For the retail market, cost of service regulation would be replaced with performance based regulation. Retail prices charged by the utilities would be capped at January 1, 1996 levels, out of concern that the CTC would cause retail prices to increase above pre-restructured levels. Without the price caps, if retail prices did not decrease by an amount greater than the CTC, then retail prices would increase to levels greater than pre-restructuring levels. Utilities were able to remove this price cap if they removed the CTC or, once they recovered their sunk costs, allowing retail prices to increase above the capped levels. The bill forced utilities to run the generation, transmission, and distribution as three separate companies, handing over control of their transmission lines to the ISO. All electricity generated by the utilities had to be sold through the PX or ISO ancillary market. This would keep the generation and distribution segments separate, negating any chances of self-dealing. Retail and distribution would be separate, leaving distribution as a regulated monopoly.

The Power Exchange, discussed above, would operate one-day-ahead and day-of-wholesale markets (Sweeney 2002). Market clearing prices for these markets would be determined on an hourly basis. The market clearing price would be determined by accepting all supply bids required to meet demand, starting with the lowest priced bid. All sellers would receive the price of the last bid accepted, this is also known as uniform pricing. Under perfect competition, the market clearing price would be the marginal cost for the last generator's bid accepted. As was noted earlier, the three IOUs were required to sell all electricity generated through the PX, but use of the PX was optional for other generators, as well as municipalities. The operation of this exchange would be overseen by the CPUC with the Federal Energy Regulatory Commission (FERC) being the final authority.

The California Independent System Operator (CAISO) would operate the transmission grid by dispatching enough supply to satisfy demand (Sweeney 2002). According to Sweeney (2002) the CAISO was also responsible for maintaining system

stability by balancing supply and demand on a real-time basis. Some generators were required to operate in order for the entire grid to function properly. This would naturally alter their bidding strategy, which will be discussed below. The CAISO collected generation and load schedules from both sides of the market on an hourly basis. These schedules were required to match up to within 2 MW of each other, although no penalty was enforced if these schedules did not meet this requirement, leaving them susceptible to manipulation. To correct for discrepancies between these schedules and actual amounts required, the CAISO would run an imbalance market, known as the ancillary services market. A set amount of reserve capacity was maintained for this market that could be brought on-line with short notice to satisfy these discrepancies. This market was operated within the hour and also used a uniform market clearing price. It should be noted that the electricity sold in this market is the same electricity that could be sold on the PX.

Potential Problems with the New Market Structure

The market design had many potential problems. Some of these problems were known in advance, while others would become known once the new markets were in operation. This section will examine the nature of these problems. Some of these problems have to do with the physical nature of electricity itself, while others were created by the market design.

While supply and demand must equate in real-time, electricity cannot be supplied instantaneously. Some generators are able to come online within a few minutes while others can take up to a few hours and this can make coordination efforts challenging. This problem becomes magnified when the various industry functions are no longer coordinated by a single firm. Facilities that require more time to come online can be costly to start up and shut down. They need to run more than just on an hourly basis to make operating worthwhile. Over time, this would naturally lead to smaller generating facilities that are able to start and stop on shorter notice and at lower cost. This effect will result in more suppliers and ultimately a more competitive wholesale market. The extent of this increase in competition could be reduced by a single firm owning multiple generating facilities.

The generation and load schedules submitted to CAISO, as mentioned above, were required to be within 2 MW of each other, although there was no penalty if this criterion was not met (Sweeney 2002). These schedules are forecasts of future demands, leaving them open to manipulation by both buyers and sellers. For example, a generator could manipulate its estimates to sell more electricity into the ancillary market if there was a chance of a higher clearing price there. The protocols for operation of this market were made public, and this could affect the generators' bidding strategies. An initial price cap of \$250/MWh was put in place on the ancillary services market. This price cap would, in effect, become a price cap on the PX market. This would result because of an operating procedure requiring that if a shortage occurs then all of the utilities would share equally the extra cost of purchasing enough electricity through the ancillary services market to eliminate the shortage. Instead of a utility bidding above \$250/MWh on the PX and paying all of the cost, they would let a shortage occur and split the costs of acquiring the needed supply with the other utilities (Sweeney 2002). Since the generation and load schedules matched up on paper, the CAISO would not know there was a shortage until the hour that this specific generation and load was called upon. This would make it challenging for the CAISO to keep supply and demand equated, in addition to maintaining system stability. If the price cap had been on the PX markets instead, then there would not have been a possibility of a shortage because any buyers not receiving their electricity on the PX markets due to the price cap would receive it on the ancillary market at the market clearing price.

Sweeney (2002) noted that the designers of the restructured system assumed that the generators would bid competitively, in other words at their marginal cost of the last unit supplied. Due to the multiple markets that generators and utilities could bid into, in addition to other factors, the optimal bidding strategy might exceed their marginal cost. The fact that markets clear hourly every day provides the utilities with a great deal of historical price and quantity information. Borenstein and Bushnell (1999) note that in a dynamic setting firms may learn to compete less aggressively with each other. Additionally, with repeated interaction the firms will be able to more easily punish a firm that does not cooperate. These effects can result in higher prices and a loss of consumer welfare. Also, the firms with a small market share might still have an incentive to exhibit

market power. For example, suppose a generator has 1000MW at a marginal cost of \$35/MWh and the market clearing price for selling all 1000MW is \$40/MWh (Sweeney, 2002). The generator would have profits of $(\$40/\text{MWh} - \$35/\text{MWh}) * 1000\text{MW}$ equaling \$5000/h. Now suppose the generator could exhibit market power by bidding 900MW at \$35/MWh and 100MW at \$50/MWh, knowing the higher bid will not be accepted. The generator is in effect withholding 10% of its generating capacity. Then, assuming the generator has a market share of 2.5%, it is withholding 0.25% [$2.5\% * 0.1$] of the overall market capacity. Assume the market is at a point on the demand curve where the price elasticity of demand is 0.1, meaning a 1% change in quantity supplied will result in a 10% change in price. This scenario is quite possible when the system is near capacity. This 10% price increase will lead to a market clearing price of \$41 [$((41-40)/40)=0.1$]. The profit earned on the 900MW accepted bid is $(\$41 - \$35) * 900\text{MW}$ or \$5400/h. By exhibiting market power in this way the generator was able to increase profits by \$400/h or [\$5400-\$5000]. This example shows that generators with only a small market share can significantly influence price. The requirement for this to happen is the system being at or nearly at full capacity, resulting in an almost vertical supply curve. If other firms were employing similar strategies, then the price could be driven up even further, even without the firms colluding with one another. The threat of new entrants could reduce the incentive of a firm to exercise market power (Borenstein and Bushnell, 1999). While this may be true, generating facilities can take years to build, resulting in extended periods of time for a firm to exercise market power before another firm is able to come on-line.

The inability of the IOUs to enter into long-term contracts could and did lead to problems for the IOUs and generators in the restructured system (Sweeney 2002). IOUs were required to purchase all electricity through the PX or ancillary market and were not allowed to enter into any long-term contracts with generators. This was done out of concern that the IOUs could give preference to their own generators and bypass the wholesale markets. Another concern was that going around the wholesale markets would reduce competition within these markets. While these concerns were justified, this created many challenges for the IOUs, which will be covered presently.

The IOUs faced many risks in the new setup, some caused by market design, some by the physical system, and others by the fuels used to generate electricity. By not

being allowed to enter into medium and long term contracts with the generators, the IOUs faced undue price risks caused by a volatile spot market (Sweeney 2002). These contracts were a way for the IOUs to reduce their price risk associated with acquiring electricity. These contracts have historically been the method used to reduce the effects of dealing exclusively with spot markets. The municipalities were exempt from state regulation and were not required to carry out transactions solely through the spot market, leaving them largely unaffected during the Crisis period (Sweeney 2002).

Nearly 60% of the three IOU's generation was fueled by natural gas, which can have highly volatile prices (Sweeney 2002). The inability of the IOUs to enter into contracts forced them to deal exclusively with the wholesale markets and the resulting spot prices. Natural gas prices in California can be more variable than other parts of the country, due to supply constraints coming into and also within the state. In addition to electricity generated using natural gas, roughly 13% of California's electricity demand is provided for through hydroelectric power, mainly from the Pacific Northwest. The amount of electricity generated through this method is tied closely to the amount of rainfall received in this area, so during times of low rainfall less electricity is available for purchase through this generation method. Hydroelectric power is one of the least costly methods of electricity generation. So, not only would less rainfall result in a reduction in supply, but also a reduction in one of the lowest cost methods of electricity generation. This could increase prices by causing a reduction in capacity and also by having to shift to other more expensive forms of electricity generation.

Many of the potential problems were problems because they caused a reduction in possible capacity. The excess generating capacity planned for under the previous regulatory regime could not be guaranteed under the restructured system. Through traditional regulation the utility was responsible for planning for and providing excess capacity to account for increasing electricity demand. Under the new system, it was thought that increasing wholesale prices would send a signal to investors that excess profits could be earned, resulting in the building of more generating capacity. This thinking fails to take into account the fact that generating facilities cannot be built overnight, with facilities generally taking three to four years from the time of permit application to coming online (Sweeney 2002). Electricity generation is capital-intensive

and goes through classic “boom and bust” cycles, resulting in periods of increasing and decreasing amounts of excess capacity. These periods of low excess capacity could cause high wholesale prices, especially during times of near peak generation (Sweeney 2002). Conversely, during times with large amounts of excess capacity, some generators might not be able to cover their operating costs.

The retail markets provided additional risks for the IOUs to deal with. The IOUs would operate distribution as a regulated franchise, based on incentive regulation (Sweeney 2002). As discussed above, retail prices would be capped at January 1, 1996 levels during the competition transition charge period, out of concern that the CTC would lead to higher retail prices. According to Joskow and Kahn (2001), this cap was set at about \$60/MWh and projections for future wholesale prices were about \$30/MWh, leaving significant head room between these prices. As noted by Sweeney (2002), the CPUC only had jurisdiction over the IOUs, leaving the municipalities to enter contracts and just as importantly, if not more so, pass increasing wholesale prices on to consumers through higher retail prices. Independent retail providers were also not bound by retail price caps. This would create an incentive for consumers to switch between the IOUs and these other providers whenever it was advantageous to do so. The IOUs were forced to offer service to any one in their region, but consumers were not required to purchase from the IOUs, allowing them to come and go as they pleased. This would not only make demand planning difficult for the IOUs, but also put them in the position of earning little profits. During periods when other firms were offering retail rates lower than the regulated IOU rates, consumers would choose these other firms. Conversely, when these other firms were offering retail prices higher than the regulated prices, consumers would switch back to the IOUs. There was no contract forcing consumers to stay with the IOUs for a specified time. While this could have been a significant issue, Joskow and Kahn (2001) observe that 90% of the retail demand through 2000 was provided by the IOUs.

If wholesale prices did exceed the retail price cap, then the IOUs could offer bonds to account for the difference. So instead of consumers paying higher rates today they will pay for the bonds for years to come. Consumers will not only pay for the difference between wholesale and retail prices, but also the interest paid on these bonds. With no method to pass these excess costs on to consumers, the retail price cap forced the

IOUs to bear the wholesale price risk. No method was in place to send a signal to consumers, by way of higher retail prices, when wholesale prices increased. This would prove to be a major downfall of the restructured system.

The California Electricity Crisis Spring 2000

The new market design began operation on March 31, 1998, about two years before the events of the California Energy Crisis took place. What went wrong? And why did the system work for two years before problems arose, ultimately leading to the abandonment of the restructured system? For starters, many of the potential problems discussed in the previous section actually took place. Generating capacity became limited during the Crisis period, but was not a factor prior to this. This allowed the system to appear to function properly before the Crisis. It will also be shown that according to Borenstein (2002) two solutions, variable retail prices and long-term contracts, would have allowed the restructured system to operate more efficiently.

Wholesale prices in California greatly increased during the Crisis period, predominately due to a culmination of factors causing a reduction in supply and excess capacity, without a reduction in demand. In the spring of 2000, wholesale prices greatly increased. Peak prices on the California PX exceeded \$400/MWh and \$500/MWh in June and July of 2000, respectively (Sweeney 2002). Figure 3.1 shows the weighted average wholesale prices on the PX, separated into northern (NP15) and southern (SP15) California, and the CAISO real-time market (Sweeney 2002). California's generating capacity had been slow to increase relative to the increases in electricity demand throughout the early 1990s. The effect of decreasing excess capacity did not become greatly pronounced until the Crisis period. The other causes of high wholesale prices would not have had the effect they did if excess capacity had not been decreasing throughout the previous decade, leading to a tight market when the system was generating at near peak capacity. The graph shows prices in northern and southern California diverge towards the end of 2000. In the winter months, Northern California becomes an exporter of electricity to the Pacific Northwest. Prices in northern California increased significantly as a direct result of less hydroelectric power being generated by these northwestern states.

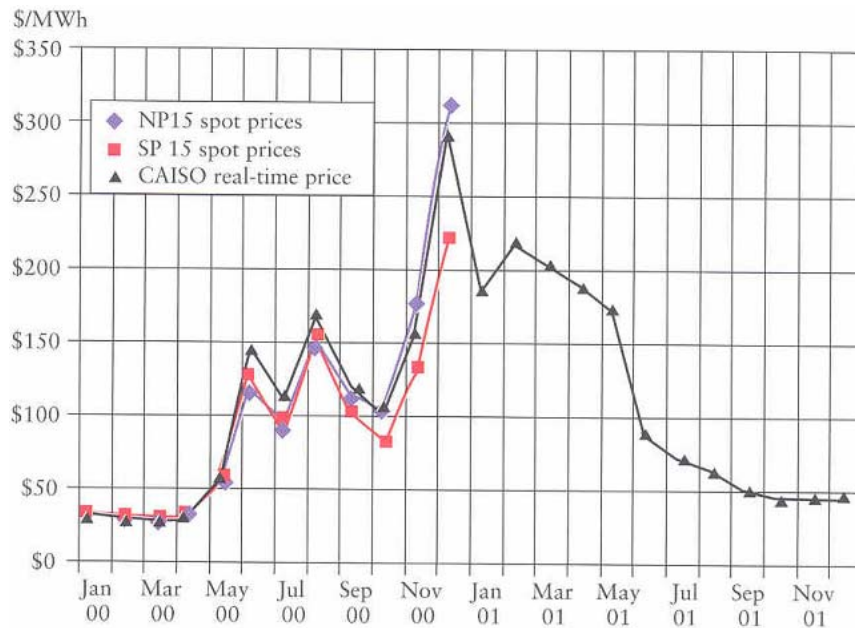


Figure 3.1 Wholesale Spot Prices, California PX (Day-Ahead), CAISO Real-Time.

Source: Sweeney 2002.

The western United States is connected by transmission lines known as the Western Interconnect, and these lines allow electricity transmission within and between these states (Sweeney 2002). As discussed by Sweeney (2002), transmission of electricity into and out of California would take place with these western states. Any wholesale price discrepancies between regions would lead to arbitrage and hence price discrepancies quickly disappearing. Price discrepancies will disappear given transmission capacity between states does not become congested. So, higher wholesale prices in California would lead to higher wholesale prices throughout the remainder of the Western Interconnect. Fortunately for utilities in other states, the majority of their electricity was procured through long-term contracts, as was done in California before the restructuring. This would largely shield them from the effects of the high wholesale prices of the Crisis period. For these utilities, the Crisis period was relatively short lived compared to these utilities contracts. This would largely negate the effect high wholesale prices would have had over the longer term. In the long term, these contracts would have to be renegotiated, probably at these higher wholesale prices. Long-term contracts would

greatly reduce the volume supplied on the wholesale markets, causing small fluctuations in supply/demand to have larger effects on price. As these contracts expire, more generating capacity would be freed up resulting in increasing imports into California.

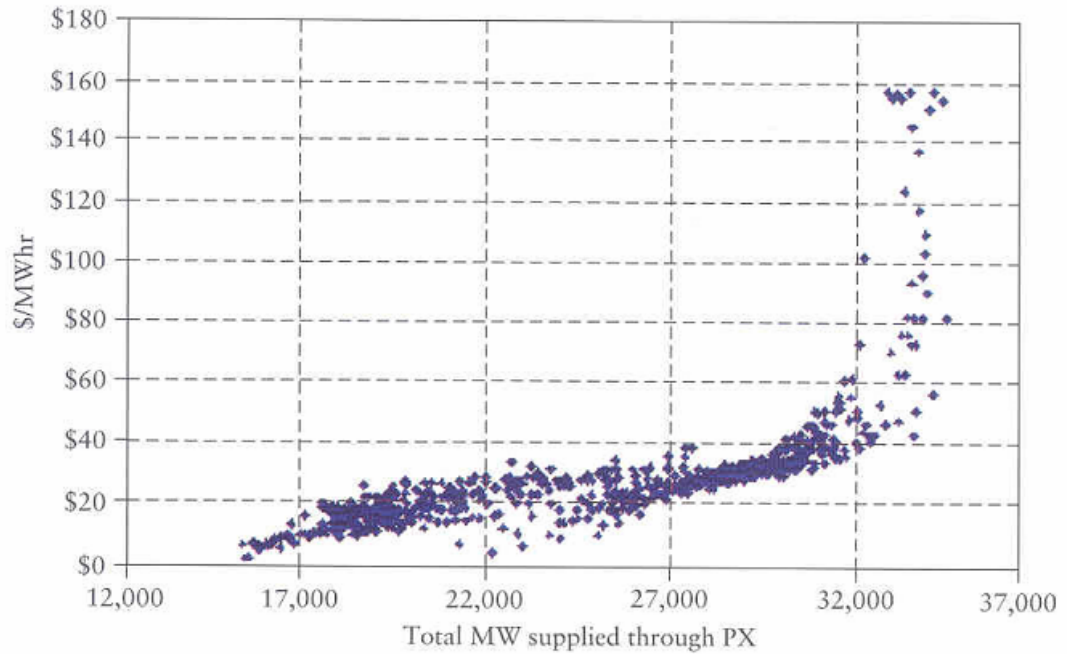


Figure 3.2 PX Day-Ahead Demand and Price, July 1999. *Source: Sweeney 2002.*

Figure 3.2, shown above, is the short-run supply curve for the CA PX in July of 1999 (Sweeney 2002). Each dot represents a market clearing price and quantity on an hourly basis during the month. Notice the hockey stick shape representing significantly increasing marginal costs, assuming competitive bidding, as the system nears capacity. The supply curve becomes very inelastic after 32,000 MWs, resulting in prices that increase dramatically with relatively small changes in quantity. The less efficient generators making up the vertical portion of the supply curve would only be willing to come online with significantly higher prices, as compared to the more horizontal portion of the supply curve. There are fewer points on this plot in the nearly vertical region, showing that most of the time wholesale prices were well below maximum capacity. This resulted in average prices that were significantly lower than peak prices. Over time, average market clearing prices were moving towards the limiting region, in direct

response to new capacity failing to keep pace with growing demand. This led to significant increases in average wholesale prices.

Once demand is in the nearly vertical portion of the supply curve, anything causing increases/decreases in supply would have a significant effect on the market clearing price. Two causes of greatly increasing marginal costs during the crisis period were natural gas constraints into and within California and limited quantities of NOx emission permits. These two factors would have a greater effect on the less efficient generators, as compared to the more efficient ones (Sweeney 2002).

Natural gas coming into and within California was limited by pipeline capacity. These capacity constraints led to higher natural gas prices, greatly increasing the marginal costs of all natural-gas-fired generators (Sweeney 2002). Recall that nearly 60% of the three IOU's generating capacity came from natural gas. Figure 3.3 and 3.4, below, show natural gas prices at two points in California. Prices increased dramatically during 2000, especially during the winter, while prices in Louisiana did not exceed \$10/MMBtu throughout this time period. California is an electricity exporter to the northwest U.S. during the winter months, increasing the demand for natural gas. Figure 3.4 shows spot market prices at the same locations during the summer of 2000. Why did these price spikes not occur to the same level during the summer of 2000, considering electricity demand was near full capacity during the summer and winter of 2000? One would think that during the summer months all natural gas facilities would be generating, given the high wholesale prices during this time. While less electricity was leaving California during the summer, more was being demanded within the state. One would logically conclude that natural gas prices were being manipulated, especially during the winter months, though not much can be concluded without a more thorough analysis of the situation.

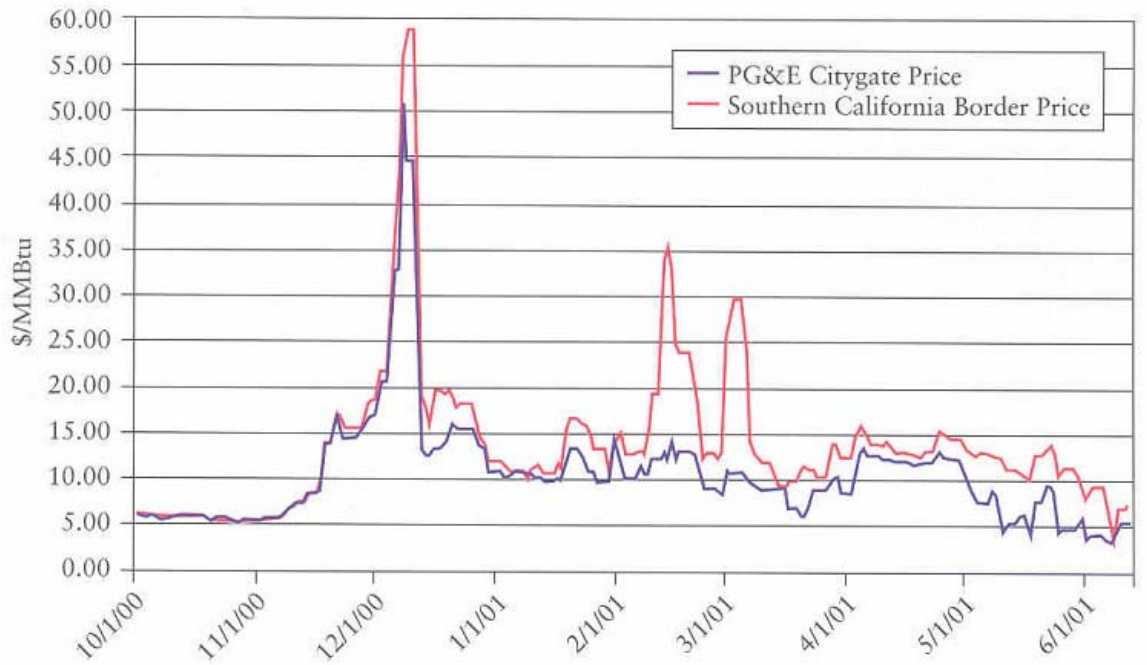


Figure 3.3 Natural Gas Spot Market Prices: PG&E Citygate and Southern California Border. *Source: Sweeney 2002.*

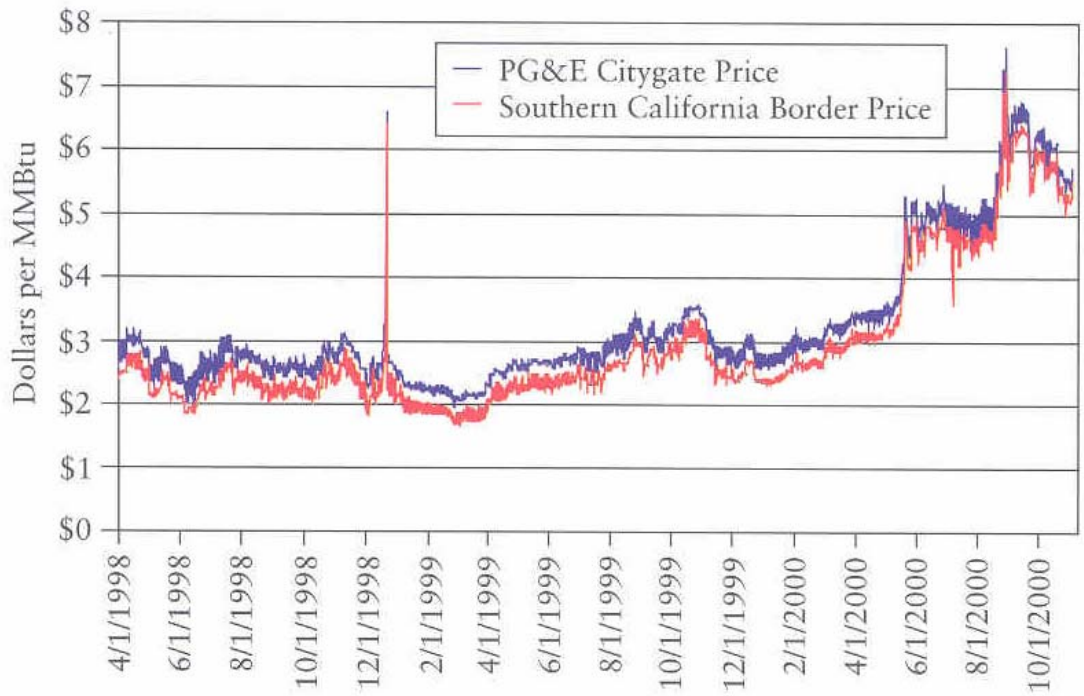


Figure 3.4 Natural Gas Spot Prices in California: PG&E Citygate and Southern California Border. *Source: Sweeney 2002.*

Increasing natural gas prices have a much greater effect on the less efficient generators, that make up the vertical portion of the industry supply curve. The following example from Sweeney (2002) shows why these less efficient generators were affected to a greater extent by higher natural gas prices. A typical combined-cycle plant would have a heat rate of about 6.8 million Btu per MWh, so when natural gas prices go from \$2.5/MMBtu to \$6/MMBtu marginal costs will change by \$24/MMBtu [$6 \times 6.8 - 2.5 \times 6.8$]. A less efficient plant, for example Southern California Edison's Highgrove 1 and 2 facilities have heat rates of 13.4 million Btu per MWh. Using the same price levels as before would result in an increase in marginal cost of \$46/MWh [$6 \times 13.4 - 2.5 \times 13.4$]. This increase would shift the generators industry supply curve upwards, but more so for the less efficient generators. With industry demand being closer to full capacity a larger portion of the time, the less efficient plants will be online a larger percentage of the time. Most of the generators purchase some quantity of natural gas through long term contracts, so the above example might be at the far end of the spectrum, but it does show significant differences between generating facilities.

According to Sweeney (2002), the South Coast Air Quality Management District started allotting a set amount of NOx emissions permits in 1993. The total emissions allowed by these permits were significantly higher than actual emissions until 1999. During that year, allowed emissions equaled actual emissions in the region. An example, similar to the natural gas one, can serve to demonstrate how the less efficient plants will be affected much more than the more efficient plants. The effect of these permits will increase the marginal cost curve much more for the less efficient plants. Two properties of these less efficient generators will make emissions significantly more costly for them. Firstly, these generators have higher heat rates, meaning that they require more fuel to produce the same amount of heat/electricity as compared to more efficient generators. Secondly, they do not burn the fuel as cleanly as more efficient plants, causing more emissions for each unit of fuel burned. Joskow and Kahn (2001) state that the least efficient generating units can have emission rates up to 50 times higher than that of the most efficient units. The difference in fuel use rates between these facilities would only be a factor of two, showing that the difference between plants is magnified much more when accounting for emissions. As discussed by Sweeney (2002), permits were trading

for \$4,300 per ton in 1999 and eventually increasing to an average of about \$45,000 per ton for most of 2000. For older, less efficient generators, marginal costs would increase by about \$90/MWh between these two periods, as a result of the emissions permits. The limiting effect of these permits forced generators to pick and choose when to generate electricity. This means that firms would no longer bid according to their marginal cost, but rather their marginal cost plus some opportunity cost associated with generating at one time as opposed to some other time. This competitive bidding would be seen by many as the exercising of market power by these generators. Borenstein (2002) noted that the inframarginal rents earned by the more efficient plants went from \$20/MWh in July 1998 to in excess of \$100/MWh during the Crisis period. In the long run, excessive rents would increase the incentive to replace less efficient plants with newer, more efficient ones, causing these rents to decrease.

Considerable debate has focused on if the generators exercise of market power was the cause of the Crisis. I believe the exercise of market power by the generators was a direct result of the tight market situation and was not the cause of the Crisis, although it further exacerbated the situation by further increasing wholesale prices. As discussed by Borenstein, Bushnell and Wolak (2002), market power can have significant effects on efficiency. With respect to electricity markets, when a lower cost firm withholds capacity then another higher cost firm has to make up for this capacity reduction. The less efficient generator will be supplying the power that, under perfect competition, the more efficient generator would. In the longer term, higher prices caused by market power can lead to new investments that may not have been made at competitive prices. Additionally, these higher prices may cause other industries that have electricity-intensive production practices not to invest in the area, or they may cause existing firms to inefficiently shift to other inputs, in place of electricity.

Multiple studies have shown that firms exercised unilateral market power in California's wholesale electricity markets (Borenstein, Bushnell and Wolak, 2002; Joskow and Kahn, 2001; Wolak, 2003). As noted by Borenstein, Bushnell and Wolak (2002), wholesale expenditures in the summer of 2000 were \$8.98 billion, as compared to \$2.04 billion in the summer of 1999. They find that 21 percent of this increase is due to production costs, 20 percent to competitive rents, and 59 percent to market power. These

investigations into market power in California have not found evidence that suppliers coordinated their efforts to raise wholesale prices (Wolak, 2003). Borenstein (2002) observes that firms with a relatively small market share (six to eight percent), can exercise significant market power during times of near system capacity. More opportunities to exercise market power would become available as the market cleared closer to capacity limits a larger portion of the time.

At times it may appear that firms are exercising market power even when they are not. Competitive bidding by the generators is one explanation that could lead to higher prices. Being able to bid into two markets (PX and CAISO imbalance market) allowed firms to bid more aggressively into the PX knowing that they could likely sell the electricity on the imbalance market the next day (Sweeney 2002). Generators were bidding their capacity into the PX twenty-four times a day, 365 days a year. This knowledge of prior market outcomes allowed the generators to know with greater certainty what they could bid and the likelihood of their bid being accepted. According to Sweeney (2002), the risk of the generators not being paid by the utilities, which became a significant issue further into the Crisis, caused them to bid above their marginal cost to account for this risk. The California Power Exchange commissioned a study to determine whether pay-as-bid pricing would reduce market power over uniform pricing. The report submitted by Kahn et al. (2001) concluded that switching from uniform pricing to pay-as-bid pricing would not have reduced market power during the Crisis period. Figure 3.5 below, shows the generating capacity that went off-line during the Crisis. Most of the generating capacity was off-line for maintenance reasons. This increase in capacity off-line in late 2000 could possibly have been due to generators being run harder than normal during the previous summer, but political officials claimed this was a strategic move to further decrease capacity. These firms would naturally want to perform maintenance on their equipment at times when it would be the least costly/most profitable time for them to do so. Political officials wanted to condemn the generators for trying to maximize profits through the use of market power. The firms were only doing what was in their best interest.

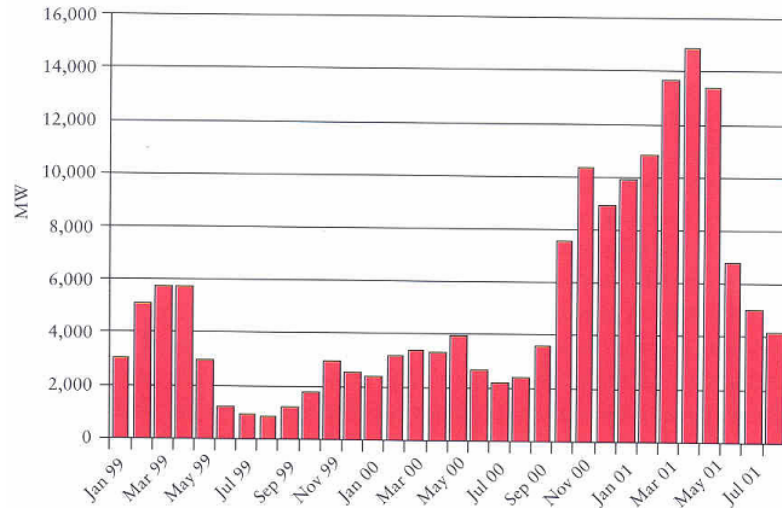


Figure 3.5 Average Generating Capacity Off-Line. *Source: Sweeney 2002.*

The price cap on the ancillary market effectively became the price cap on the PX market (Sweeney 2002). Figure 3.6 below illustrates this result. Notice that the market clearing price on the PX never exceeds the \$250/MWh price cap put in place on the ancillary market. This price cap became binding for many hours during 2000. Also, this figure does not resemble the supply curve of Figure 3.2., showing that generators are no longer bidding at their marginal cost. This could be the result of generators adding on an opportunity cost, increasing bids to account for some sort of risk, influencing prices through market power, or any combination of these. The price cap gave California generators an incentive to export if the price out-of-state exceeded the price cap, further reducing supply in California, typically during times of peak capacity. Political officials found this upsetting, but again these firms were only doing what was in their best interest (Sweeney 2002).

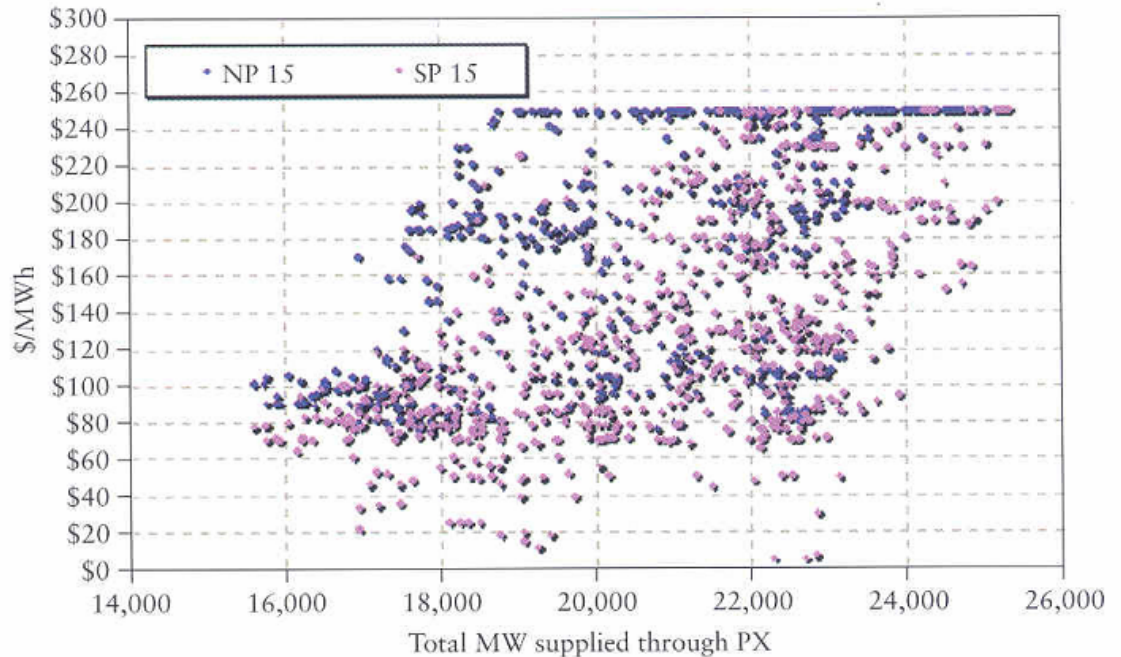


Figure 3.6 Price Versus Quantity for California PX, November 2000. *Source: Sweeney 2002.*

These retail price caps turned the “electricity crisis” into a “financial crisis” for the IOUs, generators, and eventually the state of California (Sweeney 2002). According to Borenstein (2002), variable retail prices, in addition to long-term contracts, could have mitigated many of the problems of the Crisis. With retail price caps in place, increasing wholesale prices did little to reduce demand (Sweeney 2002). This effect is represented in Figure 3.7, with nearly vertical demand imposed on the generators’ supply curve. The distances represented by ‘a’ and ‘b’ show the range of electricity demand throughout a typical day, from times of low demand to high demand. A small increase in overall demand, represented by a shift from the range covered by ‘a’ to that covered by ‘b,’ will result in a large increase in average wholesale prices. The retail price caps shielded the customers from the volatility of the wholesale market, leaving the risk solely with the IOUs. The utilities were required by the state to supply whatever quantity was demanded by retail customers, at a price no greater than the capped price, forcing the utilities to take a loss equal to the difference between wholesale and retail prices. The utilities could only take these losses for so long, eventually becoming insolvent.

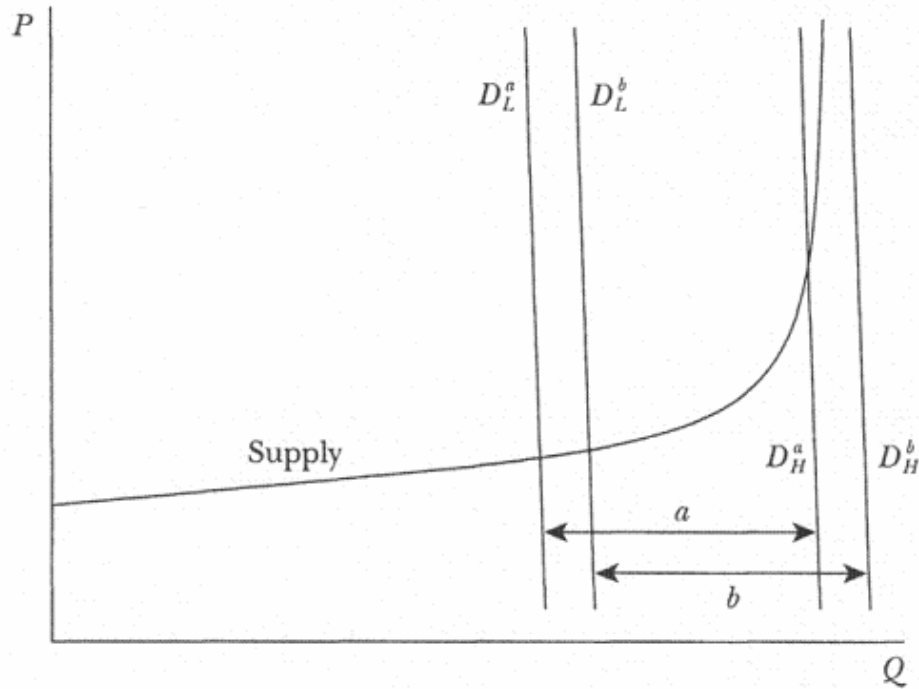


Figure 3.7 A shift in variable demand in an electricity market. *Source: Borenstein, 2002.*

The losses created by the difference between wholesale and retail prices reduced the utilities' creditworthiness with the generators (Sweeney 2002). The credit risk led the generators to sell at prices higher than their marginal cost, in order to account for the risk of the utilities not being able to pay. This is when the Electricity Crisis became a financial crisis, putting the utilities on the verge of bankruptcy. The utilities could have removed the price cap by forgoing the collection of their stranded costs and this is what San Diego Gas and Electric (SDG&E) did. SDG&E removed the price cap once they recouped their stranded costs, allowing them to increase retail prices, but the state quickly reinstated the price cap for SDG&E. The state eventually bailed the utilities out by purchasing electricity on their behalf, leaving the retail price caps in place, and causing the state to take the losses previously incurred by the IOUs. With one buyer of electricity, the state, the competitive nature of the wholesale market was greatly diminished. The state issued bonds to cover the losses, resulting in California residents paying the price difference through these bonds. In time, the state increased the retail price caps, allowing wholesale prices to be passed on through higher retail prices. These

higher retail prices were to remain in place long enough for the IOUs to pay off their debts and allow them to begin purchasing electricity again. By this time, wholesale prices had decreased back to pre-Crisis levels. It has been noted that had the rate increases occurred in the early stages of the Crisis, many of the Crisis's problems could have been avoided (Sweeney 2002).

Solutions That Could Have Reduced the Effects of the Crisis

According to Borenstein (2002), two changes in the form of long-term contracts and real-time retail prices could have greatly reduced the effects of the Crisis. Borenstein (2002, p. 203) states,

Combining long-term contracts with real-time pricing can provide the right economic incentives to reduce demand at peak times when the system is strained, while still assuring customers of relatively stable monthly bills.

In addition to contracting and real-time pricing, Borenstein, Bushnell and Stoft (2000) show that increasing transmission capacity between certain generators can significantly increase competition. While some characteristics of the Crisis period were unavoidable, these changes would have greatly reduced the negative effects.

The role of long-term contracts discussed by Borenstein (2002) could reduce the volatility associated with the spot market. Contracts on average will not lead to lower prices than the spot market, but may lead to lower price risk. California municipalities were allowed to enter into these contracts which allowed them to weather the Crisis period more easily than the IOUs (Sweeney 2002). Once the Crisis period began, it was too late for the IOUs to enter into long term contracts because these contracts would be at the higher wholesale prices. This is exactly what the state of California did after they began purchasing electricity on behalf of the IOUs (Sweeney 2002). These contracts could also reduce the incentive of generators to exercise market power (Borenstein, 2002). When a generator sells a larger amount of its potential supply through contracts, it will have less pronounced incentives to increase the price on the remaining supply sold through the spot market. The incentive for the IOUs to enter into these contracts could be

reduced due to regulators deeming the contracts “imprudent” and not allowing the IOUs to pass these costs through to customers.

As discussed by Borenstein (2007), during the Crisis California had in place a combination of flat and time-of-use (TOU) retail rates. Residential customers faced a flat rate, while industrial and commercial consumers used TOU rates. Once these flat rates are set they are invariant to the supply and demand conditions of the wholesale market. TOU rates are also set for a given period, but different rates are charged depending on the time of day and year. Most industrial and commercial customers in California use a rate structure with five rates that vary depending on time of day, week, and year. Real-time pricing (RTP) is the proposed solution to retail demand being largely invariant to price fluctuations. Borenstein (2002, p. 203) states,

Deregulating only the supply side of the market seems to be the equivalent of making an electricity market operate with one arm tied behind its back.

As discussed by Borenstein (2002), RTP would allow customers to reduce consumption when supply became limiting and retail prices were high. Customers would be billed at the hourly market rate, or whatever rate was used, when the electricity was consumed. RTP would allow consumers to know current market conditions and adjust consumption accordingly. Depending on the price elasticity of demand, price spikes associated with the Crisis period would have been significantly reduced. Customers would reduce and/or shift consumption to lower cost times when electricity generation neared capacity and prices were the highest. Real-time pricing would reduce the need for excess capacity because as the system neared capacity prices would increase resulting in a reduction in demand. Most of the electricity supplied as the system nears capacity is produced through the use of so-called peaking plants. These peaking plants largely lay idle, only running when the system nears capacity. Borenstien and Holland (2005) found that increasing the share of customers using RTP would decrease capacity requirements and increase economic welfare. Their simulations revealed that if a third of customers used RTP, then peaking capacity needs would be reduced by 44%. They also found that time-invariant pricing does not achieve the first-best or even the second-best optimum. In addition to a reduction in the need for excess capacity, real-time pricing would

decrease the incentive a generator has to exercise market power. The incentive to exercise market power would decrease due to a price increase resulting in a greater reduction in demand. While RTP has a significant potential upside, a few obstacles stand in the way of its implementation. These include political barriers due to wealth transfers between customers, the cost of installing a real-time metering system, and a more complex billing structure (Borenstein, 2007).

Borenstein (2007) finds that significant wealth transfers would occur with the implementation of RTP in California. Under the current flat rate and TOU setup, customers that consume disproportionately greater quantities during high price periods are subsidized by those that consume smaller quantities during these times. Borenstein's analysis used a sample of 1,142 industrial and commercial customers located in northern California. Between 40 and 70 percent of the wealth transfer is accounted for by moving from a flat rate to the TOU setup, which the customers in this model have already transitioned through. A smaller wealth transfer would occur when moving from the current TOU setup to RTP. Borenstein notes that TOU rates generally understate the peak price and overstate the off-peak price, resulting in a greater wealth transfer when going from TOU to RTP. Moving to a RTP setup would likely be met with resistance by the customers made worse off. To account for this opposition Georgia Power has a "two-part" system that allows customers to procure a baseline quantity at regulated TOU prices, and purchase the remaining requirement at real-time prices. This design could still maintain the cross-subsidy, likely making it more acceptable than pure RTP.

As shown by Borenstein, Bushnell and Stoft (2000), strategically increasing transmission capacity between certain points on the grid can greatly increase competition. In a free market environment, a firm may have an incentive to increase congestion on a transmission line that provides imports into its service region. A transmission line becomes congested when a line is carrying all of the electricity it is physically able to carry. Competition can be greatly increased even if the added capacity is rarely used. The threat/option of use is all that is needed to significantly reduce the incentive a firm has to exercise market power. The credible threat of imports, from rising retail prices, into a firm's region will reduce the incentive the firm has to increase retail prices through market power. Increasing capacity between two points may actually lead to a reduction

in the use of the line. Electricity cannot be transmitted both directions down a single line, so the electricity that is transmitted between two points is the difference between what is required by the two points. This would give a line the appearance of not being heavily used, when in actuality it could be providing large amounts of effective transmission. Borenstein, Bushnell and Stoft (2000, p. 297) state,

Thus, if a connecting line is of sufficient capacity to reduce market power as much as possible, it may appear to be overbuilt and underused.

Significant welfare gains could come from increasing capacity on a line that is not heavily used. In the past, required increases in transmission capacity were based primarily on the congestion level of the line, this method may not be valid in a competitive electricity market. This goes to show that the market's overseers will have to change their operating procedures, when going from a regulated market to a competitive one.

CHAPTER 4 - Conclusions

The electricity industry has gone through many changes in a relatively short period of time and I believe that this will continue into the future. While there are many risks associated with moving towards a more competitive market structure, they appear to be outweighed by the potential efficiency gains through increased competition. Regulation will always be necessary over the natural monopoly segments (i.e. transmission and distribution), but the other segments (i.e. generation and retail) have the potential to operate competitively.

Performance based incentives provide a necessary transition from traditional regulation to competitive markets. PBR helps to change the way managers who operate traditionally regulated firms behave, encouraging more efficient behavior. This is a necessary change for a firm to move towards a free market design. California could have better made the change to free markets through increased PBR. Perhaps California's problems would have become apparent through this transition before they caused the pronounced economic upheavals that they did.

California taught everyone a few things not to do when it comes to restructuring their electricity sector. But they also showed us it is possible, after all the restructured setup did function well for nearly two years before the events of the Crisis period took place. California's problems pointed us in new directions for future research. If the findings of Borenstein, Bushnell and Wolak (2002) are correct, then reductions in the extent of firms exercising market power are necessary. With electricity markets, more than just firm size goes into determining a firm's potential to exercise market power. Other factors such as demand curve elasticity and network constraints can significantly affect a firm's ability to exercise power. No longer can network constraints be solely determined by congestion. As Borenstein, Bushnell and Stoft (2000) have shown, competition in a network can be increased even if network use in an area is reduced. Real-time pricing could potentially increase the effective demand elasticity. The potential benefits of RTP can no longer be ignored. The implementation of this technology is necessary to fully achieve the benefits of a competitive market.

Finally, in order for a market design to be successful, it must be able to change. It should be looked upon as a dynamic system that should not be changed based on short term fixes, but rather on long term goals. Though the California experiment ended poorly, it should not end the pursuit of a competitive electricity sector, nor should it suggest that improvement in the current structure is not possible. The key is to learn from the mistakes in California to improve market design going forward—to practice what the great inventor Charles Kettering referred to as “intelligent failure.”

References

- Borenstein, S. 2007. "Wealth Transfers Among Large Customers from Implementing Real-Time Retail Electricity Pricing." *The Energy Journal* 28(2):131-149.
- . 2002. "The Trouble With Electricity Markets: Understanding California's Restructuring Disaster." *Journal of Economic Perspectives* 16(1):191-211.
- Borenstein, S., and J. Bushnell. 1999. "An Empirical Analysis of the Potential for Market Power in California's Electricity Industry." *The Journal of Industrial Economics* 47(3):285-323.
- . 2000. "Electricity Restructuring: Deregulation or Reregulation." *Regulation* 23(2):46-52.
- Borenstein, S., J.B. Bushnell, and F.A. Wolak. 2002. "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market." *The American Economic Review* 92(5):1376-1405.
- Borenstein, S., J. Bushnell, and S. Stoft. 2000. "The competitive effects of transmission capacity in a deregulated electricity industry." *RAND Journal of Economics* 31(2):294-325.
- Borenstein, S. and S. Holland. 2005. "On the efficiency of competitive electricity markets with time-invariant retail prices." *RAND Journal of Economics* 36(3):469-493.
- Cave, M.E., S.K. Majumdar, and I. Vogelsang. 2002. *Handbook of Telecommunications Economics*, vol. 1. The Netherlands: Elsevier.
- Comnes, G. A., S. Stoft, N. Greene, and L.J. Hill. 1995. "Performance-Based Ratemaking for Electric Utilities: Review of Plans and Analysis of Economic and Resource-Planning Issues." Berkeley, CA:Lawrence Berkeley Laboratory. LBL-37577. November.
- Joskow, P.L. 1996. "Introducing Competition into Regulated Network Industries: from Hierarchies to Markets in Electricity." *Industrial and Corporate Change* 5(2):341-382.
- . 1997. "Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector." *Journal of Economic Perspectives* 11(3):119-138.

- . 1972. “The Determination of the Allowed Rate of Return in a Formal Regulatory Hearing.” *The Bell Journal of Economics and Management Science* 3(2):632-644.
- Joskow, P.L., D.R. Bohi, and F.M. Gollop. 1989. “Regulatory Failure, Regulatory Reform, and Structural Change in the Electrical Power Industry.” *Brookings Papers on Economic Activity: Microeconomics* 1989:125-208.
- Joskow, P.L., and E. Kahn. 2001. “A Quantitative Analysis of Pricing Behavior in California’s Wholesale Electricity Market During Summer 2000.” Working paper 8157, National Bureau of Economic Research, Cambridge, MA.
- Joskow, P.L., and R. Schmalensee. 1987. “The Performance of Coal-Burning Electric Generating Units in the United States: 1960-1980.” *Journal of Applied Econometrics* 2(2):85-109.
- Kahn, A.E., P.C. Cramton, R.H. Porter, and R.D. Tabors. 2001. “Pricing in the California Power Exchange Electricity Market: Should California Switch from Uniform Pricing to Pay-as-Bid Pricing?.” Blue Ribbon Panel Report.
- Munson, R. 2005. *From Edison to Enron: the business of power and what it means for the future of electricity*. Westport, Conn.: Praeger Publishers.
- Sweeney, J.L. 2002. *The California Electricity Crisis*. Stanford, California: Hoover Institution Press.
- Warkentin-Glenn, D. 2006. *Electric Power Industry in Nontechnical Language*. Tulsa, Okla.: PennWell Corp.
- Weisman, D.L. 2002. “Is there ‘Hope’ for price cap regulation?” *Information Economics and Policy* 14(2002):349-370.
- Weisman, D.L. and J.P. Pfeifenberger. 2003. “Efficiency as a Discovery Process: Why Enhanced Incentives Outperform Regulatory Mandates.” *The Electricity Journal* 16(1):55-62.
- Winston, C. 1998. “U.S. Industry Adjustment to Economic Deregulation.” *Journal of Economic Perspectives* 12(3):89-110.
- Wolak, F.A. 2003. “Measuring Unilateral Market Power in Wholesale Electricity Markets: The California Market, 1998-2000.” *The American Economic Review* 93(2):425-430.

Wolfram, C. 2003. "The Efficiency of Electricity Generation in the U.S. After Restructuring." Working paper, Center for the Study of Energy Markets, University of California, Berkeley. CSEM WP 111R.