Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market

Severin Borenstein*, James Bushnell** and Frank Wolak***

December 2001

For helpful discussions and comments, we thank Carl Blumstein, Roger Bohn, Peter Cramton, Rob Gramlich, Paul Joskow, Alvin Klevorick, Christopher Knittel, Patrick McGuire, Catherine Wolfram, participants in the POWER Research Conference on Electricity Restructuring, and seminar participants at the Department of Justice and the NBER Summer Institute. Jun Ishii, Erin Mansur, Steve Puller, and Marshall Yan provided excellent research assistance.

* Haas School of Business, University of California, Berkeley, CA 94720-1900, University of California Energy Institute, and NBER. email: borenste@haas.berkeley.edu.

** University of California Energy Institute, 2539 Channing Way, Berkeley, CA 94720-5180. email: jimb@ieor.berkeley.edu.

*** Department of Economics, Stanford University, Stanford, CA 94305-6072, and NBER. email: wolak@zia.stanford.edu.

Borenstein is a member of the Governing Board of the California Power Exchange. Bushnell was a member of the Power Exchange's Market Monitoring Committee during 1999-2000. Wolak is chair of the California Independent System Operator's Market Surveillance Committee. The views expressed in this paper do not necessarily reflect those of any organization or committee.

Abstract

California's deregulated wholesale electricity market opened in 1998 and operated relatively smoothly for two years. Beginning in the summer of 2000, however, prices increased dramatically, more than tripling from their levels in the previous summer. Eventually, these unprecedented wholesale prices combined with a retail price freeze to bankrupt the largest investor-owned utility in the state and push the second largest to the edge of bankruptcy. While some observers said that this was simply the result of growing demand and inadequate supply, others argued that sellers exercised market power. We examine the degree of competition in the market from June 1998 to October 2000, just before the market effectively ceased operation. We find that there were significant departures from competitive pricing and that these departures were most pronounced during the highest demand periods, which tend to occur during the summer. We then parse the increased wholesale expenditures during summer 2000 into production cost increases, competitive rents, and increases due to the exercise of market power. We find that 58% of the change in wholesale market expenditures, which rose from about \$1.5 billion in summer 1999 to nearly \$8 billion in summer 2000, can be attributed to market power.

I. Introduction

In the spring of 2000, the momentum behind a dramatic restructuring of the electricity industry appeared to be irresistible. There were four regions of the U.S. with independent system operators running spot markets for wholesale electricity–California, PJM (major parts of Pennsylvania, New Jersey, Maryland, Delaware, Virginia and the District of Columbia), New England, and New York. Several other states were undertaking initiatives to restructure their electricity sector along similar lines. Beginning in summer 2000, however, soaring wholesale electricity prices in California made international news and threatened the financial stability of the state. The disruptions in California slowed, and threatened to reverse, the movement towards deregulating electricity markets in the U.S. and elsewhere.

In the aftermath of California's electricity crisis, policy-makers debated the correct lessons to take from the state's restructuring as well as the proper regulatory response, if any, to the crisis. Many of the answers to the questions being debated depend upon a proper diagnosis of the problems that plagued California's power sector during this period. Were soaring power costs the result of market "fundamentals" such as rising fuel prices, environmental cost, and a scarcity of generating capacity? Or were power suppliers able to exercise significant market power? In this paper we estimate the extent to which each of these factors - input costs, scarcity, and market power - influenced market outcomes in the California power market from 1998 through 2000. We analyze input and output prices, generator variable costs, and actual production quantities to measure the degree to which California wholesale electricity prices exceeded competitive levels. We also address the question of the efficiency impacts of market power in this market.

While market power has been studied and estimated in many industries, there has been little attention paid to intertemporal variation in the ability to exercise market power. For industries in which the good is storable, such intertemporal variation is necessarily small since inventories greatly reduce intertemporal supply variation, and possibly, demand variation. In markets for non-storable goods, including electricity and most services, this is not possible. The problem is exacerbated in electricity because demand is very inelastic in the short run and supply becomes very inelastic as production approaches the system generation capacity. Recognizing the dynamics of market power is likely to be important in both determining its causes and crafting remedies as part of the evolving public policy towards electricity restructuring.

Luckily, due to the history of regulation in electricity markets, data exist on the hourly output of all generating units and transmission power flows. In addition, information collected on the technical characteristics of each California generating unit during the former regulated monopoly regime – specifically, the heat rate (the efficiency with which the fuel burned is turned into electricity) and variable operating and maintenance costs – allows very accurate estimation of generating unit-level variable costs.

We find that all of the factors that have been identified as the cause of the California power crisis played a role in market outcomes. Due to rising input costs, even a perfectly competitive California electricity market would have seen wholesale electricity expenditures triple between the summers of 1998 and 2000. Market power, however, also played a very significant role. In summer 1998, 30% of total electricity expenditures could be attributed to market power, a figure that increased to 55% in summer 2000. The increased percentage margins combined with substantial production cost increases for marginal producers to create a drastic rise in absolute margins and push the market into a crisis in 2000.

In section II, we discuss the issues raised in estimating market power in electricity markets and the consequences of market power. We present an overview of California's electricity market in section III. In section IV, we describe the estimation technique in detail in the context of the California market, addressing each component of the market and outlining the assumptions made in implementing the analysis. We try to take a conservative approach, interpreting the data in a way that would be likely to understate the degree of market power exercised. We present the results of the analysis in section V. In addition to our estimates of premia of actual prices over the competitive levels, we also attempt to parse changes in competitive revenues between changes in actual costs and changes that reflect rents to inframarginal competitive sellers. We conclude in section VI.

II. Market Power Analysis in the Electricity Industry

During the 1990s, regulatory evaluation of short-run horizontal market power has focused on concentration measures, such as the Herfindahl-Hirschman index (HHI). Unfortunately, such measures are a poor indicator of the potential for, or existence of, market power in the electricity industry, because the industry is characterized by highly variable price-inelastic demand, significant short-run capacity constraints, and extremely costly storage.[1] It is easy to show that in such circumstances, firms with very small market shares could still exercise significant market power.

Fortunately, because of electricity's history as a regulated industry, a wealth of production cost and other data are available that allow for more direct analysis of the behavior of firms and the impacts of that behavior on market outcomes. From these data, we are able to construct a competitive counterfactual to which we can compare the actual market outcomes.

To construct the competitive counterfactual, we model each firm as a price-taker that would sell power from a given plant so long as the price it received was greater than the incremental cost of production. Of course, the cost of selling a unit of electricity can be greater than the simple production costs if the firm has an opportunity cost that is greater than its production cost, such as the revenue the firm would get from selling power or reserve capacity in a different location or market. On the other hand, a high price in an alternative market can reflect market power in that market, resulting in the transmittal of high prices across markets by the response of competitive suppliers. We discuss these alternative opportunities below.

Some analysts of the electricity industry have raised the concern that price-taking behavior on the part of every firm is simply too strict of a standard to be used as a benchmark. They argue that it is unrealistic to think that no market power will exist, because some market power exists in most markets. We recognize this fact, and that a government's attempts to intervene to mitigate market power can be more costly to society than the market power itself. While the presence of some market power should not be grounds for declaring restructuring of the electricity industry a failure or increasing the level of government intervention in the market, measures of the performance of electricity markets relative to a competitive benchmark should be the starting point for considering regulatory action.

---

[1] See Borenstein, Bushnell, and Knittel (1999) for a more detailed discussion of the applicability of concentration measures to market power analysis in electricity markets and citations to regulatory decisions that have relied on concentration indices.

Thus far, we have discussed only situations in which a firm unilaterally exercises market power. Antitrust law is most often concerned with collusive attempts to exercise market power. Unfortunately, many of the attributes that facilitate collusion are present in electricity markets: In most electricity markets, firms play repeatedly, interacting on a daily basis, so there is opportunity to develop subtle communication and collusive strategies. The payoff from cheating on a collusive agreement may be limited due to capacity constraints on production, though for the same reason, the ability to punish defectors may be limited. Finally, the industry has fairly standardized production facilities, so homogeneous costs may make it easier for firms to reach tacit or explicit collusive behavior. All that said, we have not explored the question of tacit or explicit collusion among firms in the California market.[2] Rather, in this paper we focus on market outcomes.

In focusing on outcomes, there are two indicators that clearly distinguish market power and each leads to a distinct estimation technique. First, in a competitive market, a firm is unable to take any action, including output decisions or offer prices, that significantly affects the price in a market. This suggests a method of estimation that involves studying the bidding and output decisions of each unit or firm in the market to detect successful attempts to affect prices. This is the general approach used by Wolak and Patrick (1996), Wolfram (1998), Wolak (1999), Bohn, Klevorick, and Stalon (1998), Bushnell and Wolak (1999), Puller (2001), and Mansur (2001).

The second empirical approach is at the market level, and is the one that we adopt here. We test whether the market as a whole is generating competitive prices given the production capabilities of all players in the market. As such, these tests are less vulnerable to the arguments of coincidence, bad luck, or ignorance that may be directed at analysis of the actions of a specific generator. This approach is less informative about the specific manifestations of market power, but is effective for estimating its scope and severity, as well as identifying how departure from competition varies over time. This is the general approach used in Wolfram (1999).

A potential drawback of this approach is that it captures all inefficiencies in the market, some of which may not be due to market power. If, for instance, the ISO systematically

---

[2] See Puller (2001).

held low-cost generators out of production simply due to a faulty dispatch algorithm, that would impact the estimate of market power. During the period we study, the California market clearly still had a number of design flaws that contributed to inefficient dispatch and market pricing. For the great majority of these, however, the flaw would be fairly benign if firms acted as pure price takers, rather than exploiting these design flaws to affect the market price. Furthermore, we find that over substantial periods of time, prices did not significantly differ from our estimates of marginal cost, indicating that there were no systemic inefficiencies raising prices in all periods. Still, the estimates must be taken with the caveat that they include failures to achieve competitive market prices for reasons other than market power, including bad judgment and confusion on the part of some generators or market-making institutions.

The Consequences of Market Power

In analyzing the efficiency consequences of market power in electricity, one must begin from the recognition that short-run electricity demand exhibits virtually zero price elasticity. Almost none of the customers in California, or anywhere else, are charged real-time retail electricity prices that vary hour-to-hour as wholesale prices do. Since, as we show below, the exercise of market power varies tremendously on an hourly basis, the absence of very-short-run elasticity is critical to understanding the consequences.[3] In studying California specifically, one much also consider that, during the 1998-2001 transition period, end-use consumers were insulated from energy price fluctuations by the Competition Transition Charge (CTC). The CTC is a mechanism that was implemented along with the restructuring of the industry in order to allow the incumbent utilities to recover their stranded generation costs. During the transition period, the vast majority of end-use consumers face fixed rate schedules that were imposed along with the CTC.[4] Thus, the CTC

---

[3] In California and elsewhere, time-of-use rates are common for large users. These price schedules generally have preset peak, shoulder, and off-peak rates, which are fixed for periods of months. They do not distinguish, for instance, a weekday afternoon with extremely high wholesale prices from a more moderate weekday afternoon. Borenstein (2001, 2002) argues that time-of-use rates are an extremely poor substitute for real-time electricity pricing and the real-time pricing would greatly mitigate wholesale price volatility.

[4] Even "direct access" consumers, who buy energy from some source other than their incumbent utility, were insulated from wholesale energy price fluctuations in the short-run by the CTC. This is because the stranded cost component paid by all consumers is calculated in a way that moves inversely to the energy price. The higher the energy price, the lower the CTC payment for that hour.

greatly lessened even the monthly elasticity of final consumer electricity demand.

Due to extreme short-run inelasticity of demand, market power in electricity markets has little effect on consumption quantity or short-run allocative efficiency. As described below, however, generating companies in California vary markedly in their costs and generation capacity, so the exercise of market power by one firm can lead to an inefficient reallocation of production among generating firms: a firm exercising market power will restrict its output so that its marginal cost is below price (and equal to its marginal revenue), while other firms that are price-taking will produce units of output for which its marginal cost is virtually equal to price. Thus, there will be inefficient production on a market-wide basis as more expensive, competitive, production is substituted for less expensive production owned by firms with market power. This is the outcome Wolak and Patrick (1996) described in the U.K. market, where higher cost combined-cycle gas turbine generators owned by new entrants provide baseload power that could be supplied by coal-fired plants which were being withheld by the two large generators exercising market power.

In addition, several recent analyses have demonstrated that the exercise of market power in an electricity network can greatly increase the level of congestion on that network.[5] This increased congestion impacts negatively both the efficiency and the reliability of the system. Market power can also lead firms to utilize their hydro-electric resources in ways that decrease overall economic efficiency.[6]

Lastly, electricity prices influence long-term decision-making in a way that can seriously impact the economy and efficient investment. While it has been pointed out that high prices should spur new investment and entry in electricity production, these investments may not be efficient if motivated by high prices that are caused by market power, which may indicate a need not for new capacity, but for the efficient use of existing capacity. Conversely, artificially high prices can lead some firms *not* to invest in productive enterprises that require the use of electricity, or to inefficiently substitute to less electric-intensive production technologies.

---

[5] See Cardell, Hitt and Hogan (1997), Bushnell (1999), Borenstein, Bushnell and Stoft (2000), and Joskow and Tirole (2000).

[6] See Bushnell (forthcoming).

Beyond the efficiency considerations, market power has potentially large and important redistributional effects. The California electricity crisis of 2000-2001 illustrates both the immense potential size of these effects and the difficulty of analyzing them. While the transitional retail rate freeze associated with the CTC prevented immediate passthrough of increased wholesale prices, the utilities were eventually driven to bankruptcy in one case, and very close in another, so the ratepayers or taxpayers ultimately became the bearer of much of these costs. At this writing, it is still unclear who will bear what share of the expense, and how much of the revenues paid to generators will be refunded to buyers under orders from federal regulators.

III. The California Electricity Market

Until December 2000, the two primary market institutions in California were the Power Exchange (PX) and the Independent System Operator (ISO). The PX ran a day-ahead and day-of market for energy utilizing a double auction format.[7] Firms submitted both demand and supply bids. In the day-ahead market, which was by far the largest market in California, firms bid into the PX offers to supply or consume power the following day for any or all of the 24 hourly markets. Although they were not envisioned as such, the PX markets were effectively financial, rather than physical, markets. As explained below, this is because firms could purchase or sell electricity in real time to change their day-ahead PX positions in what is essentially an energy spot market run by the ISO.[8] In addition to the PX, other institutions, known as "scheduling coordinators," (SCs) could submit the results of completed wholesale energy transactions to the ISO.[9]

The ISO is responsible for coordinating the usage of the transmission grid and ensuring that the cumulative transactions, or schedules, do not constitute a reliability risk, *i.e.,*

---

[7] The PX double auction accepted supply and demand bids from each market participant and set a market clearing quantity at the intersection of the resulting aggregate supply and demand curves implied by all of these bids.

[8] The transaction costs of trading in the PX relative to the ISO, or other institutions is a source of considerable confusion. For the purposes of this discussion we consider these differences to be negligible relative to the costs of the underlying commodity, electrical energy.

[9] In January of 2001, the PX ceased operation and the California Department of Water Resources (DWR) assumed responsibility for the bulk of wholesale purchases on behalf of all investor-owned utilities in California and began operating as its own SC.

are not likely to overload the transmission system.[10] As the institution responsible for the real-time operation of the electric system, the ISO must also ensure that aggregate supply is continuously matched with aggregate demand. In doing so, the ISO operates an "imbalance energy" market, which is also commonly called the real-time energy market. In this market, additional generation is procured in the event of a supply shortfall, and generators are relieved of their obligation to provide power in the event that there is excess generation being supplied to grid. Like the PX, this market is run through a double auction process, although of slightly different format. Firms that deviate from their formal schedules are required to purchase (or sell) the amount of their shortfall (or surplus) on the imbalance energy market.[11] During our sample period, no further penalties were assessed for deviating from an advance schedule. The imbalance energy market therefore serves as the *de facto* spot market for energy in California. During our sample period, the ISO imbalance energy market constituted less than 5% of total energy sales with the PX accounting for about 85% and the remainder taking place through bilateral trades.

The ISO also operates markets for the acquisition of reserve, or stand-by, capacity and for the relief of constrained transmission interfaces. These reserves are purchased through a series of auctions that determine a uniform price for the *capacity* of each reserve purchased. Most of the reserve capacity is also available to provide imbalance energy, and therefore will impact the spot price. A production unit committed to provide reserve capacity during an hour would therefore earn a capacity payment for being available and, if called upon in real-time, the imbalance energy price, for actually providing energy.

"Regulation reserve", the most short-term reserve, is provided by generation that is equipped to respond automatically to voltage fluctuations. Due to the nature of this reserve service, and to metering limitations, generation capacity providing regulation reserves cannot set the imbalance energy price. It does earn the hourly imbalance energy price for all net energy supplied during each hour it is providing regulation. As we describe below, we therefore consider units providing regulation services to be "held-out" of the market.

---

[10] Unlike the PX, the ISO has continued to function in approximately its original role through the 2000-2001 electricity crisis.

[11] The purchasing and selling is in fact done by the ISO itself, and accounts are settled through an ex-post adjustment.

Table 1a: California Generation Companies (MW)
July 1998 - Nameplate Capacity

Firm	Fossil	Hydro	Nuclear
PG&E	3700	5728	2160
SCE	1595	1002	2327
SDG&E	1990	0	430
Duke	2639	0	0
AES/Williams	3921	0	0
Reliant	3698	0	0
Dynegy	1585	0	0
Other	322	0	0

Table 1b: California Generation Companies (MW)
July 1999 - Nameplate Capacity

Firm	Fossil	Hydro	Nuclear
PG&E	570	5728	2160
SCE	1595	1002	2327
Duke	3343	0	0
AES/Williams	3921	0	0
Reliant	3698	0	0
Dynegy	2831	0	0
Southern	3130	0	0
Other	322	0	430

A. Market Structure of California Generation

The California electricity generation market at first glance appears relatively unconcentrated. The former dominant firms, Pacific Gas & Electric (PG&E) and Southern California Edison (SCE) divested the bulk of their gas-fired capacity in the first half of 1998 and most of the remainder in early 1999. SCE retained only a small proportion of its capacity not already covered under regulatory side agreements. These divestitures left the generation assets in California more or less evenly distributed between seven firms. The generation capacity of these firms during July 1998 and July 1999 is listed in Table 1. “Fossil” includes all plants that burned natural gas, oil, or coal to power the plant, but over 99% of the output from these plants is fueled by natural gas. The market structure during 2000 was largely unchanged from that of 1999.

As can be seen from Table 1, PG&E was the largest generation company during the summer of 1998. The seemingly dominant position of PG&E is offset to a large extent by its other regulatory agreements. All of its nuclear generation in California, for instance, is treated under rate settlements independent of market prices. More importantly, the incumbent utilities in California were the largest buyers of electricity during this time period. It was this purchasing obligation, combined with a cap on retail rates, that led to the bankruptcy of PG&E and near-bankruptcy of SCE in 2001.

B. Analyzing Market Power in California's Electricity Market

Critical to studying market power in California is an understanding of the economic interactions between the multiple electricity markets in the state. Participants moved between markets in order to take advantage of higher (for sellers) or lower (for buyers) prices. For instance, if the ISO's real-time imbalance energy price were usually higher than the PX day-ahead price, then sellers who saw this would reduce the amount of power they sell in the PX and sell more in the ISO imbalance energy market. These attempts to arbitrage the PX/ISO price difference would have the effect of raising the PX price and lowering the ISO imbalance energy price, thereby eliminating the price differential. For this reason, it is not useful to study the PX market, or any other of the California markets, in isolation. The strong forces of financial arbitrage mean that any change in one market that affects that market price will spill over into the other markets.[12]

This interaction of the different California electricity markets means that we must study the entire California energy market in order to analyze market power in the state. For this reason, in the analysis below we look at the entire generation in the ISO/PX service area regardless of whether the power from a generating plant is being sold through the ISO, the PX, or some other scheduling coordinator.

The recognition of the California power market as being effectively an integrated market due to arbitrage forces yields two other important insights. First, although the California market has a few very large buyers of electricity, the large utility distribution

---

[12] BBKW (2001) finds that although significant price differences between the PX and ISO did occur during individual months, overall, there was no consistent pattern of the PX price being higher or lower than the ISO price.

companies (UDCs), these companies do not possess monopsony power in the conventional sense. This is because these companies do not control the level of end-use demand of their customers. The UDCs cannot therefore reduce end-use consumption below efficient levels in order to lower overall power purchase costs. They did have some limited freedom as to *which* market they would purchase these needs from by choosing between day-ahead purchases in the PX and de-facto spot purchases from the ISO imbalance energy market. Because end-use demand was effectively inelastic, the downward sloping demand curves seen in the PX day-ahead market reflect a choice to wait for the imbalance energy market to purchase power if PX prices reached certain levels, rather than a choice to forgo consumption alltogether. Nonetheless, since sellers could move between markets as well, the buyers had no ability to persistently exercise monopsony power.

The second insight from a recognition of market integration involves the impact of price-caps in the various markets. Because the ISO imbalance energy market was the last market in a sequence of several, the level of the price-cap in the imbalance energy market fed back to form an implicit cap on prices in the other advance markets. The intuition for this is simple. Knowing that the maximum one might have to pay for power in real-time was capped at $250/MWh, for example, no buyer would be willing to pay in excess of that amount for purchases in advance. Thus, the aggregate demand curve in the day-ahead PX market became near horizontal at prices approaching the level of the price-cap in the ISO imbalance energy market.[13]

Many of the suppliers that compete in the ISO or PX also are eligible to earn capacity payments for providing ancillary services, as well as energy payments for generating in real time, if they bid successfully into one of the ancillary services markets. Ancillary services therefore represent an alternative use of much of the generation capacity in California. It is therefore necessary to consider the interaction between the energy and ancillary services markets. In the case of the California market, the relevant consideration is that the provision of ancillary services in most cases does not preclude the provision of energy in the real-time market. Thus for the bulk of generation the choice between ancillary service and energy markets is not an either-or choice.

---

[13] See Bohn, Klevorick and Stalon (1999).

It is important to recognize that the pool of suppliers available to ancillary services markets is very similar to that available to the energy markets. The main difference is that some generators are physically unable to provide certain ancillary services. Thus there are fewer potential suppliers for some ancillary services than there are for energy. We therefore would expect that the energy market would be at least as competitive as the ancillary services markets, and probably more so. It follows that price-cost margins would be at least as great if not greater in ancillary services markets than in energy markets. In fact the ancillary services markets, for a variety of reasons, appear to have been significantly less competitive than the energy market during the time period of our study.[14]

The other prominent opportunity for the usage of California generation is the supply of power to neighboring regions. Higher prices for energy outside of California could produce a result in which all generators within California were able to earn prices above their marginal cost, even if they behaved as price takers. For this to be the case, the California ISO region would have to be a net exporter of power. During our sample period, such conditions arose only in a total of seven hours out of the 22,681 hours in the sample. Even in these hours, the maximum net quantity of energy exported out of the ISO control area in any hour was a modest 896 MWh. Therefore, if we assume that trading in these markets was efficient, the export opportunities for producers within California were very limited relative to the California market.

Even if this were not the case, however, our analysis would fully account for the opportunity cost of exports because under the California market structure firms from other states purchased power through markets run by the PX and ISO. If there were sufficient demand for power in Arizona being purchased from the California markets, for example, this would, absent transmission congestion, raise the market clearing price within the California PX and ISO markets. If transmission were congested, then further exports would be infeasible. The market prices we utilize in our estimates should therefore already incorporate any opportunities for export from California.

---

[14] See Wolak, Nordhaus, and Shapiro (1998)

IV. Measuring Market Power in California's Electricity Market

The fundamental measure of market power is the margin between price and the marginal cost of production. As discussed above, if no firm were exercising market power, then all units with marginal costs that are below the market price would be operating. Even in a market in which some firms exercise considerable market power, the *marginal* unit that is operating could have a marginal cost that is equal to the price. When a firm with market power reduces output from its plants or, equivalently, raises its offer price for its output, its production is usually replaced by other, more expensive generation that may be owned by non-strategic firms.

In estimating a price-cost margin in this paper, we therefore must estimate what the marginal cost of serving a given level of demand *would be* if all firms were behaving as price takers. Unlike in most industries, there is enough information available about the costs of the generators to obtain fairly precise estimates of the price-cost margin. In the following subsections we describe the assumptions and data used for generating estimates of the marginal cost of supplying electrical energy in California.

A. Market Clearing Prices and Quantities

As described above, the California electricity market in fact consists of several parallel and overlapping markets. Fortunately, our assessment of the overall degree of market power is simplified by the fact that most sellers and buyers are free to participate in any of these markets. With this fluidity of market participants across markets, we would expect that the market clearing prices from each of these markets to be equal in expectation.[15]

Given that generation and distribution firms, as well as other power traders, can arbitrage the expected price of energy across these commodity markets, we rely upon the unconstrained PX day-ahead energy price as our estimate of energy prices in any given hour. We chose to rely upon the PX unconstrained price because the PX handled over 85% of the electricity transactions during our sample period and the unconstrained PX price

---

[15] One might be concerned that this arbitrage would not hold in light of the requirement during our sample period that the three investor-owned utilities buy all of their energy from the PX. Given the financial nature of the PX market, the full meaning of this requirement was somewhat ambiguous. More importantly, Borenstein, Bushnell, Knittel and Wolfram (2001) find that the PX and ISO prices track quite closely throughout most of our sample period.

represents the market conditions most closely replicated in our estimates of marginal costs. In particular, we do not consider the costs of transmission congestion or local reliability constraints in our estimates of the marginal cost of serving a given demand. The PX unconstrained price is also derived by matching aggregate supply with aggregate demand without considering these constraints. The resulting market clearing price therefore reflects an outcome that would occur in the absence of transmission constraints, just as our cost calculations reflect the outcome in a market in which all producers are price takers and there are no transmission constraints.[16]

It has been argued (see Harvey and Hogan (2000)) that the day-ahead PX price should be expected to systematically overstate the marginal cost of energy supply since sellers in the day-ahead market would include a premium in their offer prices to account for the opportunity of earning ancillary services revenues, which require that the units be not be committed to sell power in a forward market. However if this were true, then the PX price would also be systematically higher than the ISO real-time price, which would not include a such a premium since suppliers of ancillary services are also eligible to sell energy in the real-time market. Empirically this is not the case. Over our sample period, the PX average price was not significantly greater than the ISO average price (see Borenstein, Bushnell, Knittel, and Wolfram, 2001).[17]

The interaction of these energy markets also requires us to use the systemwide aggregate demand as the market clearing quantity upon which we base our marginal cost estimates. This level therefore includes consumption through the PX, other SCs, and any "imbalance energy" demand that is provided through the ISO imbalance energy market. Consumption from all of these markets is in fact metered by the ISO, which in turn allocates charges amongst SCs during an ex-post settlement process. We are therefore able to obtain these aggregate market clearing quantities from the ISO settlement data.

---

[16] We would like to emphasize again that we use the PX price as representative of the prices in *all* California electricity markets. This is not a study of the PX market and the market power we find is not limited to the PX market. It is the amount we estimate to be present in all California electricity markets.

[17] There is also a fundamental theoretical flaw in the Harvey and Hogan argument. Though option value would cause a firm to offer power in the day-ahead market at a price above its marginal cost, arbitrage on the demand side (and by sellers firms that do not qualify to provide ancillary services) would still equalize the market prices. The equilibrium outcome would just have a reduced share of power sold through the day-ahead market due to the foregone option value.

The acquisition of reserves by the ISO also requires discussion here. Since the ISO is effectively purchasing considerable extra capacity for the provision of reserves, it might seem appropriate to consider these reserve quantities as part of the market clearing demand level. However, with the exception of regulation reserve, as described below, all other reserves are normally available to meet real-time energy needs if scheduled generation is not sufficient to supply market demand.[18] Thus, the real-time energy price is still set by the interaction of real-time energy demand – including quantities supplied by reserve capacity – and all of the generators that can provide real-time supply. Therefore, we consider the real-time energy demand in each hour to be the quantity that must be supplied and capacity selected for reserve services to be part of the capacity that can meet that demand and, as such, to be part of our aggregate marginal cost curve.

The most responsive form of reserve is regulation. Units providing regulation services are required to automatically adjust their output levels in a way that allows the ISO to continuously balance supply and demand. Unlike the other forms of reserve, regulation capacity, is, in a way, held out of the imbalance energy market and its capacity could therefore be considered to be unavailable for additional supply. For this reason we add the *upward* regulation reserve requirement, which at times reaches 11% of demand, to the market clearing quantity for the purposes of finding the overall marginal cost of supply.[19]

B. Marginal Cost of Fossil-fuel Generating Units

In estimating the marginal cost of production for an efficient market, we use the fuel costs of each fossil-fuel generating unit as well as the variable operating and maintenance (O&M) cost of each fossil-fuel unit. For units under the jurisdiction of the South Coast

---

[18] Due to reliability concerns, the ISO at times has not utilized spinning and non-spinning reserves for the provision of imbalance energy (see Wolak, Nordhaus, and Shapiro (1998)). The conditions under which this occurs are somewhat irregular and difficult to predict. For the purposes of this analysis we have assumed that these forms of reserve are utilized for the provision of imbalance energy.

[19] Regulation reserve is procured for both an upward (increasing) and downward (decreasing) range of capacity. The ISO needs to be able to continuously increase and decrease the output levels of certain units in order to balance the system. Since the generation units that are providing *downward* regulation are, by definition, producing energy, the capacity providing downward regulation should not be considered to be held out of the energy market. Note also that by adding regulation needs to the market demand, we are implicitly assuming that all regulation requirements are met by generation units with costs below the market clearing price. To the extent that some units providing regulation would not be economic at the market price, this assumption will tend to bias downward an estimate of market Power.

Air Quality Management District (SCAQMD), we also include the cost of NOx emissions, which are regulated under a tradeable emissions permit system. The cost of NOx permits was not significant in 1998 and 1999, but rose sharply during the summer of 2000. The marginal cost of each fossil-fuel unit i, MC_i, is calculated by using its average heat rate multiplied by fuel cost and adding an estimate of variable O&M cost to that product. These cost estimates are detailed in Appendix A. Figure 1 illustrates the aggregate marginal cost curve for fossil-fuel generation plants located in the ISO control area that are not considered to be "must-take" generation (as described below) and shows how it increased over time due to higher fuel and environmental costs. Note that because the higher-cost plants tend to be the least fuel-efficient and the heaviest polluters, increases in fuel costs and pollution permits not only shift the supply curve, but also increase its slope since the costs of high-cost plants increase by more than the costs of low-cost plants.[20]

The supply curves illustrated in Figure 1 does not include any adjustments for "forced outages." Generation unit forced (as opposed to scheduled) outages have traditionally been treated as random, independent events that, at any given moment, may occur according to a probability specified by that unit's forced outage factor. In our analysis, each generation unit, i, is assigned a constant marginal cost mc_i – reflecting that unit's average heat rate, fuel price, and its variable O&M cost – as well as a maximum output capacity, cap_i. Each unit also has a forced outage factor, fof_i, which represents the probability of an unplanned outage in any given hour.

Because the aggregate marginal cost curve is convex, estimating aggregate marginal cost using the expected capacity of each unit, $cap_i * fof_i$, would understate the actual expected cost at any given output level.[21] We therefore simulate the marginal cost curve that accounts for forced outages using Monte Carlo simulation methods. If the generation units $i = 1, ..., N$ are ordered according to increasing marginal cost, the aggregate marginal cost curve produced by the j_{th} iteration of this simulation, $C_j(q)$, is the marginal cost of

---

[20] Our estimates assume that the rated capacities of the plants, cap_i, are strictly binding constraints. It has been pointed out to us that the plants can be run above rated capacity, but at the cost of increased wear and more frequent maintenance. If we incorporated this factor – about which there seems to be very little detailed information – it would shift rightward the industry supply curve and would increase our estimates of market power.

[21] For any convex function $C(q)$, of a random variable q, we have, by Jensen's inequality, $\mathrm{E}(C(q)) \geq C(\mathrm{E}(q))$.

the k_{th} cheapest generating unit, where k is determined by

$$k = \arg\min x \mid \sum_{i=1}^{x} I(i) * cap_i \geq q. \qquad [1]$$

where $I(i)$ is an indicator variable that takes the value of 1 with probability of $1 - fof_i$, and 0 otherwise. For each hour, the Monte Carlo simulation of each unit's outage probability is repeated 100 times. In other words, for each iteration, the availability of each unit is based upon a random draw that is performed independently for each unit according to that unit's forced outage factor. The marginal cost at a given quantity for that iteration is then the marginal cost of the last available unit necessary to meet that quantity given the unavailability of those units that have randomly suffered forced outages in that iteration of the simulation. If, during a given iteration, the residual demand (after imports) exceeded available capacity, the price was set to the maximum allowed under the ISO imbalance energy price cap during that period. Note that often the PX price in that hour did not hit this level, so such outcomes were counted as "negative market power" outcomes in the analysis. Thus, these outcomes are not driving, and if anything are reducing, our finding of market power.

We did not adjust the output of generation units for actual outages. This is in part because the scheduling, and duration of planned outages for maintenance and other activities is itself a strategic decision. Wolak and Patrick (1996) present evidence that the timing of such outages was extremely profitable for certain firms in the U.K. electricity market. It would therefore be inappropriate to treat such decisions as random events. Since we find market power in the summer months – high demand periods in California in which the utilities have historically avoided scheduled maintenance on most generation – it is unlikely that scheduled maintenance could explain these results in any case.[22] We would expect scheduled maintenance to take place in the autumn, winter and spring months, which is the time period over which we find little, if any, market power. Unfortunately, the ISO did not begin reporting either planned or unplanned outages until 2001, after the end of our sample period.

---

[22] Scheduled maintenance on must-take resources, such as nuclear plants, and reservoir energy sources was accounted for under the procedures outlined in the following sections.

The operation of generation units of course entails other costs in addition to the fuel and short-run operating expenses. It is clear that sunk costs, such as capital costs, and periodic fixed capital and maintenance expenses should not be included in any estimate of short-run marginal cost. More difficult are the impacts of various unit-commitment costs and constraints, such as the cost of starting up a plant, the maximum rates at which a plant's output have be ramped up and down, and the minimum time periods for which a plant can be on or off. These constraints create non-convexities in the production cost functions of firms. For a generating unit that is not operating, these costs are clearly not sunk. On the other hand, it is not at all obvious that it is optimal for a price-taking profit-maximizing firm to pro-rate such costs into its supply bids. In fact, it is relatively easy to construct examples where it would clearly *not* be optimal to do so.[23] For the time being, we do not attempt to capture directly the impacts of these constraints on our cost estimates, although we do discuss interpretations in light of these non-convexities.

C. Imports and Exports

One of the most difficult aspects of estimating the marginal cost of meeting total demand in the ISO system is accounting for imports and exports between the ISO control area and other control areas. Unlike must-take generation, many imports and exports are not a result of pre-existing contractual arrangements. Although we can observe the net amount of power entering or leaving the ISO system at each interface point, we do not have data on the value (or opportunity cost) of that power outside California, nor on the cost of transmitting power to the interface point.

If the power market *outside* of California were perfectly competitive, then the marginal generator that is importing into California would, absent transmission constraints, have a marginal cost about equal to the market price in California. When market power is exercised within California, this would mean that, in an effort to drive up price, some in-state generators are withdrawing (or raising the offer price on) their marginal generation

---

[23] For example, consider a generator who estimates that it will be 'in' the market for six hours on a given day and bids into the market in each hour at a level equal to its fuel costs plus 1/6 of its start-up cost. Now imagine a market outcome where the price in one hour rises well above this bid level, but in subsequent hours remains at a level above the unit's fuel costs, but below the sum of its fuel cost and pro-rated start-up. This unit thereby has committed to operate in one hour, but is 'out' of the market in subsequent hours, even though it could have cleared an operating profit at market clearing prices.

and allowing more-expensive imported power to be substituted for it. Thus, in the absence of market power, we would see lower imports. This means that the cost of serving the demand that remains after the competitive level of imports is netted out would be higher than the cost of serving the demand that remains after the true level of imports is adjusted for.[24]

Figure 2 illustrates a hypothetical marginal cost curve of the in-state generation, excluding must-take and reservoir (hydro and geothermal) energy resources. The market demand is q_{tot}, and the observed price is p_{px}. At a price of p_{px}, we see imports of q_{imp} $(= q_{tot} - q_r)$ that shift the remaining demand to the left to a quantity q_r. If the price were instead set at the competitive price of p^*, we would see imports at some level less than or equal to those seen at p_{px}. This would shift the residual in-state demand curve to a quantity q_r^*. Thus, in order to estimate the price-taking outcome in the market, we need to estimate the import supply function.

Estimating the Import Supply Functions: One of the primary responsibilities of the California ISO is to ensure the reliable usage of the system's transmission network. This requires that the ISO sometimes operate a market for rationing transmission capacity when its use is oversubscribed. This market is implemented through the use of schedule "adjustment" bids, which are submitted by scheduling coordinators to the ISO along with their preferred day-ahead schedules.

Scheduling coordinators submit their preferred import quantities and the ISO verifies that these imports do not exceed transmission capacity limits. If these proposed imports are feasible, no further adjustments are required. In the event that the net of proposed import and export schedules does exceed transmission capacity on some interface, the ISO initiates a process of congestion relief by adjusting schedules according to their adjustment bids. Adjustment bids establish, for each schedule coordinator, a willingness-to-pay for transmission usage. Schedules are adjusted according to these values of transmission usage, starting at the lowest value, until the congestion along the interface is relieved. A uniform

---

[24] Capacity constraints on both the transmission interfaces into California and the production capacity of non-Californian producers complicate this intuition somewhat. If such a capacity constraint were binding at the observed California market clearing price, then the marginal *production* cost of imports would most likely be below this market clearing price. In such a circumstance, one cannot say with certainty that a perfectly competitive price within California would yield lower imports.

price for transmission usage, paid by all SCs using the interface, is set at the last, or highest value of transmission usage bid by an SC whose usage was curtailed.

The adjustment bid process is intended to allocate scarce transmission capacity to its most valued uses, and to price that capacity based upon those values. Adjustment bids take the form of supply and demand curves located on either side of a congested transmission interface. A price-taking SC that is importing power into California, for example, would submit as adjustment bids its cost of imported power on one side of the interface, and its resale value of that power on the other side. The difference between the import cost and resale value is the schedule coordinator's value of using the transmission interface. If this value is less than the transmission usage charge, the SC would want its schedule to be curtailed. If the transaction value is still greater than the transmission usage charge, then the SC would want the scheduled import to proceed. If the quantity of imports at the unconstrained PX price, when combined with imports from other SCs, exceeds the import capacity, then import quantities from this zone are adjusted downward according to their adjustment bids. This adjustment continues until the import quantity is feasible. At this point, only the imports that are profitable, even with the transmission charge, remain.

The adjustment bids provide information about the elasticity of imports. Adjustment bids reveal the willingness-to-supply of imported energy at each interface over a wide range of import quantities, not just at the observed import quantity. By aggregating the import adjustment bids over all transmission interfaces and over all schedule coordinators, we can establish an upper bound on import supply. Let the import supply curve of schedule coordinator sc at import zone z be the net of its preferred import quantity and all of its incremental and decremental adjustment bids *into California* from z.

$$q_z^{sc}(p) = q_z^{sc,init} + \sum_{\hat{p}<p} q_z^{sc,inc}(\hat{p}) - \sum_{\hat{p}>p} q_z^{sc,dec}(\hat{p}). \qquad [2]$$

In other words, the ideal level of imports from sc at z at a price of p, would be the sum of its scheduled imports, which are independent of price, and the amount of extra supply it is willing to provide at a price at or below p less the amount of supply it does not want to produce at price p. An aggregate import curve into the California ISO system for any hour can be calculated by summing the value of $q_z^{sc}(p)$ over all interfaces and SCs:

$$q_{imp}(p) = \sum_{sc} \sum_{z} q_z^{sc}(p). \qquad [3]$$

This aggregation constitutes an upper bound because the ISO is in practice prevented from substituting import adjustments across individual schedule coordinators or across transmission interfaces, so that the actual import supply curve will be a significantly steeper function of price than the curve constructed as described. The ISO will only act in the event that the initial schedules indicate that congestion will arise, even though the adjustment bids may indicate a potential Pareto improving import adjustment. Thus, while our aggregate import supply curve assumes that all imports from all locations are perfect substitutes, and that these imports are priced at marginal cost, reality falls short of this level of import efficiency. One consequence of this is that the import quantities implied by the aggregate of the adjustment bids do not exactly equal the imports that are actually observed. To realign the import supply curve implied by the adjustment bids with the observed import-price pair for each hour, we calculate the change in imports in each hour according to the following equation:

$$\Delta q_{imp}(p) = q_{imp}(p) - q_{imp}(p_{actual}). \qquad [4]$$

The adjustment made in equation [4] ensures that at prices equal to the actual observed price, there would be no change from the observed level in imports when performing our counter-factual price calculation. A positive $\Delta q_{imp}(p)$ implies an increase in imports relative to the observed price.

D. Hydroelectric and Geothermal Generation

Energy-limited units (*i.e.,* hydro and geothermal units) present a different challenge since the concern is not over a *reduction* in output relative to observed levels but rather a *reallocation* over time of the limited energy that is available to them.

The bids of hydro units provide less information about their usage in a perfectly competitive market than one might expect at first glance. The bids of hydro units do not reflect a production cost but rather the opportunity of using the hydro energy at some later time. In the case of a hydro firm that is exercising market power, this opportunity cost would also include a component reflecting that firm's ability to impact prices in different hours.[25] It is important to note that even the observed bid prices of a price-taking hydroelectric firm would provide little information about the opportunity cost of the energy in a competitive market, because the actual opportunity cost of water for these units will be influenced by the expectation of future prices, which is in turn impacted by the ability of other firms to raise those prices.

For these reasons, we make the assumption that the actual, observed output of these resources is the output that would be produced by a price-taking firm acting in a perfectly competitive market. In other-words we take the observed releases of reservoir energy as the optimal schedule that would result in least-cost production in a competitive market.

In practice, this assumption means that, in constructing our estimate of the marginal cost of meeting load in any given hour, we apply the observed production of hydro and geothermal resources for each hour and then calculate the marginal cost of satisfying the remaining demand with the state's fossil-fuel resources.

For the purpose of calculating the impact of market power on total production cost, it is easy to see that this is a conservative assumption, one that will produce downward biased estimates on the efficiency effects of market power. The optimal hydro schedule will, by definition, lead to weakly lower production cost than any other hydro schedule. To the extent that actual production differed from the optimal schedule, it could only raise total production cost. Thus our assumption will bias upward our estimate of perfectly competitive production cost.

For the purpose of measuring market power, we need to consider the impact of our assumption on our estimates of marginal production cost. Of concern is the possibility that the observed hydro schedule (which may include a response by hydro firms to the exercise of market power by others) – when combined with a counter-factual perfectly competitive production of fossil-fuel resources – could produce a *lower marginal cost* estimate on average than the optimal hydro schedule. However, it can be shown that when system marginal production costs from non-hydro sources are convex in quantity, any reallocation of hydro

---

[25] See Bushnell (2002).

energy away from the least-cost allocation will raise marginal costs more in the hours from which energy is removed than it will reduce marginal cost in the hours to which energy is added.[26] Thus our assumption of optimal hydro production can only bias our time-weighted estimates of marginal cost upwards, and therfore our estimates of price-cost margins downward.

We also present results in which price-cost margins are weighted by the market volumes in each hour. To consider the effect of our hydro assumptions on these results, we need to address the possibility of a reallocation of hydro energy between off-peak and peak hours relative to the optimal schedule. A hydro firm that is attempting to exercise market power would likely allocate less hydro energy during peak hours than would be the case for a price-taking firm (see Bushnell (2002)). This strategic hydro allocation, when combined with competitive fossil-fuel production, would produce a higher weighted average of marginal cost than would the optimal schedule. To the extent the firms controlling hydro resources attempted to exercise market power with those resources, our results will therefore understate the overall level of market power.

However, the vast majority of reservoir resources were controlled by the PG&E and SCE. These utilities were selling power under a retail rate freeze and therefore had a fairly strong incentive to lower wholesale power costs. Therefore it is possible that these firms responded to an increase in market power with an over-concentration, relative to perfect competition, of energy during high demand periods. As argued above, this reallocation (if allowed by the flow constraints) would raise off-peak marginal costs more than it would lower on-peak marginal costs. However, since (non must-take) market volumes are likely to be higher on-peak, the impact on the quantity-weighted average of marginal cost is uncertain.

We examine this issue empirically by asking whether our estimates of marginal costs produce opportunities for a reallocation of hydro energy that would result in a lower weighted-average marginal cost. Such an opportunity would exist if fossil-fuel marginal costs during some high demand period were lower, due to overproduction from hydro

---

[26] This is because at the least-cost allocation of hydro energy, marginal fossil-fuel costs will be equalized over all hours for which hydro flow constraints allow a discretionary use of hydro energy.

sources, than the marginal cost in some lower demand period. If, by contrast, our marginal cost estimates (over a period with stable input prices) were monotonic in market demand, then no obvious opportunity for lowering the weighted average of marginal cost exists and any bias from our hydro assumption is likely raising costs and lowering market power.

Since each individual estimate, and realization, of marginal cost reflects a probabilistic outcome on unit outages, we would not expect that individual observations would exhibit a strictly monotonic relationship with demand. We instead perform a kernel regression of marginal cost on demand in order to detect whether in aggregate there are systematic deviations from a monotonically increasing relationship between demand and our estimate of system marginal cost. Such regressions for each of the three summers in our sample period show that our marginal cost estimates were on average monotonically increasing within each of these time periods. This leads us to conclude that it is highly unlikely that our assumption that the actual schedule of production reservoir resources was the cost-minimizing schedule creates a significant negative bias on the weighted average estimates of marginal costs.

E. Calculating Cost Increase Relative to Competitive Outcome

Utilizing the assumptions outlined in the previous sections, we estimated the perfectly competitive market price in the California energy markets for each hour of market operation from June 1998 through October 2000. The residual market demand q_r, to be met by in-state fossil-fuel units within the ISO system, is estimated to be

$$q_r^t(p) = q_{tot}^t + q_{reg}^t - q_{mt}^t - q_{rsv}^t - q_{imp_{act}}^t - \Delta q_{imp}^t(p). \qquad [5]$$

where, q_{tot}^t and $q_{imp_{act}}^t$ is the actual ISO metered generation (with net imports), and total net imports for hour t, respectively. q_{tot}^t includes generation scheduled through all energy markets associated with the ISO control area, including the PX, ISO imbalance energy market, and other SCs. q_{reg}^t represents the addition to demand due to the need for capacity dedicated to regulation. The quantities q_{mt}^t and q_{rsv}^t represent the amount of energy produced by must-take generation and by reservoir (hydro and geothermal) generation, respectively. These quantities are all treated as price inelastic. The level of imported energy, $q^t_{imp_{act}} - \Delta q^t_{imp}(p)$ is adjusted by the market clearing price, as described above.

We make 100 fossil-fuel generation cost estimates, each reflecting a combination of independent Monte Carlo draws for the outage of a generation unit, for each hour. For each of these draws from the system-wide marginal cost curves we compute the intersection of this marginal cost curve with the residual market demand curve $q^t_r(p)$. This yields an estimated marginal cost and an instate market-clearing quantity q^t_{rj} for Monte Carlo draw j. We denote the marginal cost associated with this quantity as C^t_j. We can then compute an estimate of the expected value of the marginal cost of meeting the instate demand that results from price-taking behavior by instate generators as:

$$\bar{C}^t = \frac{\sum_{j=1}^{100}(C^t_j)}{100}. \qquad [6]$$

Note that there are cases in which $P^t_{px} - \bar{C}^t$ is negative in our simulations. Absent an operational error or an attempt at predatory pricing, firms will not actually be willing to sell power at prices below their *true* economic short-run marginal costs. In other words, prices will not be below the perfectly competitive price. Nonetheless, during some hours, particularly June 1998 and during the winter and spring of 1999, PX prices were below our estimates of the perfectly competitive market price. At least three factors contribute to these outcomes.

First, our cost estimates can exceed the actual marginal cost because we do not consider the dynamic effects of unit commitment constraints, such as start-up costs, ramping rates and minimum down times. These constraints can create opportunity costs of shutting down units that, in essence, lower the true marginal cost of operating that plant. Of course these same constraints also can create opportunity costs that, at other times, raise the true marginal cost. This is one reason why we include the negative mark-ups in our results; we did not want to exclude the off-peak impact of these constraints on our cost-estimates, since there is an opposite effect on our estimates during peak-hours.

Second, cost information for generating units are not exact data on which all parties agree. In some cases, our estimates of a unit's marginal cost could be slightly too high and in others slightly too low. Therefore we include negative price-cost differences in order to prevent truncating the effect of data uncertainty on our cost-estimates.

Third, and probably most importantly, our calculations do not control separately for the output levels of *reliability must-run* (RMR) generation, since we focus on the PX unconstrained price. Some generation units have been declared must-run for *local* grid reliability under certain conditions. These generators get separate non-market payments when they are called under the RMR contracts they have signed with the ISO. RMR units are not dispatched as part of the system. Because they are held out and paid a different price, the resulting price in the PX can be below the marginal cost to the system if the power provided by RMR units were instead provided as part of the full dispatch of the system. In fact, due to the high level of RMR calls by the ISO during the time period we study, particularly the spring of each year, it is possible that no other fossil-fuel generation was economic during some time periods. In those cases, the highest (opportunity) cost units selling in the PX could be hydro or out-of-state coal plants, either of which have lower marginal cost than any of the fossil-fuel plants we examine. However, these periods are likely to occur when the PX price is extremely low, not extremely high. In such cases, import energy with costs below those of in-state fossil-fuel generation could be the marginal generation, and the actual PX price could be lower than the marginal costs of any of the fossil-fuel units we have examined. Because we don't account for the RMR units, our estimates could still indicate that a fossil-fuel unit is marginal and its cost is the system marginal cost, so our estimated system marginal cost would be above the actual PX price due to unaccounted-for RMR calls.[27]

If the estimated MC is above the PX price for either of the first two reasons, then it seems that the most accurate estimate of market power would come from including the "negative market power" outcomes in our calculations. However, the total startup costs for the fossil-fuel units in California is almost surely less than \$100 million during our sample period, less than 2% of the effects we find.[28] In addition, there are other reasons

---

[27] This implies that neglecting RMR calls could underestimate market power. In addition, it appears from preliminary evidence that the implementation of RMR agreements has exacerbated some of the *local* market power problems that they were designed to mitigate. See Wolak, Nordhaus, and Shapiro (1998).

[28] For most of the units in our fossil-fuel cost curve, these costs had been estimated for inclusion in the

to think that startup costs explain only a minor part of the deviations from marginal cost pricing. First, the units that turn on to meet peak demand during the summer have little or no startup costs, so the impact at the times we find the greatest market power is likely to be low. Second, we find virtually no margins in the winter, when there are more plant startup costs in aggregate than during summer months.[29] Finally, our estimates of market power are substantially greater in summer 2000 than in summer 1999, but the number of startups by instate fossil-fuel plants were virtually the same. During June-September 1999, these plants in aggregate had 5006 startups. They had 5055 startups during the same period in 2000.

Likewise, it is unlikely that much of the negative market power outcomes could be the result of cost data errors. Many PX prices in June, for instance, were well below the costs that anyone has claimed for operation of fossil-fuel generating units.[30] Thus, it is most likely that the cost estimates that exceed the PX price occur because there were no fossil-fuel generating units that were economic to run at the time. Only fossil-fuel units running under RMR contracts were active. In that case, the marginal cost of the system, and thus the market price, is being set by much cheaper out-of-state coal plants, by nuclear plants, or by hydro or geothermal plants. If this is the case, then the proper treatment would be to truncate the results, resetting any finding of "negative market power" to set marginal cost equal to price. Still, in order to avoid biasing the results in favor of finding market power, we don't truncate the negative outcomes in the primary results we report.

V. Results

We computed the expected perfectly competitive price each hour for the months of June 1998 through October 2000 using the algorithm described above. From the import

---

contract payments for RMR performance. Total annual costs for start-ups of RMR units were about $17 million for the first year of operation. Even adjusting for the greater number of startups per year that occurred in later years, the total cost is unlikely to exceed $100 million.

[29] Additionally, if marginal cost functions turn upward smoothly around the rated capacity, rather than having a strict L-shape, the typical argument that a competitive plant would bid its startup costs for the "single hour" it would run are incorrect. In that case, even the last plant turned on would run for many hours because it would be replacing higher-cost output from other plants that would otherwise be producing along the steepest parts of their MC curves.

[30] If we were to ignore any "negative market power" outcomes for prices below, say, $18/MWh, virtually all of the "negative market power" effects would be eliminated.

adjustment bids, the average hourly reduction in imports from the observed level at the PX price versus the level at our estimate of marginal cost was 5.9%, with a standard deviation of 28.8%. For each hour, we can calculate an arc elasticity of import supply from the change between the competitive and actual price and the resulting change in imports. The median arc elasticity of import supply for these hours is 0.56.

The added wholesale cost of energy due to departures from a competitive market, ΔTC, is calculated by taking the difference between the PX price and our estimate of marginal cost and multiplying it by the total ISO metered generation less the must-take energy for that hour.[31] That is, for hour t,

$$\Delta TC^t = [p_{px}^t - \bar{C}^t] * [q_{tot}^t - q_{mt}^t], \qquad [7]$$

where $\bar{C}^t$ is expected marginal cost in period t. This expectation is taken with respect to the distribution of generating unit outages.

For any set of hours $\{S\}$, our measure of market performance is

$$MP(S) \ = \ \frac{\sum_{t \in S} \Delta TC^t}{\sum_{t \in S} TC^t}. \qquad [9]$$

$MP(S)$ is the average increased wholesale cost of electricity during all hours in $\{S\}$. Defining MP(S) in this manner is consistent with the view, reflected in our competitive benchmark Monte Carlo simulation, that the observed market price is conditional on a realization from the joint distribution of generating unit outages. To reflect this fact, let $\hat{p}_{px}^t$ denote the observed PX price for hour t and $E(\hat{p}_{px}^t)$ the expectation of this magnitude with respect to the joint distribution of generating unit outages. Unlike the counterfactual case assuming price-taking behavior, we cannot draw from the distribution of generating unit outages and compute a distribution of market prices that reflect the current level of market power. This would require a model for the strategic interaction of among players in the California market. However, by defining MP(S) as shown in equation (8), we can take advantage of the law of large numbers to prove that our measure is a consistent estimate

---

[31] By taking the observed quantity as the market demand, we are, for the reasons discussed earlier, implicitly assuming that demand is price inelastic.

of the average cost increase. To show this, re-write the index as:

$$MP(S) \;=\; \frac{1/Card(S)\sum_{t\in S}[\hat{p}^t_{px} - \bar{C}^t] * [q^t_{tot} - q^t_{mt}]}{1/Card(S)\sum_{t\in S}\hat{p}^t_{px} * [q^t_{tot} - q^t_{mt}]}, \qquad [8']$$

where Card(S) is the cardinality or number of elements (hours) in the set S. For sets S with a large number of elements, the index is approximately equal to

$$MP(S) \;=\; \frac{\sum_{t\in S}[E(\hat{p}^t_{px}) - \bar{C}^t] * [q^t_{tot} - q^t_{mt}]}{\sum_{t\in S}E(\hat{p}^t_{px}) * [q^t_{tot} - q^t_{mt}]}, \qquad [8'']$$

which is equal to the ratio of the expected cost increase relative to the perfectly competitive benchmark, due to the current level of market power and market imperfections, divided by the expected cost of purchasing electricity under current market conditions.

Table 2 reports the PX price, estimated marginal cost, and the added cost of power due to prices that exceeded marginal cost for each month in the sample period. As is evident from Table 2, June 1998 produced very idiosyncratic results, with an average PX price considerably below our estimate of marginal cost. The market was only in its third full month of operation at this time, and many of the fossil-fuel generation units were at this time going through a process of ownership transfer and regulatory approval. As described above, the use of reliability must-run contracts, which provide fossil-fuel units with subsidies beyond the market price, was also widespread during June. For these reasons, we believe the market results from June 1998 do not provide much meaningful information on the state of competition in the California market. Nevertheless, for completeness, we have included these results. For the set of all hours over the entire 29 month period that we study, the MP(S) is equal to 35%, amounting to total payments in excess of competitive levels equal to \$6.09 billion with a standard error of \$741 million.[32]

Having generated estimates of price-costs margins for each hour of the 29-month sample period, we can examine subsets of the data to gain insight into the underlying dynamics of the market. One test of the credibility of our results is whether our estimates of market power vary in the way that economics would predict. We would expect market power to be quite low during the off-peak months, particularly January through April. Electricity

---

[32] Appendix B outlines our procedure for computing this standard error, which accounts only for the error associated with the randomness of forced plant outages.

Table 2: Actual Price and Estimated Marginal Cost

Month	Year	mean of actual production per hour (MWh)	mean of PX price (\$/MWh)	mean of marginal cost (\$/MWh)	sum of ΔTC (millions \$)	aggregate $\frac{\Delta TC}{TC}$
June	1998	25372	12.09	21.88	-48	-48%
July	1998	28964	32.41	26.62	114	30%
August	1998	31320	39.53	27.09	228	41%
September	1998	28061	34.01	25.56	140	35%
October	1998	25649	26.65	25.50	22	8%
November	1998	24768	25.74	26.93	2	1%
December	1998	25637	29.13	24.79	52	19%
January	1999	25971	20.96	21.76	2	1%
February	1999	25593	19.03	20.56	-7	-4%
March	1999	26050	18.83	20.21	-7	-4%
April	1999	25562	24.05	23.85	11	5%
May	1999	25736	23.61	24.74	7	3%
June	1999	28356	23.52	25.23	18	7%
July	1999	30647	28.92	26.36	76	19%
August	1999	30398	32.31	29.83	69	16%
September	1999	28466	33.91	29.53	72	18%
October	1999	27333	47.63	33.56	195	32%
November	1999	26252	36.91	28.12	115	28%
December	1999	27122	29.66	27.04	40	11%
January	2000	26684	31.18	26.92	59	16%
February	2000	26922	30.04	28.86	19	6%
March	2000	26477	28.80	30.73	-11	-3%
April	2000	26524	26.60	31.68	-37	-13%
May	2000	27412	47.22	39.64	161	26%
June	2000	31069	120.20	52.70	1220	64%
July	2000	30825	105.72	56.75	894	54%
August	2000	32862	166.24	71.87	1610	59%
September	2000	30067	114.87	74.01	631	39%
October	2000	27550	101.51	67.03	462	35%

demand is low in these months and supply is relatively large due to the resurgence in hydro production from winter rains. In January-April 1999, we find an average MP(S) of -0.2%, and in January-April 2000, we find an average MP(S) of 2.3%. Thus, we find that there was essentially no margin between prices and marginal cost during the period in which supply was most abundant compared to demand and sellers had the least ability to exercise market power. In addition, these results provide evidence that significant short-run operating costs are not missing from our cost estimates, because negative or zero margins would not be observed over such an extended period of time.

The series of events that led to the California electricity crisis in 2000-2001 began with dramatic price increases during the summer of 2000. Many policymakers and regulators have argued that the competitive performance of the market fundamentally changed during summer 2000, thereby initiating the crises. In order to make such comparisons, however, one must account for differences in the relative levels of demand during these periods. As described above, we would expect the estimated market power to increase as the demand faced by in-state non-utility sellers rises relative to the capacity of these players.[33] Figure 3 shows a kernel regression of this relationship for August and September of 1998, 1999, and 2000.[34] The horizontal axis of figure 3 is the demand faced by these firms after accounting for actual imports, must-take, and hydro production. The vertical axis is the ratio $\Delta TC^t/TC^t$, which is equal to the Lerner index for that hour.[35]

The results summarized by this figure show that market power steadily increased with the residual demand faced by the non-utility in-state suppliers consistent with the earlier discussion of nature of competition in the electricity industry. During lower demand hours and months, as well as springtime months when significant hydro energy is available, no single firm can affect prices significantly. During higher demand hours, however, competitive sources of energy begin to reach their capacity limits and the pool of potential competitors for additional supply dwindles. Because of the lack of significant storage capacity and the inelasticity of demand, firms can take advantage of the capacity limits of their competitors during these high demand hours. This is consistent with the effects detected from the oligopoly equilibrium simulations in Borenstein and Bushnell (1999).

---

[33] The utilities had little or no incentive to exercise market power because they were net buyers of electricity and the revenue from power that they sold into the PX was just netted out from their power purchase costs.

[34] These are historically two of the highest demand months of the year in California, and they can exhibit the lowest supply availability due to declining hydro resources late in the summer. We focus on these two months because the ISO energy price-cap varied during our sample period, but it was set at the same level, \$250/MWh, for August and September of all three years.

[35] Because the Lerner index is not symmetric around zero, negative values of the ratio are set to zero to maintain a reasonable scale for the figure.

Our results also indicate that, given the supply and demand conditions during that period, the performance of the market was not dramatically different in 2000 than during 1998 and 1999. Figure 4 shows the cumulative density functions for the demand met by instate fossil-fuel generation for these 2 months during 1998-2000. Although total market demand was only 6-7% higher during 2000 than in August and September of the previous two years, the demand met by instate fossil-fuel plants, nearly all of which were unregulated by 1999, increased from an average of 6798 MW/hr during 1998 and 5624 MW/hr during 1999 to 10108 during 2000. This is largely due to a substantial decline in imports from an average of 6948 MW/hour in 1999 to 3956 MW/hour in 2000. Thus, although the performance of the market controlling for the demand faced by instate fossil-fuel generation did not change significantly during 2000, the average *level* of this demand did change. Far more hours spent at higher residual demand levels created larger average margins during 2000. This combined with the fact that marginal costs also nearly tripled between 1999 and 2000, which meant that similar percentage margins reflected much larger absolute dollar margins, producing extremely large wealth transfers.

VI. Rent Division and Deadweight Loss

Even without a market power analysis, it is clear that the extraordinary prices that began in the summer of 2000 created large transfers of wealth. The analysis we have carried out, however, allows us to parse the changes in wholesale payments for electricity between three mutually exclusive and exhaustive categories: changes in the competitive cost of generating electricity, changes in the level of competitive inframarginal rents (which would have occurred without any market power), and changes in seller rents due to the exercise of market power. Some of the rents due to market power became profits of electricity producers or marketers, but some were dissipated in production efficiency losses: efficiency losses resulting from the operation of higher-cost production units when a firm with lower-cost production exercises market power and restricts output.

A. Deadweight Loss

We begin by estimating the loss in economic efficiency due to the imperfections in the market. Since the demand for electricity in the California market was effectively perfectly inelastic with respect to the wholesale market price, efficiency losses would stem primarily from the inefficient allocation of production.[36] With asymmetric producers, there will be efficiency losses from the substitution of higher-cost production from price-taking firms, or even smaller strategic firms, for the lower-cost production of the larger firms that are exercising market power.

This would be a fairly straightforward calculation if there were no imports into the state. With no imports and perfectly inelastic demand, we could simply compare the efficient production costs of a given quantity of power, using the approach described in the previous section, with the actual production cost of that quantity of power. Due to imports that vary in quantity with the exercise of market power in state, however, we also need to account for the substitution of higher-cost imports for lower-cost in-state generation.

Thus, we divide the efficiency loss into these two components: the loss due to the inefficient production of the actual quantity served by fossil-fuel plants inside the ISO system, and the loss due to the increase in higher cost imports from the reduction of output by firms within the ISO system. Recall that the vast majority of power imported into California originates from regulated or publicly-owned firms, and that most of these firms have substantial native demand obligations. We have therefore assumed that the adjustment bids from these firms reflect the actual opportunity cost of their production, *i.e.,* that they are price-takers. When calculating wealth transfers, this is a conservative assumption. However, when calculating the impact of market power on efficiency losses, it is not. By assuming that import bids reflect the marginal cost of the supplier, we assume that increased production from these imports due to market power exercised by firms within California creates an increase in total production cost. If the adjustment bids from firms outside of California contain margins over their own marginal cost, this margin will be counted as an efficiency loss, rather than a transfer from consumers to those producers.

The two components of deadweight loss are illustrated in figure 2. A detailed description of these calculations is given in appendix C. The inefficiency from the reallocation of the actual production quantity among fossil-fuel resources inside the ISO is illustrated by

---

[36] This is not true if the exercise of market power caused some interruptible customers to drastically reduce their demand. This happened on 26 occasions during our sample (5 in 1998, 1 in 1999 and 20 in January-October 2000, but we have no way of estimating the deadweight loss from these demand reductions.

the solid gray area between the competitive marginal cost curve and the "actual" marginal cost curve just above it.[37] The total in-state fossil-fuel production inefficiency for June through September of 1998, 1999, & 2000 are shown in table 3. The expected cost from additional imports is shown in the striped area of figure 2 and reflects the difference between producing the quantity $\Delta q^{t}_{imp}(p_{comp})$ from imported production and producing that same quantity from instate production along the marginal cost curve MC_{comp}. Again, we have assumed that the adjustment bids of importing firms reflect their actual marginal production costs. The total production inefficiencies from higher-than-optimal imports for the summer months of our sample are shown in table 3.

In figure 5, we illustrate the relationship between our estimated instate productive inefficiency and residual system demand with the results of a kernel density regression of hourly inefficiency estimates and residual demand faced by the in-state fossil-fuel generators. Given our findings in the previous section, it is not surprising that we observe low levels of production inefficiency at low levels of system demand when there are low levels of market power. What may be surprising at first is that productive inefficiency peaks at moderately high levels of demand and then declines as demand nears system capacity even though our estimates of market power continue to increase. When the system is near capacity, however, very small output reductions can yield enormous price increases. Thus, while exercise of market power at those times may cause large wealth transfers, the resulting productive inefficiency is small because nearly all resources are running in any case. The relationship between total productive inefficiency, including imports, and total system demand is similar although the level at which efficiency losses begin to decline occurs at a relatively higher level of overall load.

B. Rent Division

With the calculation of deadweight loss due to productive inefficiency, we are now in a position to parse the total wholesale market payments into costs, competitive rents, and rents due to the exercise of market power. In figure 6, the quantity Q_{ISO} - Q_{mt} represents the amount of power traded in the wholesale market in a given hour, with quantities to

---

[37] This is a rough representation since the cost difference need not rise with the quantity of instate fossil-fuel production, as we discuss below.

the left of Q_{mt} being must-take power that is not compensated at the market price.

Total wholesale market payments are the sum of all the shaded areas. When the areas labeled *Mkt. Power Rents* and *Import Loss* are removed from the total, the result is the total wholesale payments that would have resulted if the market were perfectly competitive. The quantity Q_{res} represents hydro and geothermal production during the hour. For purpose of calculating the change in rents during summer 2000 and how those rents were divided, we assume that hydro and geothermal power has zero marginal cost, though this assumption has no effect on the conclusions

Under competition, the quantity Q^{Cal}_{comp} is produced by instate fossil-fuel units and the quantity $Q_{ISO} - Q^{Cal}_{comp}$ is imported. The area labeled *Comp. Total Cost* is the variable production costs of instate units and of imported power for their respective shares of production. Competition generates inframarginal rents equal to the sum of the areas labeled *Comp. Rents* 1 and *Comp. Rents* 2 for instate fossil-fuel producers, including reservoir resources, and the area *Comp. Rents* 3 for imports.[38] Together, these areas – *Comp. Rents* and *Comp. Total Cost* – account for all wholesale market payments under perfect competition.

With market power, the quantity Q^{Cal}_{actual} is produced by instate fossil-fuel units and the quantity $Q_{ISO} - Q^{Cal}_{actual}$ is imported. The areas labeled *Comp. Rents* 2 and *Import Loss* are the additional variable production costs of the imported power under the assumption that imports are bid competitively. In addition to the inframarginal rents represented by the area *Comp. Rents* 1, instate producers receive the area *Mkt. Power Rents* 1, but some of these rents are dissipated through inefficient production as described previously. Imports receive inframarginal rents *Comp. Rents* 3 as well as the area labeled *Mkt. Power Rents* 2. Together, these areas account for all wholesale market payments with market power.

Between the summers of 1998 and 2000, the wholesale market cost of power rose from $1.45 billion to nearly $8 billion. Efficient production costs more than tripled between these

---

[38] We have not calculated the import elasticity for prices below p_{comp} so we assume that marginal costs decrease in a linear fashion and that the marginal costs of imports are zero when the import quantity is zero. This does not impact the analysis of the change in rents.

Table 3: Production Costs and Rent Distribution (Millions $)
June - September

	1998	**1999**	**2000**
Total Actual Payments	1450	1495	7910
Total Competitive Payments	1018	1260	3567
Production Costs – actual	625	749	2355
Production Costs – competitive	582	691	1898
Competitive Rents	436	569	1669
Oligopoly Rents	389	180	3908
Oligopoly Inefficiency – instate	22	22	74
Oligopoly Inefficiency – imports	21	34	382

periods and with the marginal unit having higher costs, competitive rents for lower cost units also more than tripled. Oligopoly rents, however, increased by an order of magnitude, from nearly $400 million to close to $4 billion between these two summers. Thus while a substantial portion of the increased market cost of power was due to rising input costs and reduced imports, these factors also increased the dollar magnitude of the market power that was exercised by suppliers. As the results in the previous section indicate, the underlying competitive structure of the market does not appear to have changed substantially between 1998 and 2000. Rather the higher residual demand levels in 2000 created more frequent opportunities for in-state fossil-fuel producers to collect large margins on increased costs, leading to the 10-fold increase in oligopoly rents to suppliers.

The inefficiencies that resulted from the reallocation of production within California were much more modest, remaining at about 3% of total production costs through all three summers. The inefficiencies due to increased imports in power did grow substantially during our study, rising from 3% to over 16% of total production costs by the summer of 2000. To the extent that prices from importing firms did not reflect their actual production costs, but their own market power, this figure will include oligopoly rents earned by producers outside of California as well as actual productive inefficiencies.

VI. Conclusions

Restructuring of electricity industries has been predicated on the belief that competitive wholesale electricity markets can be attained. The debate over whether that assumption is correct and what must be done to ensure competition in electricity generation is ongoing. We have attempted here to make a reliable first estimate of whether and the degree to which California's wholesale electricity market has deviated from the competitive ideal.

Though a great deal of cost data are available for electricity generation units, we still had to make a number of assumptions in order to reach an estimate of the extent of market power in California. In most, though not all, cases, we have made assumptions that, if anything, are likely to produce results indicating less market power than actually exists.

The results indicate that market power in California's wholesale market was a significant factor during the summers of 1998 through 2000, though somewhat less so in 1999. These estimates should serve as a reminder that the problem of producer market power that was addressed in a purely regulatory framework for most of the 20th century has not completely disappeared with the recent restructuring. Our results demonstrate that market power is most commonly exercised during peak demand periods, which is not at all surprising given the current inability of wholesale demand to respond to high hourly spot prices. This underscores the importance of designing wholesale electricity markets that maximize the likelihood that wholesale price signals will be reflected in retail electricity rates.

These estimates demonstrate the degree to which prices exceed system marginal costs, the price level that would occur if all firms behaved as competitive price takers. We have not attempted to assess the profitability of any generation firms selling in California, since such profits are not necessarily an indication of market power, just as the absence of profits is not an indicator of competitive behavior. Under very favorable conditions for electric power supply, such as the high hydro conditions experienced over at least the first half of 1998, firms may have difficulty earning profits whether or not they are able to exercise market power. In all markets with durable assets, such as is the case in this industry, there are likely to be periods of high and low (or negative) profits regardless of the competitiveness of the market. Thus, the profits of generating companies in California during the time period we study provide little information about the competitiveness of this market.[39]

Finally, we want to emphasize again that this study was intended to develop an index of the extent and severity of market power. This is separable from the important debate over what index *levels* indicate a need for some form of market intervention. Years of electricity regulation confirmed the belief that government intervention can be costly and can result in tremendously inefficient production. The balancing of the costs and benefits of such intervention will require a great deal more study in this industry as restructuring proceeds.

---

[39] It is also worth noting that we have analyzed only the energy markets in California. Most generation units were eligible to earn additional revenues under reliability must-run contracts and from the sale of ancillary services.

References

Andrews, D.W.K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, **59**, pp. 817-858.

Bohn, R., Klevorick, A., and Stalon, C. (1999). "Second Report on Market Issues in the California Power Exchange Energy Markets." Available at www.calpx.com.

Borenstein, S. (2001). "Frequently Asked Questions about Implementing Real-Time Electricity Pricing in California for Summer 2001." mimeo, University of California Energy Institute, March.

Borenstein, S. (2002). "The Trouble with Electricity Markets: Understanding California's Restructuring Disaster." *Journal of Economic Perspectives*, forthcoming.

Borenstein, S. and Bushnell, J.B. (1999). "An Empirical Analysis of the Potential for Market Power in California's Electricity Market." *Journal of Industrial Economics*, **47**(September 1999).

Borenstein, S., Bushnell, J.B., and C. Knittel (1999)."Market Power in Electricity Markets: Beyond Concentration Measures," *The Energy Journal*, **20**(4,1999).

Borenstein, S., Bushnell, J.B., and S. Stoft (2000). "The Competitive Effects of Transmission Capacity in a Deregulated Electricity Market." *RAND Journal of Economics*, **31**(Summer 2000).

Borenstein, S., Bushnell, J.B., C. Knittel, and C. Wolfram (2001)."Trading Inefficiencies in California's Electricity Markets," POWER Working Paper PWP-086, University of California Energy Institute, October.

Bushnell, J.B. (2002). "A Mixed Complementarity Model of Hydro-Thermal Electricity Competition in the Western U.S.," *Operations Research.* Forthcoming.

Bushnell, J.B. (1999). "Transmission Rights and Market Power," *The Electricity Journal*, **12**(8), pp. 77-85.

Bushnell, J.B. and F.A. Wolak (1999). "Regulation and the Leverage of Local Market Power: Reliability Must-Run Contracts in the California Electricity Market." POWER Working Paper PWP-070, University of California Energy Institute.

California Energy Commission (CEC, 1995). 1994 Electricity Report. Data appendices. Sacramento, CA.

Cardell, J.B., C.C. Hitt, and W.W. Hogan (1997). "Market Power and Strategic Interaction in Electricity Networks." *Resources and Energy Economics.* **19**(1-2), pp. 109-138.

Green, R.J., and D.M. Newbery (1992). "Competition in the British Electricity Spot Market." *Journal of Political Economy.* **100**(5), pp. 929-953.

Harvey, S.M. and W.W. Hogan, "Issues in the Analysis of Market Power in California," mimeo, October 27, 2000.

Joskow, P.L. and J. Tirole (2000). "Transmission Rights and Market Power on Electric Power Networks." *RAND Journal of Economics*, Forthcoming.

Kahn, E., Bailey, S., and Pando, L. (1996), "Simulating Electricity Restructuring in California: Interactions with the Regional Market," *Resources and Energy Economics*, **19**(1-2).

Mansur, E. (2001). "Pricing Behavior in the Initial Summer of the Restructured PJM Wholesale Electricity Market," POWER Working Paper PWP-083, University of California Energy Institute, April.

Newey, W.K., K.D. West (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* . **55**, pp. 703-708.

Pando, L. (1995). Testimony and Workpapers in Southern California Gas Company CPUC Application No. 95-06-002.

Puller, S. (2001). "Pricing and Firm Conduct in California's Deregulated Electricity Market," POWER Working Paper PWP-080, University of California Energy Institute, July.

Schmalensee, R. and B. W. Golub (1985). "Estimating effective concentration in deregulated wholesale electricity markets," *RAND Journal of Economics.* **15**(1), pp. 12-26.

Wolak, F.A. and R.H. Patrick (1996). "The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales Electricity Market," POWER working paper PWP-047, University of California Energy Institute.

Wolak, F.A. (1999). "An Empirical Analysis of the Impact of Hedge Contracts on Bidding Behavior in a Competitive Electricity Market," mimeo, Stanford University.

Wolak, F.A., Nordhaus, R., and C. Shapiro (1998). "Preliminary Report on the Operation of the Ancillary Services Markets of the California Independent System Operator (ISO)." Avaliable at www.ucei.berkeley.edu/ucei/filings.html

Wolfram, C.D. (1998), "Strategic Bidding in a MultiUnit Auction: An Empirical Analysis of Bids to Supply Electricity in England and Wales," *RAND Journal of Economics*, **29**(4), pp. 703-725.

Wolfram, C.D. (1999), "Measuring Duopoly Power in the British Electricity Spot Market," *American Economic Review*, **89**(4), pp.805-826.

Appendix A: Data Sources

Fossil-Fuel Generation Data

Heat rates for fossil-fuel generation units that were not must-take and were located within the ISO control area are primarily taken from the California Energy Commission's dataset on WSCC generation for use with General Electric's MAPS multi-area production cost model. This is the dataset used in Borenstein and Bushnell (1999). Some unit heat rates were taken from the data set used by Southern California Gas Company in its 1995 performance-based ratemaking simulation studies (Pando, 1995). This dataset was also used by Kahn, Bailey and Pando (1996) in their simulation analysis of the WSCC. For NOx emission credit costs in the SCAQMD air basin, we used the quantity weighted monthly average price paid for emissions credit trades registered with SCAQMD. Emissions rates of generation units are taken from the Environmental Protection Agency's Continuous Emissions Monitoring (CEMS) data.

An overwhelming share of California fossil-fuel generation uses natural gas for the energy source. For the time period studied, we used weekly average natural gas spot prices reported by Natural Gas Intelligence at PG&E citygate and the California-Arizona border. The former were used for generation units north of path 15 while the latter were used for generation units in the south. Both sets of prices were adjusted by the distribution rates of the gas utility serving each generator.

A small number of California generators use either fuel oil or Jet fuel as their primary fuel. We use the Energy Information Administration's reported monthly average Los Angeles spot price for Jet fuel and no. 2 fuel oil.

Unit forced outage factors are taken from the National Electricity Reliability Council's (NERC) 1993-1997 Generating Unit Statistical Brochure, which reports aggregate generation unit performance data by fuel type and nameplate capacity. The forced outage factor that we used for our Monte-Carlo simulations were derived from the NERC reported unit Equivalent Availability Factors (EAF) and unit Scheduled Outage Factors (SOF). The former gives the fraction of total hours in which a generation unit was available, including an adjustment for partial outages, while the latter gives the fraction of hours in which each unit was unavailable due to scheduled maintenance procedures. Our derived forced outage factor, which reflects the fraction of time a unit was not available for production for unplanned reasons, was

$$FOF = 1 - \frac{EAF}{1 - SOF}.$$

Demand and Generation Output Data

Total ISO quantity for every hour is based upon the ISO's real-time metered generation and is taken from ISO settlement data. The output of must-take, hydro, and geothermal generation for each hour is also taken from these data. Imports are calculated from the net of real-time metered imports and exports aggregated over all transmission interties connecting the ISO's control area with neighboring control areas. The Mohave generation plant, although located outside of California, appears in metered data as must-take generating facility and not as an import. Production from all other generation units owned by SCE, but located outside of California, appear as imports in the settlement data.

Appendix B: Calculation of Standard Error for $\Delta TC(S)$

The observed PX price in hour i and day d, $\hat{p}^{id}_{px}$, depends on a single realization of the joint distribution of generation unit-level outages during that hour. A different realization of unit-level outages for that hour could result in a very different observed PX price for that hour. Because we do not know the precise nature of the competitive interaction among market participants, specifically the bidding strategies of both generators and loads which gave rise to the observed market-clearing PX price, we are unable to replicate the actual price-setting process for a large number of draws from the joint distribution of unit-level outages in order to compute the expected value of the PX price, $E(\hat{p}^{t}_{px})$. In contrast, we can compute the perfectly competitive market counter-factual equilibrium because it entails the assumption of marginal cost bidding by all in-state fossil generation unit-owners. This allows us to compute the realized value of the marginal cost of the highest cost unit operating in that hour for a large number of draws from the joint distribution of unit-level outages and therefore compute the expected value of this marginal cost to an arbitrarily degree of precision for each hour. We found that the 100 realizations from the joint distribution of unit-level outages led to a very precise estimates of this expected marginal cost for each hour. Consequently, the remaining source of uncertainty in $\Delta TC(S)$, the total cost difference due to deviations from competitive prices over time period S, that our standard error estimate accounts for is the uncertainty in PX prices caused by forced outages.

In constructing the standard error estimate for ΔTC, we account for arbitrary correlation in $[\hat{p}^{id} - \bar{C}^{id}]$ across the 24 hours of the day and general forms of autocorrelation in these 24 prices across days. We can write the total cost difference due to deviations from competitive prices over time period S as,

$$\Delta TC(S) \;=\; \sum_{d \in S} \sum_{i=1}^{24} [\hat{p}^{id}_{px} - \bar{C}^{id}][q^{id}_{tot} - q^{id}] \;=\; \sum_{d \in S} \sum_{i=1}^{24} mkup^{id} qn^{id},$$

where $qn^{id} \;=\; [q^{id}_{tot} - q^{id}]$ and $mkup^{id} \;=\; [\hat{p}^{id} - \bar{C}^{id}]$. This expression can be re-written as:

$$\Delta TC(S) \;=\; \sum_{d \in S} \sum_{i=1}^{24} E(mkup^{id}) qn^{id} \;+\; \sum_{d \in S} \sum_{i=1}^{24} \epsilon^{id} qn^{id},$$

where $\epsilon^{id} = mkup^{id} - E(mkup^{id})$ and $E(.)$ denotes the expectation taken with respect to the joint distribution of unit-level forced outages. Therefore the variance of $\Delta TC(S)$ can be written as:

$$Var(\Delta TC(S)) = Var(\sum_{d \in S} \sum_{i=1}^{24} \epsilon^{id} qn^{id}) = Var(\sum_{d \in S} Z_d),$$

where $Z_d = \sum_{i=1}^{24} \epsilon^{id} qn^{id}$. Under suitable regularity conditions on the sequence Z_d, for example those assumed in Newey and West (1987) or Andrews (1991), we can show that $(DAY(S))^{-1/2}(\sum_{d \in S} Z_d)$ converges in distribution to a N(0,V) random variable, as DAY(S) tends to infinity, where DAY(S) equals the number of days in time period S. A consistent estimate of V can be constructed as follows:

$$\hat{V} = g_Z(0) + 2 \sum_{\tau = 1}^{q} k(\tau/(q + 1)) g_Z(\tau),$$

where $g_Z(0) = 1/DAY(S) \sum_{d=1}^{DAY(S)} (Z_d)^2$, $g_Z(\tau) = 1/DAY(S) \sum_{d=\tau+1}^{DAY(S)} (Z_d Z_{d - \tau})$, and k(t) is a weight function satisfying restrictions given in Andrews (1991). Using this asymptotic distribution result, an estimate of the variance of $\Delta TC(S)$ is $(DAY(S)\hat{V})$, which implies a standard error of $(DAY(S)\hat{V})^{1/2}$.

To operationalize this procedure we need to construct an estimate of the unobservable, ϵ^{id}, the difference between $mkup^{id}$, and its expectation. As discussed above, in order to compute the exact expectation of $mkup^{id}$ we would need to know how all market participants bid into the PX (and other markets) as a function of current conditions in the market (system load and local reliability energy levels) and the realization of unit-level outages. Because we do not know the bidding strategies of even a single market participant and we do not observe actual generation unit-level outages during our sample period, a reduced form approach to construct the estimate of $E(mkup^{id})$ is necessary to compute an estimate of ϵ^{id}. We use a linear predictor of $mkup^{id}$ constructed using hour-of-day, day-of-the-week, and month-of-sample period dummies, along with the level and square of both the forecast ISO load and day-ahead total RMR requirements for that hour of the day. Our estimate of ϵ^{id} is the residual from the regression of $mkup^{id}$ on these variables for all hours in our sample period. For same reason that the squared difference between a random variable and its conditional expectation is always less than the squared difference between that random variable and its best linear predictor using those same conditioning variables, the variance of our estimate of ϵ^{id} should be larger than the true variance of ϵ^{id}. Consequently, we view our standard error as a very conservative estimate of the uncertainty in $\Delta TC(S)$ due to unobservable forced outages on realizations of the PX price.

Applying this procedure with the Barlett kernel, $k(s/(q+1)) = s/(q+1)$, for a value of q equal to 40 yields a standard error for $\Delta TC(S)$ for our entire sample period of \$741 million on the estimated $\Delta TC(S)$ of \$6.09 billion. For the June-September periods of 1998, 1999, and 2000, the estimates of $\Delta TC(S)$ and their standard errors (all in \$ million) are \$434 (\$92), \$235 (\$54), and \$4354 (\$476), respectively.

Figure 1: Instate Fossil-Fuel Generation Supply Curve (various months)

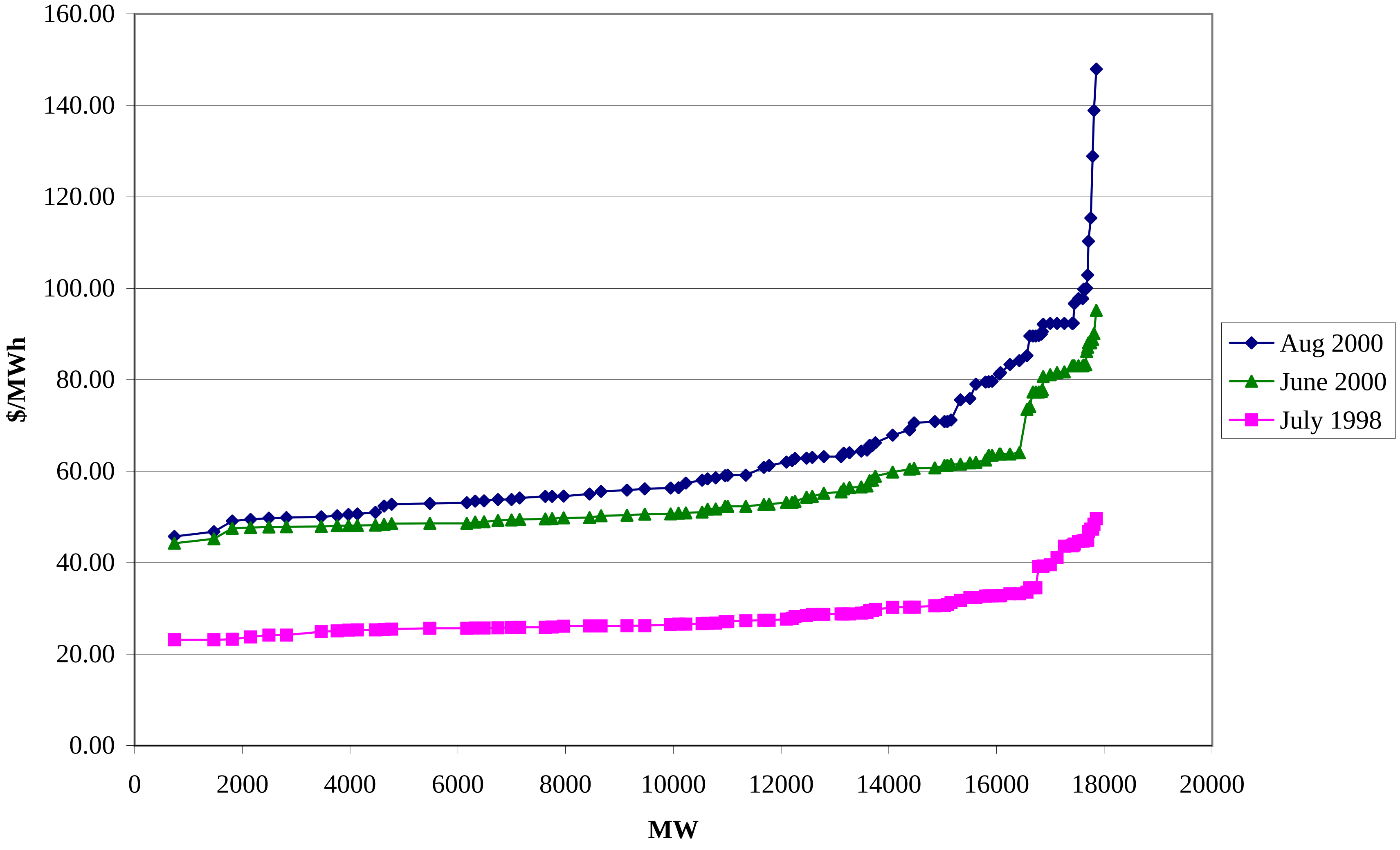

FIGURE 3

Figure 2: Import Adjustments and Efficiency Losses

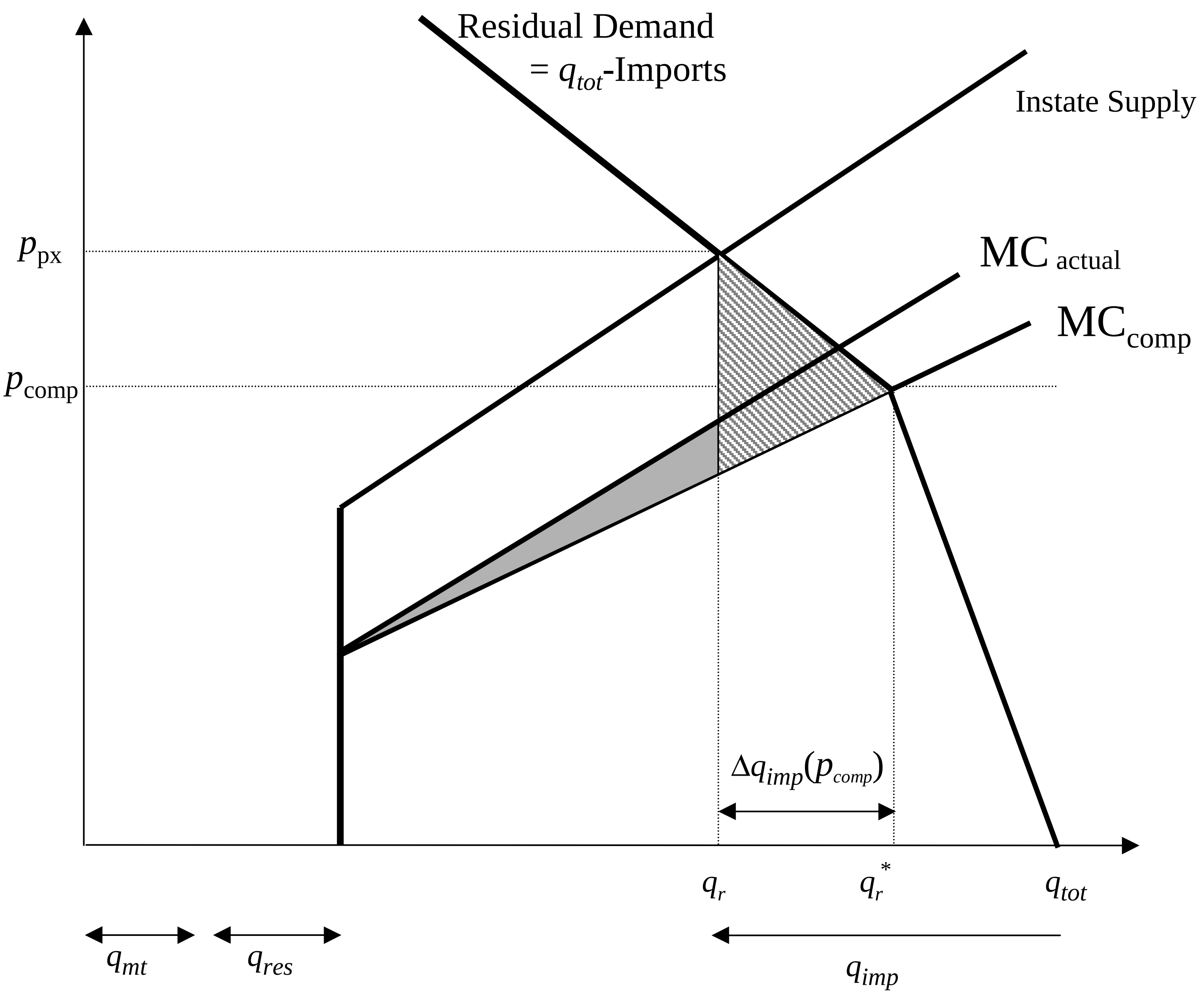

Figure 3: Kernel Regressions of Lerner Index for August & September

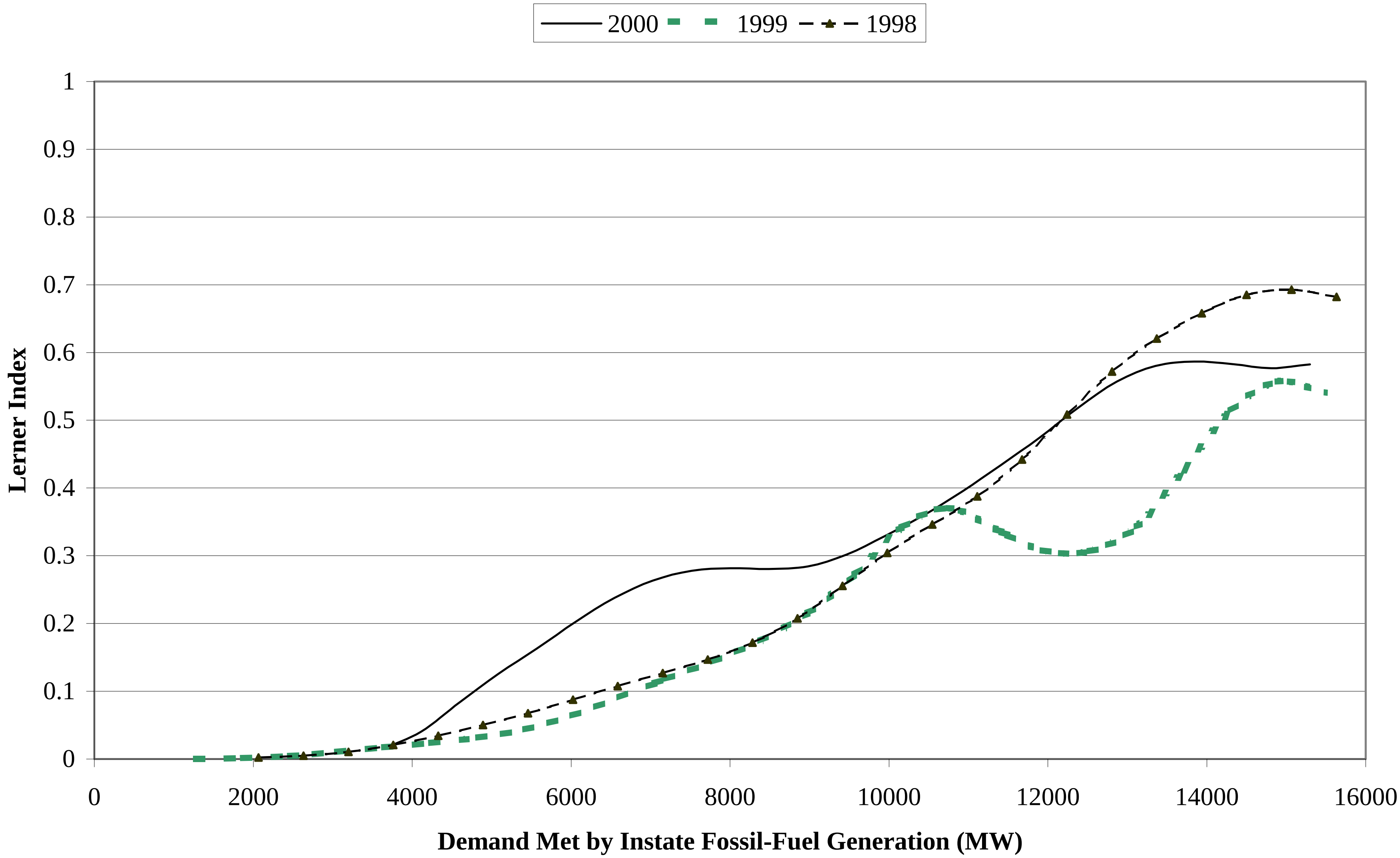

Figure 4: CDFs of demand for August & September

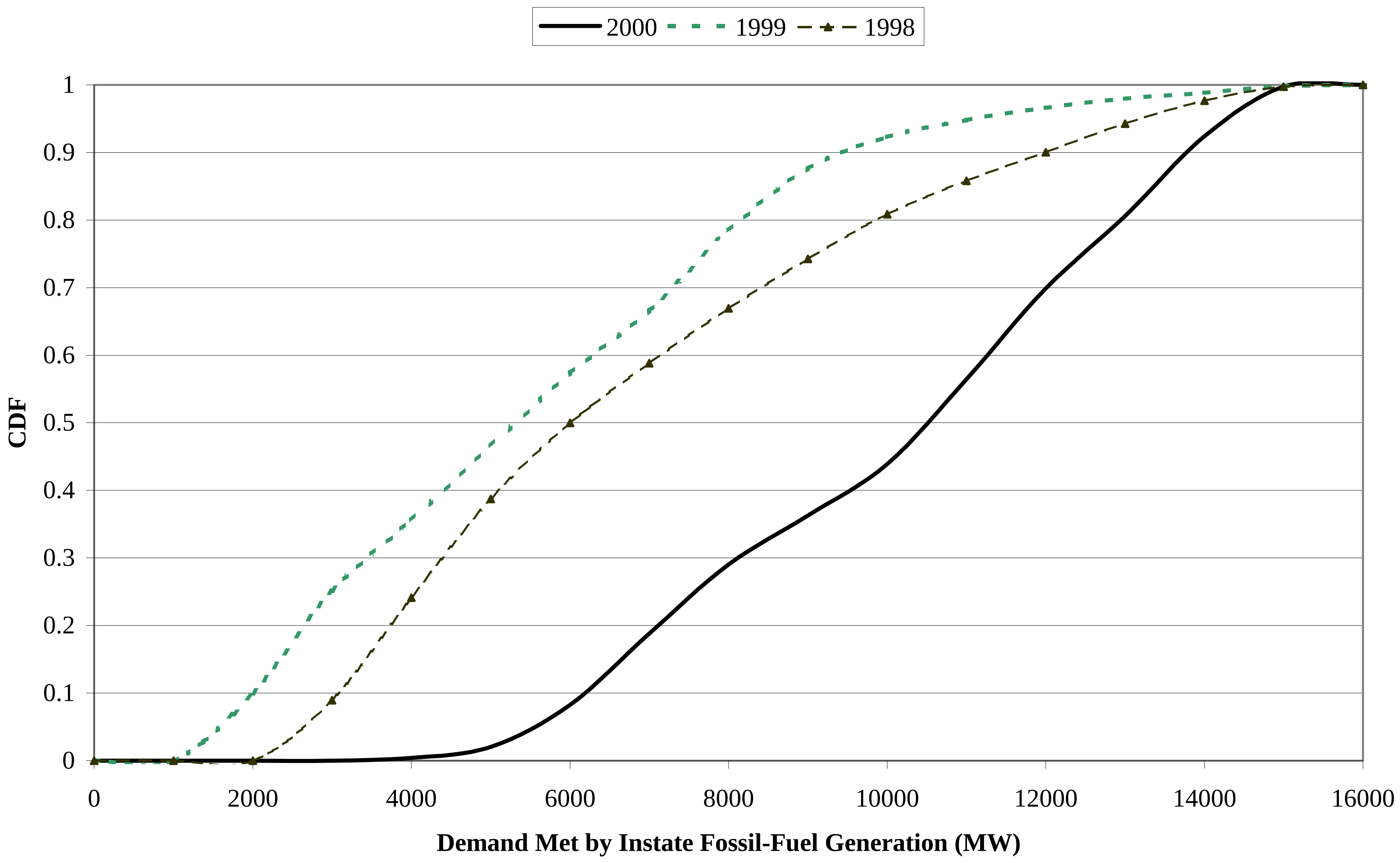

Figure 5: Efficiency Loss From Instate Fossil-Fueled Generation

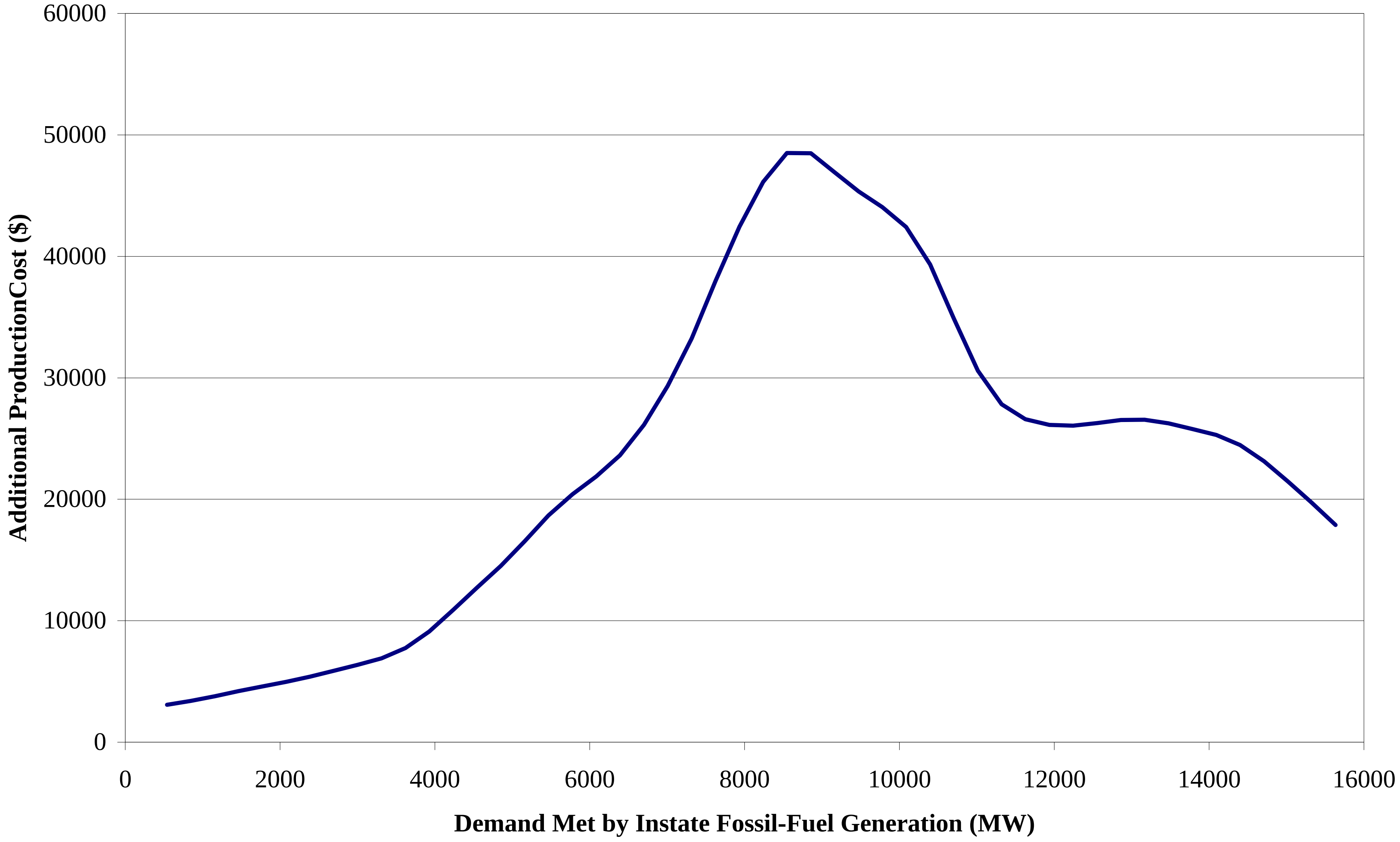

Figure 6: Calculation of Division of Rents