Of Megawatts and Men

Understanding the Causes of the California Power Crisis

Christopher F. Thornberg
UCLA Anderson Forecast

"Getting it done fast and in a way that pandered to the many interests involved became more important than getting it right."

Paul L. Joskow, "California's Electricity Crisis", National Bureau of Economic Research Working Paper 8442, August 2001, p. 11.

"If I wanted to raise rates, I could solve this problem in 20 minutes."

Governor Gray Davis, quoted in National Journal, the Hotline, March 8, 2001.[1]

The good news about California's power situation is that the widespread blackouts and extremely high wholesale and retail prices predicted for the summer of 2001 never materialized. Good weather, new power plants and new long term contracts have all helped avert another round of blackouts and debt. Yet while the crisis has become old hat in the press, it is not yet over. The bad news is that California continues to feel the hangover of the problems that emerged in the middle of 2000 and lasted into 2001. The state is saddled with buying power at what now appears to be relatively expensive wholesale rates under the long term contracts that were negotiated at the height of the crisis. The crisis quickly gobbled up the state's cash reserves that it had accumulated in the late 1990's. And the state continues at this writing a massive debt that is complicating its financial condition.

Maybe the worst casualty is the process of deregulation itself. It has been incorrectly blamed for the problems that have occurred and, as a result, the process of moving the utility industry to private enterprise has been indefinitely, and possibly permanently, delayed. This is not just true in California but in many parts of the United States where politicians are carefully considering the political risk that now appears to be attached to such efforts. Governmental oversight—whose inherent inefficiency was the initial cause for the push for deregulation—again determines the functioning of the industry. And electricity prices in California are higher now than they were prior to when the deregulation plan was initially put into place.

What went wrong? How could such good intentions lead to such dramatically bad results for the state? The short answer is that California's energy plan—while a move towards the goal of complete privatization—was really just a form of re-regulation of the industry during the short time it was in place. The retail markets remained essentially unchanged, with fixed prices and little real consumer choice as to suppliers. The primary free market feature of the plan that went into place in 1996 was that wholesale prices were allowed to float. Yet how the major utilities conducted their business still remained largely under regulatory control rather than under the competitive pressures that are generated within a truly privatized market. The wholesale market was created by mandate

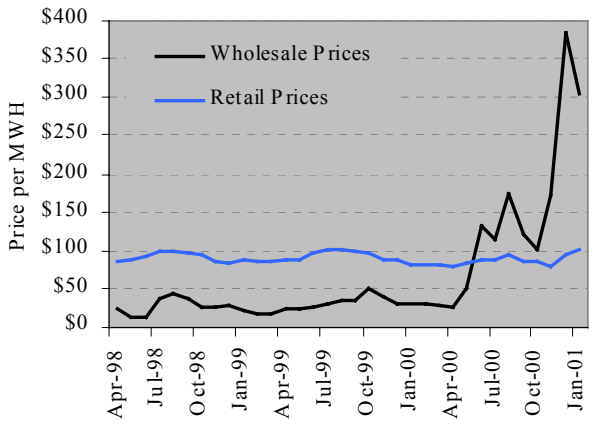
[1] Editor's note: This quote has appeared in various online and newspaper editorials, sometimes as "raise rates" and sometimes as "raise prices." Its exact origin is unclear.

under a scheme that required, among other things, the public utilities to sell off a portion of their productive capacity and then to purchase their power needs in a day-ahead spot market. Partial deregulation is often worse than no deregulation, however, given the distortions it creates in the incentives and actions of firms and individuals. As with many past cases of partial deregulation, these moves lead to an expensive breakdown in the market when these distortions are revealed through some external influence.[2]

The problems started with the exclusive use of a day-ahead spot market for power. Short-term spot transactions in a capital intensive markets ultimately lead to boom-bust cycles marked by periods of excess capacity and low prices countered by periods where capacity is constrained and prices become very high. Additionally, the short run spot market in California---combined with the lack of deregulation in neighboring states---had the effect of transferring all market risk to California. Lack of retail deregulation that kept prices for final consumers fixed then transferred all this risk to the three major public utilities. The utilities were not allowed to sign long term contracts with wholesalers, nor use financial instruments to hedge this risk.

Figure 1: California Wholesale Electricity Prices

Monthly averages



Source: CALPX

With these fundamental flaws in the deregulation plan, the question should have been 'when' will a crisis occur rather than 'if'. The market breakdown was caused by external circumstances that occurred during a capacity constrained period in the market. The external influences were a rapid increase in natural gas prices, a national power producing capacity shortage, and bad weather in both the Northwest and in California. The lack of a quick reaction to the crisis by the Federal Energy Regulatory Commission (FERC) or the California Public Utilities Commission (CPUC) then intensified an already difficult situation.

The results of this failure are well documented even if their roots are not widely understood. Wholesale prices in the day-ahead markets-- where the utilities were required to purchase their "net-short" power needs-- began to spike sharply in mid-2000.[3] Average

[2] There are a number of such cases. One of the most famous is the Savings and Loan scandal of the late 1980's. There the regulations that governed the investment behavior of the Savings and Loans were relaxed, but investors in the banks (the depositors) were still given a guarantee on their deposit by the federal government. The combination led banks to make overly risky investments in order to attract more depositors with high interest rates. Depositors, insured by the government for free, had no incentive to adequately monitor the behavior of the banks.

[3] "Net-short" refers to power demand the utilities could not generate themselves or obatin for long term contracts.

prices in December 2000 were over ten times the price two years previously. (Figure 1). During this time retail rates stayed fixed, and the losses incurred were enough to push two major utilities into financial ruin, consume the substantial financial reserves the state had built up, and still leave a substantial debt burden that is currently the focus of intense debate in Sacramento even as this paper is being written. By February 2001, the wholesale market, the California Power Exchange, was closed, and the state had become the primary wholesale purchaser of electricity for the state. To add insult to injury, California also suffered a number of widespread power blackouts. While the overall economic impact of these outages was small, nonetheless the public relations damage to the state's reputation was real.

What should California and the US learn from what happened? This paper is a retrospective on the crisis, detailing the various stages leading up to the failure of the wholesale markets, as well the factors which figured into the crisis and the cost of the crisis. First, we discuss the reasons for deregulation; then the California situation is detailed including the primary causes for failure in the system. Finally, lessons from California's experience are presented.

The Roots of Regulation and Deregulation

The Merriam-Webster dictionary defines capitalism as "an economic system characterized by private or corporate ownership of capital goods, by investments that are determined by private decision, and by prices, production, and the distribution of goods that are determined mainly by competition in a free market."[4] Most economists will also tell you that private ownership and private markets are the most socially efficient mechanism for allocating scarce resources because they push resources to their most valuable use. Markets do this through the profit mechanism. Firms or individuals that can create the greatest profit through the most efficient use of scarce resources to produce goods can 'win' the right to use these resources through the bidding process. Less efficient firms will lose the bidding war and disappear. The competition-profit system is constantly pressuring firms to innovate in ways that make more with less to ensure profits and therefore survival. In many ways, it is this competitive pressure to become individually more efficient in the use of resources that makes the overall economy advance and offers a better standard of living to everyone.[5]

Most economists will also tell you, however, that the private market system doesn't always perform optimally. There are times when the market fails to create the socially efficient outcome due to a variety of what are termed "market failures". Markets tend to operate best when there are multiple competitors and buyers, when information is easily obtainable, and when the agents in the market fully internalize all the costs or benefits of their actions. When these conditions fail to be present, it may be optimal to have some central authority, such as a government, step in to regulate the rules that govern the private market to make it operate in a more socially efficiently manner.[6]

---

[4] The Merriam-Webster Collegiate Dictionary, on-line version.

[5] It is this process that Adam Smith described in his discussion of the "invisible hand" of markets over 200 years ago. The central point of his argument is that each individual working to maximize their personal well being actually produces the optimal social outcome. In other words personal greed leads to the common good; a counterintuitive point that is so often misunderstood.

[6] For example, pollution market failures occur because the polluting agent is not fully paying for the use of the

Under some extreme circumstances the government may need to take over the provision of the service completely, as in the case of the national defense.

One reason that markets may fail is the existence of large economies of scale in the production process—the so-called 'natural monopoly' problem.[7] In this situation the market is, on the basis of cost, most efficiently served by a single firm.[8] Monopoly, however, violates the need to have multiple competitors to make a market efficient. If there is just a single firm supplying the good, it will act monopolistically and restrict quantity produced in order to push prices up above their competitive levels. In other words the market is cost efficient but not price efficient. When such situations occur, the government may step in to regulate the prices charged consumers. Presumably, the price charged will reflect what might be the price if the market were private and competitive.

When the market is best served by one firm, the government also typically restricts market entry by new firms. While the price controls suffice as a short run solution to the natural monopoly problem, the lack of entry in government operated or regulated monopolistic industries tends to cause inefficiency in the long run. These firms do not suffer the pressures of competition that typically push firms towards finding greater efficiencies. If a private firm makes a bad business decision, it may be faced with the threat of bankruptcy, and it is the firm's owners and managers that face the consequences. If a government regulated monopolistic firm makes a bad business decision, then consumers pay through higher prices or taxpayers pay through higher taxes. Thus, the decision of whether to regulate or deregulate an industry becomes a function of choosing the lesser of two costs; those that may be caused by an inefficient private market or those created by the lack of competition that comes with government regulatory oversight.[9]

In the early days of power, two important features of the industry—large economies of scale in the production and distribution of electricity, and the inability to store electricity—led to the creation of government regulated power monopolies. Early plants were more efficient as they became larger, and power losses in distribution were costly over distances, implying that one power plant could most efficiently serve a specific geographical region. Similarly, having multiple parallel distribution systems would have clearly been wasteful. Also important was the issue that electricity is not a storable good. This feature implies that at any given point of time the amount of power being generated by production plants and fed into the distribution network must be approximately equal to the amount of electricity being used by consumers on the other end. If not, the entire system will break down.

---

environment as a place to dispose of unwanted byproducts of the production process. Examples of the regulations to control these problems include the use of emission control laws on autos and factories, pollution taxes, zoning laws, etc.

[7] Economies of scale occur when each additional unit lowers the average cost for all units. Under these circumstances one large firm can produce products more efficiently than 2 small firms.

[8] A natural monopoly is technically characterized as a market where the average cost curve for a single firm is declining as it crosses the demand curve. In other words, the average cost of production is lowest when just one firm is producing.

[9] One famous case of deregulation is the AT&T monopoly on long distance telephone calls that was removed in the early 1980's. Since that time, long distance telephone rates have dropped substantially, largely due to a large amount of investment in new infrastructure and technologies. There are many forms of inefficiency created by the lack of competition. Another is the use of too much capital relative to labor due to the guaranteed returns afforded to capital investments under a regulated monopoly situation.

That defining feature of the market creates the need for a high degree of coordination and communication between power producers and between the production and distribution of power. The process initially worked best when there was common management over producing plants within a specific geographic area. While the exact government regulation model used varied widely from place to place, the typical structure was that the rates paid by the various consumers were determined by independent, government-run commissions. In return utilities were licensed to be exclusive operators and guaranteed some annual rate of return on their capital. In California the three big public utilities that formed were Pacific Gas and Electric, the Southern California Edison Company, and the San Diego Gas and Electric Company. These firms worked under the regulatory oversight of the California Public Utilities Commission. It should be noted that there are also a number of other power providers, primarily municipal generators including most significantly the Los Angeles Department of Water and Power. Because there were government-owned utilities, deregulation did not reach into every corner of the state's power market.

Over the past few decades, a number of important technological advances have diminished the need for regulated monopoly in the power industry. High power transmission lines reduce the losses associated with the movement of power across long distances. Smaller power plants are now as efficient or more efficient-- on a per unit basis as large power plants. Advances in communication technology have made it possible to coordinate production more easily across multiple power plants and also vertically disconnect ownership of the power lines from ownership of the actual production facilities. As a result, there now is the possibility of having competition in the power production industry even while maintaining the overall integrity of the system.

Change was slow to come, however. While consumers are happy with industry competition for their purchases of cars, food, and the myriad of other services and goods, they have grown accustomed to state regulated utilities and view deregulation in this area with some suspicion.[10] Regulatory commissions are often unwilling to hand their powers over to the private sector. And, of course, the utilities themselves have developed a corporate culture that revolves around the concept of a guaranteed rate of return and lobby hard to maintain this benefit. It took a number of widely publicized commercial errors on the part of local public utilities and an overall atmosphere of deregulation in the 1980's to lead the Federal Energy Regulatory Commission into passing a series of rules allowing competition at the wholesale level.[11] This change allowed for the process of deregulation to begin to take place within individual states. Yet twenty years later little had actually occurred, except for a number of experiments, the most noteworthy being in Texas, Pennsylvania, and--of course--California.

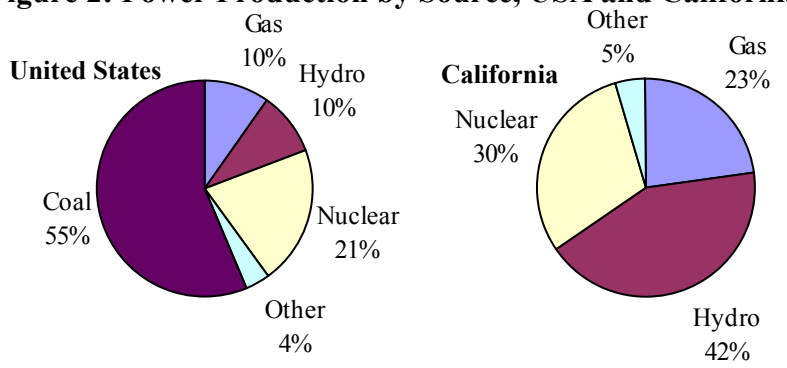
California's move to deregulation began in the early 1990's. Industrial users of electricity were the driving force behind the lobbying for price relief, with the primary

---

[10] It should be noted that in part this is due to the fact that the utilities have been relatively successful in their goal of supplying electricity, with only little of the flagrant abuses seen in other previously regulated sectors such as trucking.

[11] While local regulatory agencies controlled the retail and distribution side of the industry, FERC is charged with the regulation of wholesale production. This is likely due to the existence of many large multi-state power production plants that necessitated the need for federal regulation. It was this multi-tiered level of regulatory control that lead to the complex political maneuvering during the crisis days. The CPUC and the state were asking the FERC for regulatory relief at the wholesale level, even while FERC was looking to the state to fix the problems at the retail level.

Figure 2: Power Production by Source, USA and California, 1998



Source: Energy Information Agency

complaint being that they were paying significantly higher prices than industrial users in neighboring western states.[12] These high prices were in turn blamed on inefficient regulatory bodies, costly investments in expensive nuclear power plants, investments in 'unnecessary' excess capacity, and expensive long term contracts by the regulated utilities with independent suppliers. The complaints had various degrees of validity. Less discussed but also playing an important role-- is the issue that other states use more coal or hydro power that significantly lowered the average price of power as seen in Figure 2. California had embarked upon a plan to encourage investment in expensive--but environmentally cleaner-- co-generation gas facilities, renewable energy and nuclear power plants. It is one of the few states with no major coal consuming plants.

Whatever the motivations, the CPUC finally responded to this political pressure by publishing the first overall plan for deregulation in 1994. It was based roughly around the model used for utility deregulation in Wales and Britain, although the end result looked completely different. The primary set of ideals around which deregulation was outlined included the following points:

- Open access to power transmission systems by competing producers, with oversight provided by an independent regulator body (the CALISO or California Independent Systems Operator),
- consumer choice of the retail provider of electricity, and deregulated prices,
- creation of a wholesale spot market for electricity (CALPX or California Power Exchange Corp.) to create competition at the wholesale level,
- divestiture of generation plants by the three big state utilities in order to create the necessary liquidity in the wholesale markets and to rid the market of potentially problematic vertical integration.

Essentially the long run plan was to move from the traditional two-tiered regulated market (suppliers and consumers) to a three-tier competitive market (wholesale suppliers, retail distributors, and consumers).

The final restructuring report was released in 1996 and was approved

[12] California industry was paying an average of 7.25 cents per kilowatt hour in 1990, while the national average was 4.75 cents. It is worth pointing out that industrial users have always enjoyed substantially lower prices than residential and commercial users. In California the difference was about 2 cents.

unanimously by the state legislature that same year. Going from this initial plan to final implementation was very rushed. Although there were unresolved issues and problems, the final plan was put into place less than two years later in early 1998. Yet for all the potential promise of this plan, it was pointed out by at least one expert that "getting it done fast and in a way that pandered to the many interests involved became more important than getting it right. The end result was the most complicated set of wholesale electricity market institutions ever created".[13]

A Strange Sort of Re-regulation

Although the motivations behind the deregulation plan were surely worthy, the initial results bore little resemblance to the ideals of market based competition. California's energy plan—while a move towards privatization—was simply a form of re-regulation of the industry during the short time it was in place. This outcome was due in part to the phased-in approach the state took to implementing deregulation, where only certain aspects of the market were initially deregulated, with the idea of expanding the process over time. It was also due in part to the above-cited pandering of the many special interests that were involved with the process.

One of the most important features of the plan was to split the market into distinct wholesale and retail segments.[14] Power suppliers in California had typically been vertically integrated, with a large portion of capacity directly owned by the retail distributors. California utilities also had a variety of long term contracts with public utilities in other states to provide additional power. This contracting was in part a seasonal arrangement, where California would import power from the Northwest during the hot summer and export power back during the winter months when local demand was low. However, the trades did not completely balance, and the state imported about 10% of its power needs in the average year prior to deregulation.

The existence of independent wholesale firms in the power industry is relatively new, and they still represent only a small portion of national capacity, about 12% of total national capacity in 1998. The plan in California was to create three levels of participation, the final consumers, the wholesale producers, and the utilities acting as the middlemen between the two. Initially the deregulation plan called for the public owned utilities to divest themselves of all producing capacity. What happened instead was that the three major firms only sold off their expensive fossil fuel powered plants, and held on to their hydro and nuclear facilities. Nevertheless, the utilities found themselves in a net-short position for reasons that will be presently discussed.[15] They were therefore put in the position of needing to buy a portion of their power from independent wholesalers.

The primary 'free' market feature of the plan that went into effect in 1998 was that wholesale prices were allowed to float on a day-ahead spot market. Beyond this, however, how the major utilities conducted their business still remained largely under the

---

[13] Joskow (2001), "California's Electricity Crisis," National Bureau of Economic Research Working Paper 8442, August 2001, p. 11. This report is an excellent summary of the various portions of the final deregulation plan.

[14] It is not clear as to why this goal was so important to those who were writing the deregulation plan. It might be hypothesized that the large industrial users who were behind the initial push for deregulation felt that the split would give them the ability to purchase cheaper power by negotiating directly with wholesalers. Additionally, it might have been an effort to create competition, but still shield the small residential consumer from the resultant price fluctuations.

[15] See footnote 3.

control of the CPUC rather than under the competitive pressures that are generated within a truly deregulated market. Retail prices for most ratepayers remained fixed. The utilities were required to purchase their net-short power needs through the one-day-ahead market rather than through other acquisition mechanisms such as long-term contracts. Two of these features, fixed retail rates and the required use of a day-ahead spot market, played the largest role in setting the stage for the crisis that was to begin in late 2000. They created a situation in which the three major California public utilities essentially absorbed all supply and demand side risk for the entire Western United States. These two features will each be examined in detail below. Finally, the use of futures markets or other hedges—necessary to reduce the short run risks inherent in such markets as well as to provide the market with long term signals of market needs—were actively discouraged.

Fixed Retail Rates

One of the major debates that arose during the negotiations over the deregulation plan had to do with the issue of stranded costs. The California independent utilities feared that they would be left at a competitive disadvantage after deregulation was complete—compared to incoming retail and wholesale competitors—if they were saddled with the high average costs of the expensive power production plants in which they had invested prior to deregulation.[16] The merits of these claims are debatable. Moreover, these bad investment decisions were not the fault of the California consumer. Nonetheless, the deregulation allowed the utilities to receive some form of compensation for their sunk or "stranded" costs.

To generate the necessary profits to pay off the stranded costs, the deregulation plan called for initially maintaining fixed retail rates while allowing just the wholesale rates to float. Regulators apparently assumed that wholesale rates would almost certainly go down, creating the necessary excess profits to pay off stranded costs. They were even confident enough to lower retail rates at the outset of the deregulation plan to put rates back to where they were in 1996. Once the stranded costs were recovered, retail rates would then be allowed to float, presumably also downward. It was simply never considered that various forces could cause wholesale prices instead to rise to levels above these fixed retail rates. In part, this deficiency in planning may have been due to the fact that if wholesale prices went up, then deregulation might be seen as a failure. The financial risks of such a failure were clearly not considered.

The critical impact of fixed retail rates was that demand for power at the wholesale level became perfectly inelastic.[17] In California the suppliers in the retail markets (the public utilities) were the primary consumers on the wholesale markets and legally obligated to meet all retail power demand at the prices set by the CPUC. When retail demand went up, retail prices did not. Hence, this increase in demand was passed on fully to the wholesale markets regardless of the prices being charged there. Similarly,

---

[16] The way to think about these sunk costs is to imagine that the utilities had indeed sold off all their producing capacity. The claim being made was that the market value of these assets in a deregulated market was lower than what the utilities had paid for them. Thus, the utilities would have ended up carrying a substantial debt burden that would have prevented them from competing effectively in the new market.

[17] In a standard market when demand increases or supply decreases (costs increase), prices are typically driven up. This rise has two effects on the market. First, the higher prices encourage greater supply. Second, higher prices serve to remove marginal consumption, reducing the total quantity consumed. This latter rationing effect doesn't work if prices are fixed, however. The result can be a product shortage. A famous example of such shortages was experienced during the oil crises in the 1970's when gasoline prices were fixed by the federal government.

when the cost of producing power began to rise, consumers had no incentive to conserve, since they bore none of the burden of these higher costs. Under these circumstances, the costs or benefits of any change in the supply or the demand for power is completely absorbed by the middlemen (utilities) in the market, as opposed to being spread over all three market participants.

This issue might not have been a problem if the utilities had not been pushed into the position of needing to purchase power on the wholesale markets. At the same time that the wholesale market plan went into effect, retail consumers were also given the choice of retail power provider. Those who chose not to switch to some new provider would remain with their existing public utilities as the default provider. But this portion of the plan was undermined completely by the stranded cost provisions. New retail providers were necessarily constrained in the prices they could charge retail buyers, since they were required to pay a portion of the stranded costs as well. As such, there was little incentive to enter the retail market. And there was little incentive for retail buyers to switch even if there had been much entry, since there was no price competition mechanism to encourage them to select a new provider. In the end only 3% of retail consumers switched to new retailers, representing roughly 12% of the consumer base.[18]

The initial deregulation plan had called for the public owned utilities to divest themselves of all producing capacity. It must have been clear to the utilities that they would be facing a significant risk if they had indeed been completely reliant on these wholesale markets to buy the power necessary to meet demand, given the fixed retail rates. The three major firms ended up selling off only their expensive fossil fuel powered plants, and held on to their hydro and nuclear facilities, presuming to mitigate the risk. Nevertheless, due to the lack of retail competition these firms as the default providers suddenly found themselves in a net-short position, where a portion of the state's power needs had to be purchased from independent wholesalers in the day ahead market. This shortage—combined with the lack of pass-through of wholesale rates to retail prices—put the public utilities in an extremely vulnerable position.

Day-Ahead Spot Markets

The second major feature of the deregulation plan that led to the market meltdown was the requirement placed on the utilities that they purchase their electricity on a day-ahead spot market, the California Power Exchange (CALPX).[19] The overall system was run by the CALISO (the California Independent System Operator) who took over the actually management of the power grid. The basic formation of the CALPX was relatively simple. Wholesalers would put in bids for the power that they had available for a particular hour of the following day. The utilities would put in their bid for how much power they needed to cover their position for the following day. Utilities were required to also 'bid' their own power production out. The market simply crossed the demand and supply lines, and the price that represented the marginal unit sold was applied to all

[18] Joskow, op. cit., pg. 25. These numbers reflect the fact that the only major switchers were large industrial consumers, not surprising given that they were the primary force behind the move towards deregulation.

[19] Spot markets for power have been in existence for some time. Traditionally, most power that is supplied to consumers by the distributing utilities is produced either through direct ownership by the utility or under long term contract with an independent wholesaler. Forecasts for power consumption always have some degree of error and long term contracts cannot fully account for all these day to day and hour to hour variations in demand. Spot markets have typically played the role of load balance, where utilities may purchase the extra power necessary to meet demand, or may sell off excess capacity for other uses.

power sold in that hour.[20] This was essentially a twenty-four hour forward market since those who won the bids to supply power didn't necessarily have to be the owners of the plants that actually produced the power. They simply took on the financial obligation of making that power available.

The first problem with relying on a day-ahead wholesale market stems from the capital intensive nature of the production process. Power plants take many years to build in the best of times. Shutdowns and startups of these plants are costly and time consuming and fixed overhead costs are also expensive. Such a market can be thought to be in one of two states during operation; with excess capacity or with capacity constrained. The net result of such market structure is that a boom-bust cycle is created.[21] Such a market will be characterized by periods of extremely low prices when the market has excess capacity and extremely high prices when the market is capacity constrained. Indeed these boom-bust cycles were exactly what California power market experienced during its first 30 months of operation, although the 'boom' periods were relatively few and far between.

In an excess capacity competitive market (the 'bust'), short run prices are pushed down to the marginal cost of producing power. Typically this will be below the average cost given the high fixed costs of operation. In an excess capacity market there is therefore little incentive to invest in new production facilities, or even to keep existing facilities operational. Wholesalers have an incentive to take existing power plants offline for maintenance or simply to save on overhead, as was seen during the winter months of 2001.[22] Hence markets with excess capacity will find operational capacity shrinking.

The other state of the market is when it is capacity constrained (the 'boom'), i.e., when demand is running close to available supply. When the market demand approaches use of all available capacity, it runs into a portion of the supply curve that is essentially vertical. Regardless of price, no more product is available. As a result, prices tend to shoot up to high levels as buyers compete for resources that simply don't exist. The market will likely not operate competitively in this state, because individual supplying firms have a degree of pricing power. They can simply bid a higher price than would be the competitive level and the fact that market is at capacity implies that that they may still be able to sell the power. In other words, individual competing suppliers can wield market power even if they do not actually directly collude to withhold supplies from the market. The exact degree of market power and the ability to push prices up depends critically on how close the market is to capacity and on the elasticity of demand. In

---

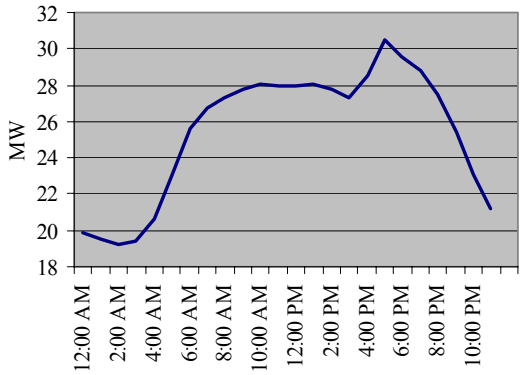
[20] The English system and that put into place in Texas differ in that they rely on bilateral contracts between individual retailers and producers rather than having one central market as in California. Along side the day-ahead markets in the CALPX were also hour-ahead markets and certain specialty reserve markets. All these different markets play the role in constantly balancing system demand with system production. There was enough of a degree of arbitrage between different markets to prevent prices from varying substantially. If a producer thought the hour-ahead price would be higher than the day-ahead price, then they had the incentive to keep their power off the day-ahead market and so on and so forth.

[21] Such boom-bust cycles have been observed in other capital intensive industries such as airlines and real estate.

[22] Many have claimed that the high amount of capacity offline during the first crisis represented a form of market manipulation by wholesalers. However, less discussed is the fact that the utilities themselves also had a significant portion of their power production facilities offline. This behavior is inconsistent with the manipulation explanation since these were the entities who had the most to lose when the market became capacity constrained. The alternative explanation is that firms rationally shut down plants because wholesale prices were simply too low to justify keeping the plant up and running.

California, fixed retail rates and the legal obligation to supply the demanded power by the utilities increased the ability of suppliers to manipulate the market even more. The utilities did not have the ability to pass costs on to consumers or 'vote with their feet' by simply not buying over-priced power.[23] Indeed most studies of the pricing in the California power markets seem to indicate that prices rose to levels that were not even remotely consistent with actual costs.

Figure 3: Hourly Power Load, California



Source: CALISO, Data for November 7, 2001

The second major flaw of the day-ahead market plan was that it created a situation whereby California absorbed all the energy risk for the entire west coast of the US. To understand this flaw, we need to understand how power markets operate. The first thing to understand is that power demand fluctuates widely over the course of a single day as well as over time.[24] Figure 3 shows how consumption in California can vary by nearly 60% over the course of a standard day, and much more during particularly hot days. Since supply must equal demand at all times on the power grid, this variation implies that power plants must be shut down and started up on an hour by hour basis. Power comes in all different 'flavors' due to the varying nature of production. Nuclear power plants, for example, have high average costs, very long start-up and shut down times, but very low marginal costs. These plants tend to be used as base-line producers, running twenty four hours per day every day. Hydro is the cheapest power of all and is easily turned on and off as need be. Gas fired plants have low average costs, but high marginal costs, and are quick to get up and running. These plants were originally designed to serve as simply as peak load plants, to be put in operation during the hours of the day when demand reached its heights.

It must be remembered that California is not a power system unto itself, but rather part of a larger system that includes eleven other states, along with a portion of Canada and Mexico. Power in these other states and areas continues to be purchased the way it was in California; through long term contract or by direct ownership by public utilities. Before the deregulation plan, California would also have had long term contracts with wholesalers or the utilities in other states. In the event of shortage or crisis situation, the

[23] It is interesting to note that while the CPUC did not give the public utilities the option of simply not buying power and thereby causing blackouts, after the state become the buyer of last resort it threatened to do just that if wholesale prices rose to too high a level.

[24] These wide fluctuations are in large part due to the fact that prices billed to consumers represent a charge based solely on total monthly consumption. The time and day of consumption is not measured or reflected in billing. Economists have been advising for many years that pricing mechanisms be put into place to allow prices for power to fluctuate over the course of the day and week. By charging higher prices during peak demand periods, the large fluctuations in demand over the course of the day and week can be smoothed out, reducing the need for excess capacity to cover peaks.

consuming utilities in the various states and areas would have had to work together to determine how the shortage could be handled.

Under the new system everything was changed. With California purchasing power in short term spot markets, all the demand in other states would be met first. Any power left over in these states would then be available to be supplied in the California wholesale markets. As a result, California found itself purchasing primarily from the most expensive power sources. Wholesalers had every incentive to use their cheapest forms of power first. The more expensive power would then be offered on the California wholesale markets. Furthermore, any shortages in other states would reduce the amount of excess capacity available to be sold on the California markets, and push California up the cost curve even farther. If power was simply not available, then it would be California that would face blackouts. California began to bear all supply side risk in the Western US power grid.[25] And of course since the public utilities were bearing the entire risk for the state, any power problems in the whole western US were bound to fall upon the three public utilities.

To make this point clearer, consider a simple example. If plywood, water, batteries and other emergency supplies suddenly increased in price by 100% in Florida as a result of an incoming hurricane, entrepreneurs from surrounding states would quickly work to take advantage of this profit opportunity by buying up available supplies in their states and exporting them to Florida as quickly as possible. However, when prices were rising in California to levels 500% to 1000% higher than in neighboring states, the same effect wasn't seen. The regulated utilities in other states were required to meet their long term contractual obligations first, despite the potential profits to be made in the California market. Consumers of power outside California did not share in shortages experienced in California.

The Inevitable Crisis

During the initial 24 months of operation, the new market system in place in California operated quite well. It appeared as if such the regulators were correct in their assumptions about the efficiency of the wholesale markets. Consumers had received a promised drop in prices when the plan went into effect, but wholesale prices were low enough that the utilities were on schedule to pay off their stranded costs on time. In San Diego the stranded costs were actually paid off early. [26] There were, not surprisingly, numerous flaws and problems in the system given how quickly it was put into place. But these were handled on a case by case basis. The boom-bust cycle was detected early, and the tendency for prices to go very high when the market became capacity constrained was noted as early as July 1998. However, these episodes were far enough apart that little was done to try and reform the fundamental flaws in the market that were being revealed.

---

[25] Much has been made of the fact that California is a net importer of power. This really makes very little difference, however. If a local power plant had a long term fixed contract with other states, it would have still been financially obligated to supply the out-of-state demand first prior to selling excess capacity to the California spot markets.

[26] San Diego's SDG&E floated retail prices in spring 2000. Hence, the consumers in that city faced massive retail price hikes when the wholesale markets began to exhibit the first sharp spikes in prices in summer 2000. It might be remarked that floating retail rates would not have mattered, since the CPUC stepped in to re-fix retail rates at their previous level in San Diego as soon as they began to rise. It is worth noting, however, that if prices had been floating across the entire state, it is doubtful that prices would have risen by as much in San Diego. The large price hikes suffered there were a function of fixed retail rates elsewhere.

The true crisis began in June 2000, as can be seen on Figure 1. After wholesale prices had been averaging under $50 per megawatt, prices suddenly began to surge upwards, with average prices peaking as high as $200 per MWH. In September and October 2000, prices began to settle back towards the lower levels seen previously. However, in November of that year prices surged again, and average prices in December were above $350 per MWH. Given that retail prices were still far below this level, it was only a manner of time before the utilities became financially strapped and unable to pay the debt they were quickly incurring. Eventually, to prevent complete financial meltdown the state finally stepped in and became the primary purchaser of power in January 2001.

There is still much misunderstanding about what actually caused the crisis in late 2000. The bad portions of the partial deregulation plan, namely fixed retail rates and the day-ahead market system, provided the fuel for crisis. California was absorbing all supply side risk for the western US, and the public utilities were in turn absorbing all the power risk within the state. This risk was realized due to a number of exogenous factors. These factors included increasing costs due to rising natural gas and NOx emission credit prices and an overall national power market that was becoming capacity constrained due to strong growth in demand. Unfortunately, the problems occurred during a time in the California market when much of the operational capacity was offline. Lack of decisive action on the part of state and federal regulators simply exacerbated an already difficult situation.

Rising Costs

The first blow to the California system was a sharp increase in the cost of producing electricity in natural gas operated plants. For reasons that are still being debated, natural gas prices began to rise in 2000. After a long period of time during which prices varied between $2 and $3 per million cubic feet, they began to rise steeply in January 2000. By the summer, prices had risen to over $4. In January 2001, prices had risen to an amazing $12 per million cubic feet.

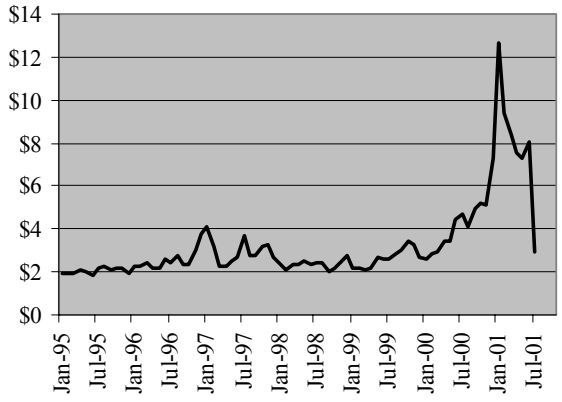
On top of this gas price hike, the twin issues of increased demand for power in California and decreased available capacity (see next section) led to a situation where many natural gas plants were being forced to operate twenty four hours per day, rather than only during peak periods as originally intended. Another market reform put in place in California was the RECLAIM pollution credit system for controlling NOx emissions from power plants and other large sources such as oil refineries. A certain number of pollution credits were to be made available to these firms each year on a declining schedule. These permits could then be traded and sold. As the natural gas plants began to operate on a more regular basis, the demand for permits naturally rose. As a result, the price of permits just about tripled during 2000. Estimates put the increase in cost on a per unit of power basis from $40 per MWH to $120 per MWH.[27]

Natural gas plants are relatively expensive to operate even when prices are low. Hence, they tend to be used as the variable capacity in a power system. The increase in input costs raised prices even higher. A good estimate is that the marginal cost of producing power with natural gas powered plants rose by a factor of 4 to 5. Even though natural gas represents only a portion of power production in California, the rules that govern the market are such that the price paid to all wholesalers was that generated by the

[27] Joskow, op. cit., p. 30.

marginal unit of production. This increase in cost was translated into higher prices for *all* power purchased on the wholesale markets.[28]

Figure 4: Natural Gas Prices, California
City Gate Prices, $/Mcf



Source: Energy Information Agency

Capacity Constraints

While rising costs certainly played a role in the power crisis, it is still quite clear that prices on the wholesale market rose to levels significantly above even these high marginal costs, indicating that the market was no longer operating competitively.[29] As discussed above, a capital intensive short-run market can begin to operate inefficiently as demand comes close to capacity. During the late 1990's, capacity was not growing as fast as demand for power both in California and the nation as a whole and this trend played a substantial role in the breakdown of the California system.

One of the major misconceptions about the power crisis is that state regulators had shut down investments in new power plants in California during the 1990's due to excessive regulatory action. While it remains true that obtaining the necessary regulatory permissions to open a new plant can be difficult and time consuming, the barriers are not insurmountable. The true issue was more one of timing and regulatory uncertainty.

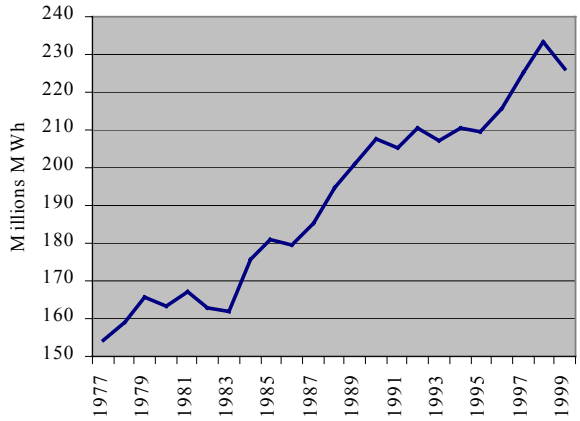
A tremendous amount was invested in new generation in the 1980's, much of it in natural gas fired co-generation facilities. Natural gas powered plants are expensive relative to hydro or coal, and the abundance of cheap power from other sources outside the state made many of these plants essentially redundant. The cost of this stock of excessive power plants was one reason that the deregulation plan was pushed in the first place. Moreover, as seen in Figure 5, the recession in the early 1990's affected the state's economy so drastically, that demand for power in California flattened. The flat demand reduced the short run need for additional power plants.

[28] It is interesting to speculate what might have happened under these circumstances had the utilities indeed sold off all their capacity. The debt burden carried by the utilities and indeed by the state might have been much higher than it actually is.

[29] See Joskow and Kahn, 2001 and Hildebrandt, 2001. Full citations may be found in the bibliography.

Starting in 1995 with the economic recovery of the state, demand for power in California began to rise dramatically. The primary barrier to investment was regulatory uncertainty. The public utilities—that had traditionally had the role of determining capacity needs—were told to reduce their producing capabilities. New wholesalers were understandably reluctant to enter into a market that had yet to have a defined form. Thus, it really comes as no surprise that capacity in the state actually decreased by a few percent between 1988 and 1998. When in 1998 the regulatory situation was settled, a number of new power plants immediately went into the planning process. One reason that summer 2001 did not experience the predicted blackouts was because a number of these plants had gone online. But the net result of the earlier slowdown in investment was that imports of power from other states by California went from 10% of demand in 1988 to 16% in 1998.

Figure 5: California Power Consumption



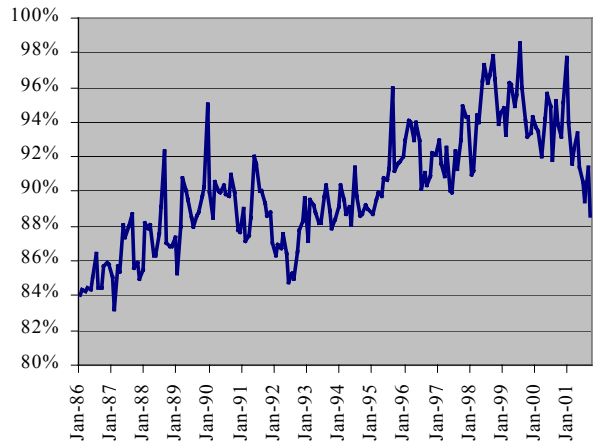
Source: Energy Information Agency

The trends in power production and consumption in California were not unique. Nationally growth in demand was outstripping growth in available capacity. Utilization rates, after hitting a low of 85% in 1992, began to move steadily upwards. As can be seen in Figure 6, by 1998, utilization rates began to hit all time highs, peaking in some months above 98%. This high utilization meant that regular production was coming increasingly from plants meant to be used primarily for peak load periods. The situation in the western US can be seen clearly in Table 1. Between 1988 and 1998, capacity in California fell by 5% while demand grew by 18%. In the other 10 states capacity grew by over 5% but demand grew by 27%! Proportionally the relative difference is roughly the same.

Table 1: Capacity and Demand, Western US

	1988	1998	Change
Electricity Production Capacity (MW)			
California	55134	52349	-5.2%
Balance from West	86798	91682	5.5%
Annual Consumption (millions of MWH)			
California	195	233	17.8%
Balance from West	269	352	27.0%

Figure 6: Capacity Utilization, Electric Utilities



Source: Federal Reserve Board of Governors

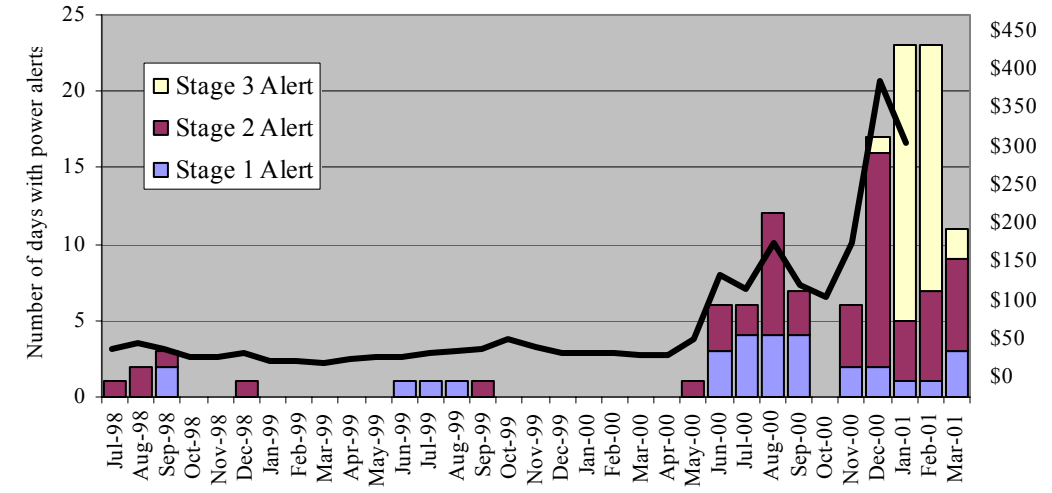
California had the misfortune of putting its deregulation plan into effect in a time period when it was purchasing a larger amount of power imports from other states that were finding it difficult to fulfill their own needs. As noted, California's day-ahead market system put it in a position of only being able to buy residual power. This residual in turn was purchased from the most expensive forms of power, in particular gas powered plants that were originally meant to be used only during peak periods. California, by being the buyer of last resort, was purchasing more and more of its power from the most expensive sources.

The capacity constraint became particularly severe in summer 2000 due to weather conditions. Heat waves moved across most of the Western US, creating large increases in demand for electricity. At the same time, a drought in the Pacific Northwest reduced availability of hydro-electric power for export to California. Since California was the buyer of last resort, the wholesale markets came closer to being capacity constrained than ever before. The CALISO issues power alerts when capacity reserves fall below five percent of current consumption. The number of days per month of power alerts is shown in Figure 7. Overlaid on this chart are the wholesale prices seen on the CALPX. The relationship between wholesale prices and capacity issues can be clearly seen.

More interesting was what happened at the end of 2000. Winter is typically a low demand period in California, which should have created excess capacity. Yet the state again ran up against such severe barriers that in January 2001, despite the extremely high prices being charged on the wholesale market, system operators could not meet demand. Rolling blackouts ensued in California for the first time. Blackouts again occurred in March and a third time in May. The state ran into capacity constraints despite low demand because of the continued shortage of power in neighboring states and the fact that a significant portion of in-state operating capacity was offline.

All power plants need to be taken offline periodically for maintenance, a function of the fact that power production is still largely mechanical. While the facts are still being sorted out, it appears that as much as 15,000 MW—twice the normal amount—was taken offline for a significant portion of the winter. Some observers claim that this was due to overuse of facilities during the previous summer that necessitated substantial mechanical overhauls in the low demand winter months. Others claim it was a deliberate attempt to 'fix' the market. It may also be that the increasingly perilous financial situation of the big utilities made wholesalers begin to doubt whether they would ever be paid. Wholesalers

Figure 7: Days of Power Alerts



Source: California Independent System Operator

might have chosen to shut down rather than incur the cost of generating power for uncertain revenues.

A final possibility is suggested here. Wholesale market participants may have believed that prices were soon going to fall again as capacity became unconstrained. They therefore might have felt that it was best to shut down high average cost power plants in advance. Whatever the circumstances, the results were clear. On top of the already high costs of production, prices rocketed to extremely high levels and the state was forced to step in as the power purchaser for the private utilities.

Lessons to be Learned

Summer 2001 was forecast to be very bad for California's power consumers. Many additional hours of rolling blackouts were predicted as power shortages continued. Forward rates for power were extremely high.[30] In response the state finally instituted a series of real fixes. Natural gas generators were removed from the RECLAIM program, lowering their cost of generation. The FERC imposed a marginal cost price cap that was tight enough to remove much of the pricing power being exerted by wholesalers. The legal and public scrutiny of the behavior of wholesalers may also have played a role in reducing any efforts to game the market. Conservation efforts were put into place, and retail electricity prices were raised by an estimated 40%. The state then started entering into long term contracts with wholesalers.

California was also aided by fortuitous events. While the troubles of the earlier period were created by a number of adverse conditions occurring simultaneously, these same factors all turned around by summer 2001. Some new power plants that had been in the works came online. Many plants that had been offline finally came back on. Natural gas prices fell. The warm weather of early 2001 was followed by an unseasonable cool summer in the West. There was no significant long heat wave. The result was very low wholesale power prices and no blackouts. Average retail prices have even fallen because

[30] There are a number of small markets that sell power in forward markets. These markets are small and not very liquid, but nonetheless give some indication of where participants believe prices will be heading.

of reduced demand and pricing methods used by the CPUC.

The current situation in California is still difficult. Long term contracts signed by the state under duress will keep average wholesale prices high. Power rates are currently low, but a price structure is now in place that will cause rates to rise rapidly if a significant heat wave hits the state. The debt the state is carrying as a result of the crisis is substantial and has yet to be financed. But these issues should be kept in perspective. As the state moves forward, the debt will eventually be paid off by consumers, and electricity rates will fall back to more normal levels. Deregulation will again be considered as an option, not just in California but also in the rest of the United States. A growing understanding that the crisis in California was due to flaws in the California plan rather than due to the deregulation concept is an important first step. Careful planning and debate are essential. Any deregulation plan that receives a unanimous vote of the legislature—we have learned—must surely be a bad one.

Here are a number of other important lessons that should be drawn from California's experience.

- States are not individual power markets but interact with each other substantially. As such, deregulation is best undertaken if committed to across states. California's problems would have been much reduced had other states been deregulated. An interstate power arbitrage could then have taken place that would have reduced California's wholesale prices while providing utilities in other states with enhanced revenues.
- Deregulation should take place all at once, rather than on a step by step basis. While floating retail rates would not have completely solved the California crisis, it would have reduced the impact substantially by introducing demand elasticity into the wholesale markets and by reducing the market power wielded by producers. Frozen retail rates prevented real competition at the consumer level and put the utilities in a net-short position.
- A day-ahead spot market system is fine for day to day needs. But it lacks a mechanism for long term capacity planning. This deficiency leads to long run boom-bust cycles that are ultimately bad both for individual consumers and the overall economy. Long term forward markets or long term contracts should be in place to provide the proper market signals. These long-term arrangements will also smooth retail prices over time. Capacity planning might also be put directly into the deregulation plan (as was done in Texas).
- Real time pricing should be a long run goal of any deregulation plan. Pricing power to reflect its true marginal cost is essential to the efficient operation of any market. The technology currently exists to easily put such systems into place, and they should be.
- All deregulation plans will have glitches and problems that need to be dealt with. Rapid and logical responses can reduce problems tremendously. More oversight with better response times should be put into place, and there should be one central regulatory body that takes ultimate responsibility for all aspects of the market rather than just one.

Bibliography

"When the lights go out," *The Economist,* January 18, 2001.

Borenstein S. "The Trouble With Electricity Markets (and some solutions)." University of California Energy Institute, *POWER Working Paper* - 081, January 2001.

Bushnell J. "West Coast Power Blues," *New York Newsday,* Op-ed Section, January 7, 2001.

Cambridge Energy Research Associates. "Short Circuit: Will the California Energy Crisis Derail the State's Economy?" June 2001.

Hildebrandt, Eric. "Further Analysis of the Exercise and Cost Impacts of Market Power In California's Wholesale Energy Markets." California Independent System Operator, March 2001.

Joskow, Paul L. "California's Electricity Crisis," *National Bureau of Economic Research Working Paper* #8442, August 2001.

Joskow, Paul L. and Edward Kahn, "A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market During Summer 2000," *National Bureau of Economic Research Working Paper* #8157, January 2001.

Maloney, M. "Economies and Diseconomies: Estimating Electricity Cost Functions," forthcoming, *Review of Industrial Organization*.

Maloney M., R.E. McCormick, and Raymond Sauer. "Customer Choice, Consumer Value: An Analysis of Retail Competition in America's Electric Industry." Washington, D.C.: Citizens for a Sound Economy, 1996.

Puller, S. L. "Pricing and Firm Conduct in California's Deregulated Electricity Market." University of California Energy Institute, *POWER Working Paper* - 080, January 2001.

Stoft S. "The Market Flaw California Overlooked" *New York Times,* Op-ed Section. January 2, 2001.

Thornberg, Christopher F. "Hey, Who Turned Out the Expansion?" *UCLA Anderson Forecast,* March 2001.