Identifying the Exercise of Market Power: Refining the Estimates

Paul Joskow* and Edward Kahn**

July 5, 2001

1. Introduction

Pricing behavior in the California wholesale electricity market has attracted considerable attention. We made an initial assessment of the evidence for firms' exercising market power seven months ago.[1] Scott Harvey and William Hogan have written an extensive response to our analysis.[2] Although Harvey & Hogan make some interesting and useful points, we find that their arguments as a whole are unpersuasive, that they are applied inconsistently, and that in sum they do not contribute to identifying or measuring when or how much market power is being exercised, the ostensible topic of their paper. In this paper, we respond to their more substantive criticisms and endeavor to reflect them in our revised analysis where relevant. We have also taken this opportunity to improve the data used in our analysis and reevaluate our treatment of ancillary services in calculating benchmark competitive prices. We have re-estimated the competitive benchmark prices to reflect these changes and have analyzed withholding behavior in more detail. Our primary conclusions are unchanged. Indeed, they have now been reinforced with more complete and accurate data and analysis: (a) actual market prices far exceeded the competitive benchmark during Summer 2000 and (b) there is now even stronger evidence of withholding behavior for three of the four generators in SP15.[3] The remaining generator, Duke Energy, had incentives to produce and behaved differently.

* Elizabeth and James Killian Professor of Economics and Management, Massachusetts Institute of Technology, Cambridge, MA.
** Analysis Group/Economics, San Francisco, CA.

We appreciate the comments of Harold Ray, Kevin Cini, Carl Silsbee and Gary Stern. Matt Barmack, Ryan Grant, Peter Griffes and Donna Lau provided excellent research assistance. This work was supported by Southern California Edison.

1 P. Joskow and E. Kahn, "A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market During Summer 2000," NBER Working Paper 8157, March 2001 (available at http://www.nber.org/papers/w8157); hereafter referred to as Joskow-Kahn. An earlier version of this analysis was filed on November 22, 2000 by Southern California Edison as Appendix A in its filing in FERC Docket EL-95-00. We made slight revisions subsequently, none of which involve quantitative issues.

2 S. Harvey and W. Hogan, "On the Exercise of Market Power Through Strategic Withholding in California," April 2001; hereafter referred to as Harvey-Hogan.

3 Throughout this paper we will use the term generator to refer to the agent controlling the output of plants. In reality, output decisions are made by marketers as well as generators. We note this point here, but for ease of exposition, will not make it repeatedly throughout the discussion.

The organization of this paper is as follows. Section 2 places this analysis in context with our previous work and Harvey & Hogan's response to it. Section 3 explores the issues Harvey & Hogan raise about our benchmark cost method. Section 4 is devoted to the numerical example that Harvey & Hogan present involving unit commitment issues. Section 5 takes up the physical withholding analysis.[4] We offer our conclusions in Section 6.

2. The State of the Debate

2.1 First Joskow-Kahn Analysis

In our initial paper, we employed uncontested economic principles to show that generators operating a portfolio of assets can increase their profits by withholding some of their capacity from the market when supply and demand are inelastic. We then devised two empirical methods using publicly available data to test whether generators were acting on this incentive during Summer 2000.

One method tested for the expected result of withholding behavior: higher prices. We therefore constructed competitive benchmark prices to predict what market prices would be in the absence of withholding behavior. We were well aware of this approach's shortcomings. First, because actual hourly prices depend on many specific and unforeseen factors, we can only use these benchmarks to estimate average prices. Events or conditions not reflected in the benchmark analysis can drive actual prices away from these estimates. Second, incomplete data and the complexity of the market forced us to make assumptions and simplifications which—to the extent that they are incorrect—will again skew our benchmark price in comparison to observed prices. We explained our assumptions and our reasons for them in our first paper. Despite these shortcomings, the benchmark prices give us a starting point for deciding whether actual prices are higher than expected in a competitive market. We found that actual prices in June, July, and August 2000 were higher than the benchmark prices by 90%, 56%, and 36%, respectively. These price discrepancies were consistent with the hypothesis that the exercise of market power (raising price above competitive levels and holding supply below competitive levels) had played an important role in determining market prices for those months. These qualitative results and associated conclusions were robust under a range of sensitivity cases involving NOx pollution reduction controls, NOx permit prices, and energy imports from out of state.

Our second method looked for direct evidence of capacity withholding. Withholding potentially profitable capacity can be profitable for a portfolio generator when the withholding increases profits from the generator's non-withheld capacity by more than the foregone profits from the withheld capacity. Withholding is more likely to be profitable during high-demand periods when most or all lower-cost generators are already

[4] We use the term "physical withholding" in a slightly different sense than Sheffrin, who is examining confidential bid data. We are looking at data only on hourly physical output, and refer to output levels below the competitive amount as physical withholding.

in use. We therefore examined summer hours in which the market price was higher than certain threshold levels. For each major generation owner, we calculated the difference between total generation capacity and actual generation supplied using two separate data sets. In both instances, we found a significant “output gap” which could not be explained by the need for ancillary services or by transmission congestion. This output gap can only be the result of either unusually high forced outages or withholding behavior. While it is impossible to prove that any given generating unit declared as a forced outage could have been available, the incentive to withhold is powerful and the observed behavior exceeds historic outage norms.

2.2 The Harvey-Hogan Critique

The Harvey-Hogan paper is philosophically dedicated to the proposition that market structure inefficiencies and the uncertainties they create for generators potentially explain all behavior that might otherwise look like the exercise of market power. Harvey & Hogan never address the fairly obvious theoretical proposition that under conditions of inelastic demand and supply, generators selling into a spot market with California’s market structure have substantial unilateral incentives and the ability to exercise market power; that is, they have the incentive and ability to exercise market power without any formal collusion between them.[5] Nor do Harvey & Hogan provide alternative estimates of benchmark competitive prices or empirical measures of market power. Rather than engage these basic principles of oligopoly theory, for example by presenting and applying an alternative theory of oligopoly behavior or alternative measures of market power, Harvey & Hogan present a litany of largely unsupported arguments that ignore what economic theory and common sense suggest about behavioral incentives. Their approach focuses on raising questions rather than providing answers. Their paper does not offer alternative estimates of competitive benchmark price, alternative measures of the gap between competitive prices and actual market prices, or alternative measures of withholding behavior. We find this philosophical approach to understanding and measuring market behavior and performance to be unproductive. Harvey & Hogan make their most useful contributions when they discuss the data that we have used in our study and its interpretation.

2.3 The Joskow-Kahn Update

Accordingly, we have updated and improved our analysis to respond to Harvey & Hogan’s questions and critiques. In the course of our review, we also improved the data used in our analysis.

2.3.1 Competitive Benchmark Price Analysis

Harvey & Hogan suggested that the market prices we relied upon in our competitive benchmark price analysis were too high and that the estimated benchmark prices were too low. If this were true, we would have overestimated the “price gap” indicative of market power. Specifically, Harvey & Hogan criticized our focus on day ahead prices. We have now examined the full range of wholesale prices: day ahead and real time energy prices,

[5] We are not saying collusion did not occur, since we cannot know that it did not, but only that collusion would not be necessary for market power to be exercised.

and ancillary services prices. We find that total market prices were actually higher than those we used in our original study.

Taking into account a number of Harvey & Hogan's points, we re-estimated our benchmark prices after (1) lowering our estimate of the effect of the demand for ancillary services on energy prices, (2) using unloaded hydro capacity to provide Regulation Up, (3) estimating competitive capacity prices of $2-$3/MWh for ancillary services, (4) allowing more hydro energy to be used to meet peak demand, (5) lowering firm capacities of wind turbines to reflect expected availabilities, (6) removing units which were unavailable during last summer, (7) including the full capacity of nuclear units (instead of derated capacity) to accord with actual Summer 2000 performance, (8) reducing net imports from out of state to match best available data, and (9) increasing the number of load points to capture convexity in the supply curve. Some of these revisions increase benchmark prices while others lower them. In aggregate, our new benchmark prices are higher in May and June and lower in July, August, and September than our old ones.

We have already indicated that the market prices we used previously were too low. Putting actual and benchmark prices together, our new estimates of the price gap are larger in every month. Accordingly, our conclusion is even firmer that prices during Summer 2000 were higher than could be sustained by competition alone. These estimates are summarized in Table 1. These estimates take new benchmark prices from Table 3 and new market prices from Table 4 below. The details are in Section 3.

Table 1. Summary of Changes in Benchmark Price Analysis ($/MWh)

	First Analysis			Revised Analysis		
Month	Market	Benchmark	Price Gap	Market	Benchmark[6]	Price Gap
May	47	47	0	61	59	2
June	120	63	57	167	77	90
July	106	68	38	118	60	58
August	166	122	44	180	81	99
September	115	104	11	126	83	43

2.3.2 Capacity Withholding Analysis

Harvey & Hogan raise a number of concerns with our withholding analysis as well, many of which overlap with their critique of our price benchmarking method. We take their criticisms with respect to the "reservoir" properties of certain generation resources and environmental constraints most seriously. Harvey & Hogan argue that these constraints were increasingly binding over the Summer of 2000. We agree. Therefore, we confine our withholding analysis in this review to the month of June, when these constraints are most likely to have been minimal because generators could not have known that

[6] The Revised Benchmark Price is calculated by adding our upper AS capacity price estimate of $3/MWh to the energy benchmark cost (100 load points) we estimate in Section 3.3 below.

unprecedented high prices would continue throughout the summer.[7] Harvey & Hogan make the same choice. We do not believe that the reservoir aspects of operating and environmental constraints make market power analysis impossible, only more difficult.[8] We thus defer for now a more extensive analysis of this issue and of the remaining summer months.

Our withholding analysis seeks to shed light on the ultimate question: Why did generators not operate units at times that would have been highly profitable for those units? The standard answer from economic theory is that generators are attempting to raise the market price and their own total profits by idling some of their units. Our initial tests showed that there was a significant amount of idle generation during high-price hours—enough to lend validity to the theoretical explanation of how generators exercise market power. Harvey & Hogan never directly address this theory or the evidence for it. Instead, they offer two main justifications for observed withholding.

First, Harvey & Hogan posit that some withheld units would not have been profitable at the times we examined. They provide indirect evidence for this hypothesis, namely that at least one unit that was not withheld was unprofitable on a run-cycle basis on one of the days in question. Taking into account start-up and minimum-load costs, which are not easily captured in hourly measures of cost, Harvey & Hogan show that Alamitos 2 was not profitable on June 17.[9] By design or chance, this instance turns out to be the only example of unprofitable operation by any generating unit over a full day in any day in our June sample: a 1 in 660 fluke.

A subtler version of Harvey & Hogan's run-cycle profitability argument invokes uncertainty: given that units sometimes must be committed hours before they are actually needed, generators' failure to accurately forecast demand may result in behavior that is observationally equivalent to withholding. If it is truly so difficult for generators to predict demand at the time that unit commitment decisions are made, we would expect to see many more cases where generators over-commit units and lose money on a run-cycle basis.

We address the Alamitos 2 example in Section 4 below.

---

[7] With perfect information about future prices, a generator will apportion a scarce resource (e.g., a gas turbine unit with annual emissions limits) across the highest-price hours in the entire year to maximize profits. In actual practice, the generator must estimate future prices and determine a threshold price above which it is willing to operate. In June 2000, without foreknowledge of the summer's events and using best available forecasts, most generators probably had a lower threshold than they did later in the summer as they increased their annual price estimates in response to revealed prices. We believe that this threshold is likely to be below our chosen cutoff of $120/MWh for most generators.

[8] A useful paper on this topic is Bushnell (1998), which shows what a market power strategy with reservoirs looks like.

[9] Harvey & Hogan are strangely inconsistent in their emphasis on the importance of the full range of market prices. In responding to our price benchmarking analysis, they emphasize that real time and ancillary services prices are important for understanding generators' incentives. However, Harvey & Hogan do not discuss these prices and revenue sources in their examination of Alamitos 2.

Harvey & Hogan also attribute output gaps to "outages."[10] In other words, they posit that engineering problems rather than economic or strategic considerations "forced" the units to be unavailable. Unfortunately, defining a true forced outage can be difficult. Our discussions with plant operators underline how the decision to declare a forced outage can be subjective. Although some engineering problems require immediate and total shutdown, plants frequently experience relatively minor operating problems. These problems may reduce efficiency or increase wear and tear without immediately affecting the plant's ability to produce power. The outage decision depends on when these problems accumulate sufficiently to pose larger risks to equipment and safety which outweigh the benefits of running the plant. But the benefits of running the plant include economic factors. Why would strategic considerations not be among them?

Because outage decisions are influenced by economic incentives, we believe that self-reported outages cannot be taken at face value. The FERC Staff Report on plant outages in California, cited by Harvey & Hogan is neither comprehensive nor definitive. Thus an important task is to analyze alleged forced outages. This task is not simple or easy. As Harvey & Hogan point out, the data necessary to perform such an analysis are not publicly available. In the absence of unbiased review of generators' operating decisions, our withholding analysis is designed to answer an indirect question: Is the amount of unavailable generation in the California market consistent with the amount we would expect to be unavailable for engineering and other known reasons? Our measures of the output gap show that these amounts are not consistent. As we have calculated it, the output gap shows the amount of unavailable generation above and beyond what would be expected if historical outage rates held true. Harvey & Hogan's points as well as common sense remind us that we cannot prove that any generating unit was withheld for strategic rather than engineering reasons. All of the outages *may* have been forced. However, the size of the output gap corroborates the theoretical expectation of strategic withholding behavior and places the burden of proving that outages were forced on the generators.

We find an important distinction among the generators in SP15, on whom our withholding analysis concentrates. Duke Energy, which appears to have been fully contracted in forward markets for 90% of its potential output, behaved much differently from Reliant, Dynegy and AES/Williams. Duke's production in SP15 was proportionally higher than that of these other firms. It reports much lower forced outage rates than what the other firms appear to claim. We believe that the outage rates and production levels reflect economic incentives. If generators are not contracted, their incentive is to withhold capacity and raise price.

[10] In our original report, we used the term "outage" to refer to any time a plant was not operating for any reason—i.e., as a synonym for unavailable generation. We recognize that others (including Harvey & Hogan) use outage more narrowly to refer to unavailability due to unforeseen engineering problems—i.e., as a synonym for forced outage. As the following discussion makes clear, we do not believe that outages can be so neatly labeled or circumscribed. We therefore remain agnostic with respect to the reasons for unavailable generation. We attempt to make the distinction between forced outages (for engineering reasons) and unavailable generation (for unknown reasons possibly including withholding behavior) clearer in this paper than in the previous one.

In addition, we document the inefficient dispatch that has resulted from withholding. We present suggestive evidence that the withholding of steam units and the consequent operation of GTs is consistent with efforts to manipulate both energy and NOx permit markets, at least partly to provide cost-based justifications for high prices.

Section 5 presents our revised withholding analysis.

3. Benchmark Costs and Prices

In this section we review a variety of issues associated with constructing a benchmark price. Harvey & Hogan voice a number of concerns with our benchmarking method, but as is frequently the case in their paper, the points raised are not reflected in any numerical estimates that can be compared to those contained in our report. Indeed, Harvey & Hogan do not quantify the effects of any of their criticisms in any comprehensive fashion. This makes it difficult to respond to their arguments. However, we try to quantify the issues we think are valid to the extent that it is possible to do so. This necessarily involves approximations and judgments. Without making such approximations and judgments, however, it is impossible to tell if the effects identified really matter. Section 3.1 discusses conceptual issues. Section 3.2 reviews data issues. Section 3.3 reports new results.

3.1 Conceptual Issues

We tried to be as fair and objective as possible in our original benchmark cost calculation. In retrospect, we feel that certain assumptions were too conservative, raising our benchmark prices unrealistically. Specifically, we have changed our treatment of ancillary services (AS) demand. In our previous estimate, we followed Hildebrandt (2000) by adding 10% to the observed demand to satisfy AS requirements. Harvey & Hogan agree with our current view that this is too high (pp. 16-17). We now add only 3% corresponding to Regulation Up to observed demand. We argue that unloaded hydro and extra-marginal peaking units are sufficient to meet the demand for other ancillary services. Hence, we should not incorporate the demand for those services in our calculation of the competitive energy price. In addition, we assert that, in a competitive market with abundant hydro capacity, the start-up costs of a GT are a rough upper bound on the price of capacity necessary to provide ancillary services. These start-up costs imply that reserves add $2-3/MWh to the cost of energy on a pro-rated basis.[11]

Our analysis adopts a de-rating approach to forced outages, meaning that the capacity we assign to all units is reduced by the assumed forced outage rate.[12] When we stack up resources against demand, we are assuming that these resources meet not only the final demand, but also the “outage loads” associated with the expected level of forced outages,

[11] This is a capacity price, or reservation fee. Energy is paid at the real time price if AS units are dispatched.

[12] We take values for these outage rates from the Henwood EMSS database, which is based on long-term historical averages. Some limited exceptions to this approach, facilitated by better data, are discussed in Section 3.2.4 below.

to the extent that outages occur on-the-day and are unknown ahead of time These forced outages are the reason that AS reserves other than Regulation exist in the first place. In our procedure, we are already including the effect of these "dispatched" AS reserves on energy prices. Therefore, we add only 3% to demand for Regulation Up. This is similar in spirit to the arguments made by Borenstein, Bushnell, and Wolak (BBW) for adjusting demands to reflect Regulation Up but not other ancillary services in this type of supply curve analysis.

Even this 3% adjustment may be too large. Both BBW and Harvey &Hogan neglect the fundamental role of hydro capacity both within the California ISO control area and the WSCC. The BBW argument about Regulation Up relies fundamentally on the same technology metaphor as the example that Harvey & Hogan offer. Harvey & Hogan show a hypothetical dispatch (p. 15, Table 3) in which inframarginal thermal generators are partially loaded to reserve unloaded capacity for "ramping." By reserving inframarginal capacity in this way, higher-cost resources get dispatched than would otherwise be the case. To model this effect, demand is increased by the amount of inframarginal capacity reserved. (Final demand does not change, but the increase in demand simulates the effect of reserves on energy market prices.)

We argue instead that most AS demand can and should be provided by unloaded hydro capacity. This is how the California system has operated historically, and there is no reason to believe that it should not operate this way under a market system. There is enough hydro in the California ISO control area to provide most of these services, both in reality and in our modeling of the energy market (see Section 3.2.1).[13] Other AS reserves can be supplied from resources both within the ISO area and outside of it. There are adequate resources in the region to meet these reserve needs. Resources available for reserves outside of the ISO control area include BPA, LADWP's Castaic pumped storage plant (1,200 MW), and out-of-state gas turbines. Since BPA has adequate capacity (if not energy) for the winter in the Northwest, then it should have had at least several thousand MW in the Summer 2000 to sell as AS reserves. Regulation Up is the only possible exception, because the ISO provides it separately for each zone from resources inside the ISO control area. SP15 may not have enough unloaded hydro to meet this requirement in high-load hours, so other (thermal) units may be required for Regulation at those times.[14]

This physical supply issue intersects with another question: how the competitive price for AS should be determined in concert with our estimate of energy prices. We suggest that the competitive price for ancillary services is the start-up cost of a gas turbine. When we revised our original benchmark cost calculation by removing AS demand from load, there was undispatched gas turbine capacity in California and the WSCC. Therefore we consider gas turbines marginal in the capacity market. The start-up costs of a GT is an

---

[13] We recognize that according to the *CAISO 2001 Summer Assessment*, not all of the unloaded capacity would be usable.

[14] This is a technology-based claim about how hydro has and should function in the energy and ancillary services markets. In other systems, where large amounts of hydro capacity (without energy behind it) are not available, then the approach based on unloaded thermal capacity is more sensible.

upper bound on the reserve capacity price because hydro will be sufficient to meet all AS demand in low-load periods. In high-load periods, particularly in SP15, these undispatched gas turbines may be needed to meet Regulation Up requirements.

The basic requirement for a unit to provide ancillary services is that it either be operating already or be induced to start up. The Henwood database suggests that start-up fuel costs for gas turbines are approximately $10/MW at gas prices of about $5/MMBtu. In addition, there are "wear and tear" non-fuel start-up costs associated with the acceleration of major overhauls. This is difficult to estimate. We have seen estimates that range from $10-20/MW.[15] If we assume that $20-30/MW is the AS capacity price, then when we allocate that over the total demand, the resulting adder to the energy price is $2-$3/MWh.[16] This is about what the ISO paid for AS in May 2000, which was a month where actual prices were close to benchmark competitive prices.

3.2 Data Issues

Harvey & Hogan raise a number of questions about how resources were counted in our competitive benchmark price estimates. We address these (and related) questions here.

3.2.1 Hydro Capacity

In our initial report, we used 8,000 MW of hydro capacity as the maximum available to serve energy demand in the ISO control area in any hour. Harvey & Hogan question whether this assumption is properly conservative (pp. 40-41). After reviewing the data, we believe that our previous estimate was *too* conservative and use 8,500 MW in our revised estimates.

There are 13,541 MW of hydro capacity in California (see Appendix 1). Some is owned or dedicated to Northern California municipal utilities which are in the CAISO control area. California entities with hydroelectric capacity outside the ISO's control area include LADWP, which has 1,764 MW of its own hydro; Pacificorp, which has 76 MW; and Imperial Irrigation District, which has 49 MW. This leaves 11,652 MW of hydro capacity in the ISO's control area. The CAISO (2001) lists 11,801 MW of hydro capacity available in the CAISO control area (p. 18, Figure II-H). The same document also estimates "Hydro Capacity Limitations" of 1,000 MW for Summer 2001 (p. 7). There is a subsequent discussion which suggests that these limitations were greater in 2000 (p. 8). Exactly what these are is unclear. However, the report includes a graphic (p. 9, Figure I-A) of the peak day dispatch for August 16, 2000, which shows approximately 8,500 MW of hydro dispatch at the maximum demand hour. Thus our previous use of 8,000 MW understates the actual capacity known to be used in the energy market in 2000. Accordingly, our revised analysis uses a maximum of 8,500 MW of hydroelectric capacity in any hour, even though total available capacity may be as high as 10,800 MW.

[15] RDI (1999) cites a $10,000 non-fuel cost per start-up for new peaking stations of 500 MW.

[16] We need 10% AS for each unit of dispatched capacity. This is Hildebrandt's (and our previous) estimate of the demand for AS. As argued above, only the demand for up regulation should affect energy prices. The capacity payment for AS is associated with all AS demands—i.e., 10% of load. Therefore, we allocate 10% of the $20-30/MW start-up cost to AS.

3.2.2 Interruptible Demand

Harvey & Hogan indicate in the context of the withholding analysis that we failed to include interruptible demand as a supply-side resource (p. 54). Had we done so, we would have estimated even lower competitive benchmark prices. There are about 800 MW of available interruptible demand that could be added to our supply stack.

2,980 MW of interruptible demand was available to utility distribution companies (UDC) during Summer 2000.[17] Interruptible demand, when called by the ISO, reduces observed demands. Table 2 shows ISO estimates of this effect for June through August 2000.[18] We can see that the maximum interruption was not invoked. Even at the highest use of interruptible demand on August 2, about 790 MW remained available. At other times even more was available. Because of issues associated with "energy limits," which are discussed below, we exclude this additional resource from our revised estimates.

Table 2. Effect of UDC Interruptible on ISO Measured Demand

Date	Forecast Peak (MW)	Actual Peak (MW)	Interruptible (MW)	Measured Peak (MW)
Jun-14	45,329	44,239	609	43,630
Jun-26	41,063	43,300	300	43,000
Jun-27	43,042	43,793	1,000	42,793
Jun-28	43,953	43,911	1,000	42,911
Jul-19	39,752	42,610	930	41,680
Jul-31	45,391	45,245	1,995	43,250
Aug-1	46,245	45,281	1,778	43,503
Aug-2	45,723	45,069	2,190	42,879
Aug-14	42,635	43,087	746	42,341
Aug-15	42,830	42,927	681	42,246
Aug-16	43,617	45,494	1,710	43,784

3.2.3 Energy Limits

Harvey & Hogan question whether our original analysis reflects constraints on energy-limited resources, such as hydro and geothermal resources. Our analysis accounted for energy-limited hydro by allocating observed hydro energy to high-demand hours. Although our original analysis did not say so explicitly, we also imposed energy limits on geothermal capacity in California by derating the capacity. This method does not account for what we would expect to be the disproportionate on-peak use of geothermal, so it overestimates benchmark prices. Energy limits also arise in the case of interruptible

[17] According to p. 10 of CAISO (2001), 2,800 MW were contractually available at the start of the year. The ISO also acquired 180 MW of interruptible demand in the Spring of 2000 according to CAISO (2000).

[18] See http://www.caiso.com/docs/09003a6080/08/8a/09003a6080088aa7.pdf. We calculated the Measured Peak by subtracting the Interruptible from the Actual Peak.

demand. This resource has a limited number of hours available—150 hours for the UDC contracts.[19] Because we did not address how these hours might have been used, we exclude them from our current analysis.

Perhaps most importantly, while Harvey & Hogan did not identify it as a problem, we have concluded that we did not treat wind resources from Qualifying Facilities (QF) correctly. The procedure used to create the supply curve from the Henwood database treated the QF wind capacity as if it were thermal capacity and only derated it by factors related to mechanical unavailability. The CAISO applies an 80% unavailability factor to account for the fickleness of wind power. Therefore, we now convert 1,876 MW maximum capacity into 375 MW of firm capacity.[20] Our original study counted 1,732 MW as firm capacity, or 1,357 MW more than we now include.

3.2.4 Outages and Availability

Harvey & Hogan assert that Hunters Point 2 & 3 were out for the entire Summer of 2000, and that Hunters Point 4 and the Riverside (formerly Highgrove) units were out in June 2000 (pp. 33-34). We included all of these in our original supply curve with derated capacity of 480 MW (326 MW for Hunters Point and 154 MW for Riverside). Riverside began to operate in late August. In our current analysis, we only include it in our supply stack for September. We remove Hunters Point 2 & 3 from the analysis for all summer months and Hunters Point 4 for June.

Our derating procedure underestimates supply if expected outages did not occur.[21] Although it is difficult to verify actual outages for many resources, plant-level monthly energy data are available from EIA Form 900.[22] We found that both Diablo Canyon and San Onofre ran at full capacity over the summer months. Accordingly, we place them in our supply curve at full capacity.

3.2.5 Net Imports

Our data review also revealed a problem with the net import data that we used in our original analysis. We relied on data posted on the University of California Energy Institute website purporting to represent real time flows into the ISO.[23] These data have been revised since our original assessment, and we have determined that they actually represent hour-ahead schedules. Additionally, we now understand that the output from SCE's share of Mohave is not counted in imports, so it must be included explicitly in our in-state supply stack. Finally, there are resources from outside of the ISO that are

---

[19] See CPUC Decision 01-01-056, January 2001.

[20] See CAISO (2001) pp. 13-15.

[21] Both BBW and Hildebrandt use actual hourly output for all must-take resources. This data is not available publicly. Since must-take resources are likely to be operated in a price-taking manner, it is appropriate to use actual production in their case. Strategic generators should not be treated in this fashion.

[22] EIA has proposed to stop reporting these and other data due to generators' confidentiality concerns. We use them also in Section 5.2. Given the importance of these data to analyzing market power, we oppose the EIA proposal and recommend continued reporting.

[23] See http://www.ucei.berkeley.edu/ucei/datamine/datamine.htm.

dispatched in real time. The real time dispatch of many of these resources is reported in the ISO's BEEP stack data. We add real time dispatched energy from these resources to hour-ahead scheduled net imports to construct our new measure of real time net imports.[24] The cumulative effect of these changes is to reduce net imports from our previous estimate by 600-700 MW/hr on average in June and July. The August estimate increases by about 170 MW/hr on average and the September estimate decreases by about 100 MW/hr on average.[25] We are still not counting Out-of-Market (OOM) imports, since prices and quantities for this category of supply are still treated as confidential by the ISO.

3.2.6 NOx Costs

Our unit commitment analysis in Section 5.3 suggests that observed RTC prices for NOx emissions are higher than they would be in a perfectly competitive market. If so, then we should lower the NOx prices used in our comparisons, which would increase our estimates of the price gap. However, since we are not able to quantify the strategic effect on RTC prices, we use the same estimates we did in our first assessment.

3.3 Re-Estimation

We have re-estimated our competitive benchmark prices incorporating all these changes and corrections. In those very few cases where our estimated supply curve has insufficient resources to meet demand, we invoke the price cap applicable to the month in question. In addition, we test for the effects of supply curve convexity that may have been neglected by our original load approximation. Harvey & Hogan discuss this issue at some length (pp. 36-40). We conduct two calculations with our supply curve, one with 10 load points per month (our original procedure) and one with 100 load points per month. The second calculation should capture some of the convexity effects identified by Harvey & Hogan.

These results are reported in Table 3. We find that there are indeed convexity effects on average. In every month the 100 point estimate produces higher prices than the 10 point estimate, though the "convexity effects" are small in comparison to the disparity between benchmark and actual prices. Nevertheless, after incorporating all of the other relevant corrections and adjustments discussed above, the average monthly benchmark price is about $12/MWh lower than our previous estimate. The new estimates for May and June are higher, mostly due to the price cap's effects in a few periods.[26] The new estimates for

[24] BEEP is the CAISO's real time Balancing Energy and Ex-Post Pricing software. Data from BEEP are publicly available through the ISO's OASIS website. Hildebrandt (2001, p. 21) constructs a measure of real time net imports similar to ours using BEEP data and hour-ahead schedules. In theory, it should be possible to check this measure of net imports against data on real time flows. Unfortunately, the ISO no longer posts data on real time flows.

[25] When we compare the 1999 to 2000 changes in net imports, the decline in 2000 is between 1,000 and 1,600 MW/hr less than estimated in Table 3 of our original assessment.

[26] The June results for the 100 load point estimate are somewhat counter-intuitive, with benchmark prices first rising as RTC prices go from $0/lb to $10/lb and then falling as these prices go from $10/lb up to $30/lb. The lower price at high RTC costs is due to the import elasticity effect, which brings in more resources as in-state fossil costs increase. These increased imports then reduce the effect of the $750/MWh price cap as well.

July, August, and September are lower. In these months, our new estimates have more net supply than previously. The contribution of RTC costs in SCAQMD declines compared to our previous estimates.

Table 3. Summary of Benchmark Costs

Month	Average PX Price ($/MWh)	Average MC ($/MWh) Assumed NOx Price $0/lb	$10/lb	$20/lb	$30/lb	$35/lb
		100 load points				
May	47.23	**55.78**	59.38	60.54	62.20	62.85
June	120.20	70.81	**74.03**	71.65	62.66	64.16
July	105.72	56.23	54.80	**56.69**	58.26	59.04
August	166.24	64.88	69.58	72.75	75.59	**78.13**
September	114.87	68.96	72.47	75.35	78.26	**79.67**
		10 load points				
May	47.23	**44.01**	45.81	48.49	50.37	51.34
June	120.20	52.01	**53.98**	56.70	58.04	59.24
July	105.72	48.31	50.30	**51.91**	53.90	54.38
August	166.24	60.86	64.46	68.47	72.79	**74.74**
September	114.87	66.41	70.77	74.06	76.58	**77.51**

Table 4 summarizes our new market price estimates. The table includes prices for Ancillary Services (AS) and an aggregate of day ahead and real time energy compiled by the ISO's Department of Market Analysis (DMA). It shows that a more comprehensive measure of market prices—reflecting all energy and AS purchases—was generally higher than the measure that we used in our original study.[27]

Table 4. Measures of Average Market Prices ($/MWh)

Month	PX Price	ISO DMA AS	ISO DMA Energy	ISO DMA Energy + AS
May	47.23	3.16	58	61
June	120.20	20.19	147	167
July	105.72	5.71	112	118
August	166.24	12.18	168	180
September	114.87	7.39	119	126

Using this new measure of market price, the gap we measured originally between actual prices and competitive benchmark prices increases in every period (see Table 1). Even if we retain our old market prices, the price gap remains large in June and increases in July, August, and September. The price gap in September, though marginally positive in our original assessment, now seems large enough under either measure to warrant attention.

[27] From ISO FERC Filing, Attachment A, "Energy Cost per MWh," "AS Costs - $ per MW Load" and "Total Costs per MWh."

4. Profitability: The Alamitos 2 Example

Harvey & Hogan use only one genuinely new numerical example to support their many allegations.[28] Its purpose is to show that generating units may have been idle during the high-price hours we examined not because they were withheld to drive up prices, but rather because they were unprofitable to run once start-up costs and other factors are taken into account. They provide indirect evidence for this hypothesis, namely that at least one unit that was not withheld was unprofitable on a run-cycle basis on one of the days in question. Harvey & Hogan find that the Alamitos 2 unit lost money on June 17. They imply we would have found many more such examples had we analyzed unit profitability in more detail.

We have now conducted that more complete analysis for the high-price days in June 2000. Far from leading us to change our conclusions, our study shows that the Alamitos 2 example is anomalous and does not support any inference that "unprofitability" explains the withholding behavior we identified. Harvey & Hogan focus on the *one* instance of negative run-cycle profitability over a full day of operation for the units and days examined by our original withholding analysis.

We address Harvey & Hogan's Alamitos 2 example in Section 4.1. In Section 4.2, we make our run-cycle profitability calculations more comprehensive by including NOx costs and ancillary services revenues.

4.1 A Rare Day in June

Harvey & Hogan focus on the month of June, in which average day ahead energy prices were \$120/MWh. This average masks considerable variation. There were 137 hours when PX prices were above \$120 and 583 hours when they were below \$120. We also used \$120 as the cutoff for defining high-price days in our withholding analysis. Table 5 shows the 14 days in which these hours occurred, the average price for those days, and the number of hours when the price was above \$120. We also show the SP15 average real time prices on these days and the number of hours above \$120.

[28] Harvey & Hogan conduct some variations on the Joskow-Kahn withholding analysis, which we discuss separately in Section 5.

Table 5. Days with Hourly Prices Above $120/MWh

Date	Average PX Price ($/MWh)	Number of Hours > $120	Average Real Time SP15 Price ($/MWh)	Number of Hours > $120
Jun-12	76	5	71	3
Jun-13	87	6	255	11
Jun-14	263	16	364	15
Jun-15	331	15	164	8
Jun-16	254	15	56	1
Jun-17	72	1	53	0
Jun-21	63	1	217	8
Jun-22	100	7	119	8
Jun-23	93	6	61	1
Jun-26	123	11	296	12
Jun-27	228	12	391	15
Jun-28	326	14	378	17
Jun-29	365	15	260	12
Jun-30	333	13	69	4

Harvey & Hogan examine the operation of the Alamitos 2 unit on June 17[29] (pp. 25-33). Their analysis shows that on that day it was unprofitable, according to the criteria they invoke, to run that specific generating unit. They state that "These calculations for Alamitos 2 are only illustrative and we have not repeated this calculation for every unit for every day of June" (p. 32). Readers may gain the impression that the Alamitos 2 example is typical of other units on other days and that "unprofitability" is an important explanation for the withholding behavior that we identified. However, Harvey & Hogan do not actually display similar calculations for any other units or days. We have now performed profitability calculations for every day listed in Table 5: the days which have the high-price hours that we focus on in our analysis of withholding behavior. This task was not unduly onerous, and Harvey & Hogan could have performed it had they wished. We find that there was only *one* unit on *one* day in June which was unprofitable under the Harvey-Hogan criteria using hourly CEMS data to account for minimum-load costs.[30] This unit is Alamitos 2 on June 17, which Harvey & Hogan happened to choose as their example.[31]

There are 660 merchant unit/days in CEMS on the high-price days in question. It is truly remarkable that Harvey & Hogan stumbled onto the one time when a unit was

[29] June 17th was a Saturday, following three days of very high prices. It is also interesting to note that the 17th had the lowest real time prices of all 14 high-price days.

[30] These criteria are CEMS heat rates and hourly production, zonal day ahead prices, no ancillary services revenue, and Joskow-Kahn gas prices of $4.59/MMBtu in the North and $4.99/MMBtu in the South (see Joskow-Kahn Table 2). CEMS is the Environmental Protection Agency's Continuous Emissions Monitoring System. We discuss some particularities of this data set in Joskow-Kahn Section 6.

[31] As a check to our work, our profit calculation for Alamitos 2 (see Table 6) is within $3 or 0.2% of Harvey-Hogan Table 7 "Total All Hours."

unprofitable to run and then discussed it in detail, leaving the reader with the impression that the example had been chosen at random and that an exhaustive analysis would reveal many other cases of unprofitability.

Since Harvey & Hogan also claim that generators' uncertainty about profits could lead to inefficient withholding, we note that there were very few "near misses" with respect to unit profitability. Table 6 shows all of the very small number of instances in our sample when units made profits below $25,000 over the course of a 24-hour day.[32]

Table 6. Units with Low Daily Margins on High-Price Days

Unit	Date	Profits ($)
Alamitos 2	Jun-17	-1,721
Alamitos 1	Jun-21	2,340
Alamitos 1	Jun-23	2,513
Redondo Beach 6	Jun-21	5,742
Huntington Beach 2	Jun-17	12,065
Alamitos 2	Jun-21	12,174
Coolwater 2	Jun-17	23,910

It is notable that six of the seven cases occurred on either June 17 or June 21. Table 5 shows these were days when the PX unconstrained price was above $120/MWh for only one hour. No other day had fewer than five high-price hours. Consequently, June 17 and June 21 are substantially less important to our withholding analysis than the other 12 days in Table 5.[33] Even if we do not dismiss unit profitability as a red herring, it is only a realistic concern for two of the 137 hours in our withholding analysis.

We accept as valid and helpful Harvey & Hogan's use of CEMS heat rates to calculate profits instead of the incremental heat rates we used in our first study. In the course of our work, however, we found some interesting comparisons between CEMS average heat rates and the incremental heat rates in Klein (1998). We explain these issues and examine other units besides Alamitos 2 in Appendix 3. Both data sets lead to the same qualitative results and the conclusion that profitability is not a valid cause to doubt our withholding analysis.

4.2 Other Costs and Revenues

Harvey & Hogan expand their analysis to include NOx credit costs as an additional component of profits (p. 32, Table 9). We consider including NOx credit costs in a unit profitability analysis to be legitimate. However, it is also appropriate to include other

---

[32] There were a few other units on some of these days that had brief unprofitable periods. These units were either starting up for the next day or shutting down from the previous day. They were all quite profitable on the full days when they operated, and these profits offset whatever losses occurred in the start-up or shut-down periods.

[33] Compared to June 21st (the other day with only one high-price hour), June 17th had more South to North congestion and much lower ancillary services revenues. Both of these effects would reduce profitability of SP15 generators on the 17th compared to the 21st.

sources of revenue. In this section we consider two additional sources of revenue, real time energy revenues and ancillary services revenue, as well as NOx costs. We then re-estimate profits for the low-margin plants (Table 6) using this fuller picture of cost and revenue streams.

4.2.1 Real Time Energy Revenues

Harvey & Hogan emphasize the importance of using real time prices in connection with the CEMS data because CEMS output includes real time output. They focus only on hours when real time prices are low compared to the day ahead prices, perhaps leading to the inference that this is always the case. As Table 4 above indicates, the opposite also occurs. June 21, for example, was a day with real time prices substantially above day ahead prices. The ISO made very large purchases during peak hours. We have not included these revenues in Table 7 below because they would not change the qualitative results.

4.2.2 Ancillary Services Revenues

Absent the availability of confidential data, there is no definitive way to measure AS revenues by unit. This does not mean that plausible approximations are impossible. We construct such an approximation for generators in SP15 and give the details in Appendix 2. We estimate each generator's capacity available to provide each service, then allocate the realized revenues earned in the region in proportion to that available capacity. As Table 7 shows, these revenues can be substantial.

4.2.3 RTC Costs

We used a June NOx cost of $10/lb in our first study. This price may understate profits, as Sheffrin cites a quote from Cantor Fitzgerald for June of $7.50. Harvey & Hogan allow prices between $5 and $20, but the high end here is hard to believe. For consistency's sake, we use $10/lb in Table 7.

4.2.4 Results

We summarize in Table 7 the effects that NOx costs and AS revenues have on the profits of "near miss" cases. [34] Alamitos 2 remains unprofitable on June 17, but we confirm our conclusion that this case is abnormal. All of the other units remain profitable. For four of the other five cases, profits increase because AS revenues outweigh NOx costs.

Table 7. Profits of Low-Margin Units with NOx Costs and AS Revenues

Unit	Date	PX Profits [1]	NOx Costs @$10/lb [2]	AS Revenues [3]	Total Profits [1]+[3]-[2]
Alamitos 2	Jun-17	-1,721	34,657	15,139	-21,239
Alamitos 1	Jun-21	2,340	20,044	46,209	28,505
Alamitos 1	Jun-23	2,513	15,271	99,982	87,224
Redondo Beach 6	Jun-21	5,742	23,974	23,832	5,600
Huntington Beach 2	Jun-17	12,065	10,405	18,935	20,595
Alamitos 2	Jun-21	12,174	38,016	43,216	17,374

[34] We omit Coolwater 2 from this analysis since it is not in SCAQMD and therefore has no NOx costs.

This entire discussion assumes that unit profitability is the sole criterion by which the behavior of a portfolio generator should be measured. We know from previous work on market power in other electricity markets that portfolio generators do not treat all units identically. Wolfram (1998) is the best-known study of this type. We have not investigated portfolio effects in this discussion.

4.3 Conclusion

Harvey & Hogan's profitability analysis is incomplete and conveys the misleading impression that capacity withholding can be explained by profitability concerns. If the behavior of the generators were truly affected by uncertainty as Harvey & Hogan continually argue, then there should be many examples of unprofitability. Under competitive conditions, generators should be expected to make commitment errors that turn out to be unprofitable after the fact. The remarkable record of profitability we have just reviewed shows instead that the generators were able to triumph over the uncertainties that allegedly plague them. They were either very smart, or able to influence prices by their behavior, or most likely, both.

To illustrate how prices could be driven to their high-observed levels, we now return to our analysis of withholding.

5. Withholding Analysis

In this section we return to the withholding issue by reevaluating the data we used previously for the month of June.[35] We demonstrate that our previous measures of withholding are robust if we select hours for our analysis based on real time rather than day ahead prices. We correct a few minor issues raised by Harvey & Hogan regarding ISO pricing zones and measures of capacity. We measure dispatched AS capacity to eliminate it as an explanation of the output gap. We conduct the analysis on a chronological basis and sharpen our analysis of previously excluded units. We find that generators withheld capacity far in excess of what can be explained by historical outage rates or demand for ancillary services. High-cost and high-emissions units ran while more efficient units remained idle. The output gap was especially large on the highest-demand days. These results are consistent with the withholding behavior for strategic rather than engineering reasons that are expected given generators' incentives. Our previous conclusions are reinforced.

5.1 Conceptual Issues

5.1.1 Real Time Prices

Before proceeding with our analysis, we take to heart Harvey & Hogan's warning that using real time prices instead of day ahead prices may lead to a different picture of the

[35] June is a good month to examine because we have two sources of generator output data and Harvey & Hogan's arguments about "opportunity costs" and the implications of "20/20 hindsight" are most unlikely to be of any relevance.

output gap. In our first study, we looked at hours when PX unconstrained prices were above $120/MWh. We now repeat our calculations of the output gap compared to ancillary services requirements using real time prices. Table 8 compares to Table 8 in our initial assessment.[36] Using real time prices but changing no other data inputs, the output gap in SP15 increases from 3,351 MW to 3,602 MW. In NP15, the gap drops from 983 MW to 895 MW.[37]

Table 8. Mean Level of the Output Gap: Real Time Prices > $120/MWh

Zone	Owner	Mean Output (MWh)	Max Output (MWh)	Mean Output Gap (MWh)	Mean AS Demand (MWh)	Mean AS Net of Replacement (MWh)
NP15	Duke	1,414	1,526	112		
	Mirant	1,936	2,719	783		
	NP15 Total	**3,350**	**4,245**	**895**	**1,553**	**1,137**
SF	Mirant	155	213	58		
	SF Total	**155**	**213**	**58**	**46**	**29**
SP15	AES/Williams	2,482	3,681	1,199		
	Duke	594	733	139		
	Dynegy	1,002	2,000	998		
	Reliant	2,221	3,487	1,266		
	SP15 Total	**6,299**	**9,901**	**3,602**	**1,606**	**986**
ZP26	Duke	950	1,037	87		
	ZP26 Total	**950**	**1,037**	**87**	**67**	**34**

As a further comparison of the two sets of prices, we also recalculated our "Test 1" for SP15 using real time prices. (Test 1 measured maximum output for a given hour by including only units which were partially loaded in that hour. It was our most restrictive test, leading to the lowest estimate of the output gap.) The resulting Table 9 compares to Table 10 in our initial assessment. The output gap increases from 1,954 MW to 1,997 MW.

[36] We now list the company formerly known as Southern Energy by its new name, Mirant.

[37] For each unit and hour, EPA's CEMS data report average load during operation and the fraction of the hour that the unit ran. Output can be calculated as the product of these two numbers. In our previous analysis and in the tables that follow, our reported output figures represent average load during operation rather than actual output. This measure overstates true output and hence understates output gaps. Harvey & Hogan appear to have followed this convention. In practice, CEMS contains data for very few fractional hours of operation. Our results would not substantively change, and in many cases would not numerically change, if we reported figures for actual output rather than average load.

Table 9. SP15 Test 1 on Real Time Prices > $120/MWh

Owner	Mean Output	Max Output	Mean Output Gap	AS Total	AS Net of Replacement
AES/Williams	2,482	2,962	480		
Duke	594	687	93		
Dynegy	1,002	1,541	539		
Reliant	2,221	3,106	885		
SP15 Total	**6,299**	**8,296**	**1,997**	**1,606**	**986**

Relying on real time prices rather than day ahead prices would not cause us to change our conclusions from our first study. In SP15, the real time prices make the output gap larger. In NP15, although the output gap decreases slightly with the switch to real time prices, we did not infer market power in our first assessment because the output gap was already lower than ancillary services demand.

The choice of relevant prices also affects the more disaggregated withholding analysis that follows. Harvey & Hogan claim that the only relevant hours to consider are those when *both* day ahead and real time prices are above a threshold level.[38] Since generators can make both start-up and dispatch decisions up until real time, we think that there are equally compelling arguments for examining hours in which only the real time price exceeds some threshold or—if at least some irreversible operational decisions are made day ahead—hours in which either the day ahead or the real time price is above some threshold. In what follows, we examine hours which meet this second criterion.

5.1.2 Geographic Locations

Table 8 confirms yet again that nothing conclusive can be said about NP15. Because the demand for ancillary services exceeds the output gap, it is possible that all unused capacity was reserved for ancillary services. We conjecture, however, given the argument in Section 3.1 above, that hydro supplied the vast majority of the AS in NP15. If this conjecture is correct, then Mirant's behavior looks like the kind of withholding that we examine more intensively in SP15. With publicly available data it is impossible to verify or refute our conjecture. In our current study, we concentrate exclusively on SP15, where our previous analysis suggested that the exercise of market power is more prevalent.

5.1.3 Ancillary Services

In our previous analysis, we used tests that compared the aggregate zonal output gap to the zonal the demand for AS capacity. We considered the possibility that Replacement Reserve was dispatched, in which case its generation would appear in the production data and could not explain any remaining output gap. Further investigation has confirmed anecdotal evidence that Reserves—not only Replacement but Spin and Non-Spin as well—were often dispatched last summer. Therefore, in order to measure the output gap net of AS, we now include only undispatched AS. We compute hourly dispatched Replacement, Spin, and Non-Spin from the BEEP stack data. In June, the dispatched AS

[38] The threshold levels chosen in our original analysis were somewhat arbitrary and based on a misinterpretation of some ISO analysis. See Joskow-Kahn, footnote 27 (p. 22).

in SP15 accounts for 600 MW on average during the high-price hours that we originally considered (and about the same in Tables 8 and 9). Harvey & Hogan appear to accept that Replacement Reserve was dispatched, but continue the unfortunate practice of focusing mainly on the total AS in each zone and fail to include dispatched AS capacity in their withholding analysis.

5.1.4 "Outages"

Our previous discussion of the output gaps introduced some unfortunate terminology that has confused subsequent discussion. In hindsight, our use of the word "outage" to describe unavailable generation was misleading. Our study did not examine any data regarding the physical availability of generating units. We simply computed different estimates of the output gap depending on what units were operating at particular times. By referring to capacity that was not operating at particular times as being "out," we appear to have given the mistaken impression that these units were not available to supply due to mechanical or other technical problems. This choice of words has been adopted by Harvey & Hogan and conflated in their discussion with the kinds of mechanical problems described in the FERC Staff's "Report on Plant Outages in the State of California."

Contrary to Harvey & Hogan's assertions, or any inferences that they may have drawn from the FERC Staff Report, with current information we don't know what judgmental factors went into the decision to declare capacity unavailable. Because of the size of the output gap, we infer that some of the "unavailable generation" could have produced power and was withheld for economic reasons, not because of mechanical failures. We are in no position to claim this of any particular unit, but our argument is based on probability and the big picture in the market. Detailed operational data not available to the public would allow us to test our conjectures at the unit level.

In this study, we examine the 14 high-priced days in June more carefully to identify commitment and dispatch behavior and to see if it is consistent with what we now know about real "forced" outages. In Section 5.3, we begin by focusing on the decisions made by generators to start up or turn off generating units, usually referred to as unit commitment decisions.

5.1.5 Chronology

Finally, we agree with the spirit of chronological analysis that Harvey & Hogan introduced with their Alamitos 2 example. We present a more complete version of our previous findings in this style. We believe that it helps to clarify the issues, allows a more accurate representation of the withholding behavior, and allows a finer level of analysis.

5.2 Data Sources

5.2.1 Output Data

Harvey & Hogan rely exclusively on the CEMS data. They completely ignore our analysis which relies on the WSCC Extra High Voltage (EHV) database. These data are available to the sponsors of the Harvey & Hogan study, and therefore they could have replicated our calculations. The WSCC EHV data include units not in the CEMS data.

Admittedly, because of aggregation at the plant level, the WSCC EHV data are not amenable to some tests that are possible with CEMS data. They are not therefore irrelevant. Indeed, our original analysis shows that the evidence for withholding is much greater in this broader setting, even if it cannot be subjected to every test that Harvey & Hogan would like. In our new analysis, we rely on the EHV data for the output of Dynegy's Long Beach plant, which is not covered by CEMS.

We also introduce here the use of the BEEP stack data to enhance our understanding of operational details in several ways. First, as mentioned above, the BEEP stack data allow us to account for the dispatched reserves. In addition, they are particularly helpful with respect to the dispatch of gas turbine units. These units are not included in CEMS at all. The output of some GTs, what we refer to below as Big GT, are included in the station-level outputs captured by the WSCC EHV database. Others, which we refer to as Small GT, are not covered by the WSCC EHV database. The Big GT category includes units at Alamitos, Etiwanda, Huntington Beach, and Mandalay, each of which have capacity greater than 100 MW. The Small GT category refers to the Dynegy units in San Diego plus Ellwood. The total capacity of Big GT is 476 MW and Small GT is 403 MW.

The BEEP data enable us to observe when GTs provide real time energy or dispatched reserves. Because it is generally not efficient to run these units partially loaded, we assume that whenever we observe output for a GT in the BEEP stack, that it is producing at capacity. This gives us a crude measure of its output.

We check the EHV data for Long Beach and our use of the BEEP stack to infer GT output against EIA Form 900 data. Using BEEP to infer the output of GTs tends to overestimate actual output.[39] By using this approximation in our withholding analysis, we are underestimating that portion of the withholding that is due to the Big GT component. Table 10 compares the two measurements.

Table 10. Validation of Big GT Output in June

Unit	Real Time Hours	Capacity (MW)	Capacity x Real Time Hours (MWh)	EIA Form 900 (MWh)
Alamitos GT	37	126	4,662	3,817
Etiwanda GT	55	118	6,490	6,947
Huntington Beach GT	44	114	5,016	4,386
Mandalay GT	41	118	4,838	4,120

We also checked the EHV data for Long Beach against EIA Form 900 data. Table 11 compares the output totals from the EHV data and EIA Form 900 from June through August.[40] Since the agreement is quite good in each month, we have confidence to use the EHV representation for this unit in our withholding analysis below.

[39] This statement assumes that whenever Big GTs run, some portion of their output appears in BEEP.

[40] We only have the EHV data through September 20, 2000.

Table 11. Validation of Long Beach Output

Month	EHV Total (MWh)	EIA Form 900 (MWh)
June	49,492	48,654
July	64,768	63,324
August	169,566	162,610

5.2.2 Capacity Data

There are many definitions of unit capacity, which result in quantitative differences that typically are small. Harvey & Hogan take issue with ours. Accordingly, we adopt for this analysis the capacities in Klein (1998) for all steam units. Since Klein does not include gas turbines, we rely on the EIA *Inventory of Electric Utility Power Plants* (2000) and *Inventory of Nonutility Electric Power Plants* (2000). Tables 8 and 9 above retain our former definition for the sake of comparability with our previous results.

5.3 Unit Commitment Decisions

We examine, on a daily basis, which units were turned on during high-price periods. It is important to separate withholding of output from units that are already producing at some level from withholding that results from not turning units on at all. The withholding discussion of Harvey & Hogan tends to suppress this distinction and to ignore the decision to turn units on, even though they make much of the issue in the context of the Alamitos 2 example. By looking at both withholding of units that are running (i.e., committed) and those that are not, we can ask more focused questions about the behavior at issue. We find that many units were not running on high-price days. So much capacity was not committed that in sum it exceeds any reasonable benchmark of forced outages associated with mechanical problems.

Table 12 summarizes the amount of capacity running at peak times by day. Using CEMS, we divide this into steam capacity running all day and steam capacity running part of the day. We measure the part-day capacity at Hour 17, so it excludes units that shut down in the morning. The table also includes a load measure and the ratio of total SP15 capacity to load.

Table 12. Total SP15 Capacity Committed On Peak (MW)

Date	Steam All Day	Steam Part Day	Long Beach	Big GT	Small GT	Total Capacity	ISO Peak	Total Capacity/ ISO Peak
Jun-12	6,271	734	560	0	19	7,584	37,132	0.204
Jun-13	6,855	1,330	560	476	271	9,492	42,288	0.224
Jun-14	7,902	763	560	476	271	9,972	43,447	0.230
Jun-15	7,422	356	560	236	260	8,833	43,146	0.205
Jun-16	6,727	2,273	560	0	0	9,560	39,823	0.240
Jun-17	7,520	480	0	0	0	8,000	33,800	0.237
Jun-21	6,101	603	560	476	223	7,963	41,414	0.192
Jun-22	6,491	906	560	350	260	8,566	40,089	0.214
Jun-23	7,506	283	560	0	0	8,349	37,228	0.224
Jun-26	7,568	952	560	476	271	9,826	42,672	0.230
Jun-27	7,889	471	560	476	271	9,666	42,693	0.226
Jun-28	7,729	806	560	476	256	9,826	42,303	0.232
Jun-29	8,064	296	560	476	260	9,655	41,606	0.232
Jun-30	7,714	500	560	0	0	8,774	38,187	0.230

Table 12 shows that the generators adjusted their operating capacity both from one day to the next, but also within the day. The Steam Part Day capacity shows that many units were started daily but only ran for part of the day. The Big GT and Small GT entries reflect capacity that is running no more than seven or eight hours per day. Given the inefficiency and high emissions rates of the Big GT and Small GT units, in a competitive market we would expect to see them operating only on days when all generating units with lower costs and emissions were already supplying the market to the extent possible.

However, many cheaper units were not supplying the market. For each of the 14 high-price days, Table 13 lists steam units that did not run at all.[41] Together with Table 12, this table shows that significant amounts of steam capacity did not operate on days when more costly and more polluting capacity did run. Among the units listed in Table 13 are some with low emissions rates, such as Redondo Beach 8, Alamitos 5 & 6, and El Segundo 4. If the operators of these plants were acting like price-takers, they would have run as much low-cost and low-emissions capacity as possible before turning to the highest-cost units. Even if the units had minor operating problems, they would still have run at some level if the suppliers were behaving competitively.

[41] Duke's four steam units at South Bay do not appear in Table 13 because they ran on each of the 14 days. As a point of comparison, the data in Table 13 on Reliant's generating units is roughly consistent with the characterization of their units in the FERC Outage Report. In particular, the availability of the Coolwater units is reported in FERC's Figure 9. Units 1 and 2 are reported to have been available 52% and 68% of the time respectively. Table 13 shows both of these units did not operate on 8 of the 14 days we study. Additionally, Coolwater 1 did not operate on another 2 of these days. The availability of Unit 3 (consisting of Units 31 and 32) was reported to be 30%. There are 8 days in our period where one or both of Units 31 and 32 did not run. Thus, the units not running according to Table 13 are unavailable with roughly the frequency that they are characterized in the FERC report. This does not imply there was no strategic element to these outages, but simply that the two data sources are consistent.

Table 13. Uncommitted Steam Units: By Generator and Day

Date	AES/Williams	Reliant	Dynegy
Jun-12	A1, A6, R5, R6	Et3; C1,31-2	Ec3, Es1-4
Jun-13	A1, A6, R5, R6	C31-2	Es3
Jun-14	A1, R5, R6	C31	Es3
Jun-15	A1, R5, R6		Es3
Jun-16	A1, R5, R6	C31-2	
Jun-17	A1, R5, R6	Et1-2; C1,31-2	Es1,2; Ec1-3
Jun-21	R5, R8	M1; Et1,2,4; C1-2	Es1,3; Ec1-3
Jun-22	R5, R8	Et1-2; C1,2,32	Es1; Ec1-3
Jun-23	R5, R8	Et1-2; C1,2,31-2	Es1; Ec1-3
Jun-26	R5, R8	C1-2	Es1-2
Jun-27	R5, R8	C1-2	Es4
Jun-28	R8	C1-2	Es4
Jun-29	A1, R8	C1-2	Es4
Jun-30	A1, R5	C1,2,31-2	Es4

A = Alamitos R = Redondo Beach Et = Etiwanda
C = Coolwater M = Mandalay Ec = Encina Es = El Segundo

In light of the inefficient dispatch that we observe in Tables 12 and 13, we question how the strategy of generators in SCAQMD affects its market fundamentals. The amount of Big GT dispatch may have been sufficiently large to raise the RTC price, in which case it is not exogenous. This would mean that our competitive benchmark analysis should employ lower RTC costs than we previously estimated based on actual history.

Finally, we may ask if the inefficient dispatch was somehow caused by the inefficiencies in the market design alleged by Harvey & Hogan. With the amount of daily capacity adjustment illustrated in Table 12, it seems difficult to argue that generators were somehow prevented by market rules from turning on their capacity when it was economic. The more plausible hypothesis is that generators were withholding at least some of the capacity listed in Table 13.

This hypothesis is supported when we examine the behavior of Duke Energy's SP15 generation. Duke has frequently asserted that it sold 90% of its output forward.[42] Our previous analysis shows that Duke had the smallest percentage output gap of any SP15 generator. Duke has also published some summary forced outage statistics. These are reproduced below in Table 14. This table compares Duke's forced outage rate to similar data for the same plants under IOU ownership and for similar plants in other regions of the US ("NERC" column). These outage rates are very low compared to the "outages" of other SP15 generators. Is Duke so much better at running plants than these other generators? This is possible, but the difference in incentives seems a better explanation.

[42] See "Duke reports below-average prices for Calif. Sales," *Megawatt Daily*, June 4, 2001.

Table 14. Duke Forced Outage Factors[43]

Plant	Duke	IOU	NERC
Moss Landing	3.7%	3.7%	4.27%
Morro Bay	2.9%	11.1%	4.33%
South Bay	1.1%	1.8%	3.61%

5.4 Daily Output Gap

In this section we characterize the output gap on a daily basis. In general, to make a judgment about whether any observed gap represents the potential exercise of market power through withholding, we need a standard of "efficient commitment." If demand is low, for example, it is not necessary to commit all units. If some units did not run on such a day, that should not be construed as market power but simply as a response to demand. We do not attempt to define here an appropriate standard for commitment by individual firms in every situation. Instead, we concentrate on the five days in June when demand was so high that all available capacity should have been running: the 14th, 15th, 26th, 27th and 28th. Table 2 shows that the ISO called for customer interruption on four of those days. This is a good indication that all available capacity was required. The only one of these days on which there was no interruption required was the 15th, but Table 12 shows that peak demand was higher on that day than on the 26th, 27th and 28th and was only 300 MW below the demand on the 14th.

Table 15 shows the gap calculations for the high-priced hours (day ahead or real time) for these five days. The Dispatched Ancillary Services (calculated from the BEEP stack data) is netted off of the AS Total and the resulting Undispatched AS is shown in bold next to the gap. For each of these days the difference between the measured Output Gap and Undispatched AS is between 2,400 and 3,000 MW. At demand levels this high, it is not surprising that prices are high when this much capacity is not supplied.

[43] See http://dena.duke-energy.com//california/releases/nr051401.asp.

Table 15. Daily Output Gap Calculations

Date	Unit Category	Mean Values (MWh)				Dispatched Ancillary Services (MWh)			
		Output	Capacity	Gap	AS Total	Replacement	Spin	Non-Spin	Total
Jun-14	Big GTs	202	476	274					
	CEMS	6,102	9,670	3,568					
	Long Beach	291	560	269					
	Small GTs	102	271	169	2,030	497	23	50	570
	Total	6,697	10,977	**4,280**	**1,460**				
Jun-15	Big GTs	63	476	413					
	CEMS	5,994	9,670	3,676					
	Long Beach	413	560	147					
	Small GTs	54	271	217	2,251	401	24	26	451
	Total	6,524	10,977	**4,453**	**1,800**				
Jun-26	Big GTs	174	476	302					
	CEMS	7,107	9,670	2,563					
	Long Beach	90	560	470					
	Small GTs	131	271	140	1,548	326	48	133	507
	Total	7,502	10,977	**3,475**	**1,041**				
Jun-27	Big GTs	293	476	183					
	CEMS	6,518	9,670	3,152					
	Long Beach	311	560	249					
	Small GTs	113	271	158	1,765	380	5	59	444
	Total	7,235	10,977	**3,742**	**1,321**				
Jun-28	Big GTs	172	476	304					
	CEMS	6,330	9,670	3,340					
	Long Beach	276	560	284					
	Small GTs	71	271	200	1,395	191	7	2	200
	Total	6,849	10,977	**4,128**	**1,195**				

We can compute from Table 13 how much of the observed gap is due to uncommitted units, versus how much is unloaded capacity that is on line. On June 15th, for example, the gap is 80% explained by uncommitted steam units. Based on our earlier analysis there is no reason to believe that it would have been unprofitable to commit and operate these units if the suppliers were behaving competitively.

6. Conclusions

6.1 Competitive Benchmark Price

We revisited our previous estimate of competitive benchmark prices to take account of criticisms made by Harvey & Hogan. We have improved our data and changed our treatment of ancillary services. We have re-estimated the competitive benchmark prices to reflect these changes. We also address the convexity issue raised by Harvey & Hogan in connection with our load approximation. Using more load points in the estimates raises our measure of the competitive benchmark by a small amount, other things equal. The net effect of all changes is relatively small, particularly in comparison to the observed prices in the market, but it goes in the direction of finding more market power and not less. The benchmark cost implication of high RTC prices in SCAQMD is lower than our previous estimate.

6.2 Withholding Analysis

Our withholding analysis broadens the scope of our previous analysis by considering both the CEMS data on steam plants and other data sources on GTs and Dynegy's Long Beach plant. We also consider chronological effects. For the five highest demand days in June, we identify 2,400 to 3,000 MW of capacity that did not operate in SP15 out of a total of about 11,000 MW. This is far above any expectation of forced outages and it occurred long before the "tired plant" effect frequently invoked by the generators in defense of their behavior could have come into play.

It is worth observing that Duke's behavior is qualitatively different from the other SP15 generators. Table 8 is consistent with our previous assessment that shows substantially greater production relative to capacity for this firm as opposed to the other three SP15 generators. Table 14 shows that Duke's forced outage rates are far below what we observe for the other SP15 generators on the high-price days. Duke has asserted many times that they sold 90% of their output forward. It is well known in theory and clear from common sense that a generator that is fully contracted has no incentive to withdraw capacity. This shows that incentives matter.

6.3 Summary

The analysis of Harvey & Hogan contributes little to the debates over the behavior of generators in the California wholesale electricity market. The data and analysis that we present here confirm and strengthen our previous conclusions. It does not address all questions of interest, nor does it cover the entire time period of interest. There is much work that remains to be done. While we agree with Harvey & Hogan that competitive electricity markets are a desirable outcome, it is also important to develop and apply methods to identify and measure the existence and quantitative significance of market imperfections. This is the only way to determine whether electricity markets are delivering on their promise and to identify and implement reforms if they are not.

7. References

Borenstein, S., J. Bushnell and F. Wolak, "Diagnosing Market Power in California's Deregulated Wholesale Electricity Market," University of California Energy Institute working paper PWP-064, August 2000, at http://www.path.berkeley.edu/ucei/PDFDown.html.

Bushnell, J., "Water and Power: Hydroelectric Resources in the Era of Competition in the Western US," University of California Energy Institute working paper PWP-056r, March 5, 1998.

California Power Exchange Corp, Compliance Unit, "Price Movements in California Electricity Markets," September 29, 2000.

California Independent System Operator, Summer 2000 Demand Relief Program, memo to Grid Reliability/Operations Committee, April 2000.

California Independent System Operator, CAISO 2001 Summer Assessment, March 22, 2001.

California Public Utilities Commission, Decision 01-01-056, January 26, 2001.

Duke Energy North America, "Duke Energy: Aggressive efforts to deliver top-ranked reliability at California power plants," at http://dena.duke-energy.com//california/releases/nr051401.asp.

"Duke reports below-average prices for Calif. sales," *Megawatt Daily*, June 4, 2001.

Energy Information Administration, Inventory of Electric Utility Power Plants in the United States 1999, DOE/EIA-0095(99)/1, September 2000.

Energy Information Administration, Inventory of Nonutility Electric Power Plants in the United States 1999, DOE/EIA-0095(99)/2, November 2000.

FERC Office of the General Counsel Market Oversight & Enforcement and Office of Markets, Tariffs and Rates, Division of Energy Markets, Report on Plant Outages in the State of California, February 1, 2001.

Griffes, P., "A report to the California Power Exchange: the benefits of a simultaneous versus sequential PX market for energy and ancillary services," March 15, 1999. Filed with FERC March 30, 1999 by California Power Exchange Corporation, Docket Nos. EC96-19-001, et. al.

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

Harvey, S. and W. Hogan, "On the Exercise of Market Power Through Strategic Withholding in California," April 24, 2001.

Hildebrandt, E., Declaration of Eric Hildebrandt, Federal Energy Regulatory Commission Docket Nos. EL00-95-000 and EL00-98-000, October 2000.

Hildebrandt, E., "Further Analyses of the Exercise and Cost Impacts of Market Power in California's Wholesale Energy Market," Attachment B to Comments of the California Independent System Operator Corporation on Staff's Recommendation on Prospective Market Monitoring and Mitigation for the California Wholesale Electric Power Market, FERC Docket No. EL00-95-012, March 2001.

Joskow, P. and E. Kahn, "A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market During Summer 2000," NBER Working Paper 8157, March 2001, at http://www.nber.org/papers/w8157.

Klein, J., "The Use of Heat Rates in Production Cost Modeling and Market Modeling," April 17, 1998, at http://www.energy.ca.gov/papers/98-04-07_HEATRATE.PDF.

Resource Data International (RDI), Outlook for Power in North America, 1999 Annual Edition, 1999.

Sheffrin, A., "What Went Wrong with California Electric Utility Deregulation," presentation to New Mexico State University Center for Public Utilities Conference, March 25, 2001.

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

Appendix 1. California Summer Hydro Capacity by Owner (MW)

Utility Name	Unadjusted for Partial Ownership	Adjusted for Partial Ownership
Bureau of Reclamation	1,998	2,179
California Dept-Wtr Resources	1,699	1,519
East Bay Municipal Util Dist	39	39
Escondido City of	2	2
Imperial Irrigation District	49	49
Kings River Conservation Dist	165	165
Los Angeles City of	1,764	1,764
Merced Irrigation District	109	109
Metropolitan Water District	102	102
Modesto Irrigation District	0	64
Nevada Irrigation District	86	86
Oakdale & South San Joaquin	97	97
Oroville-Wyandotte Irrig Dist	94	94
Pacific Gas & Electric Co	3,896	3,896
PacifiCorp	76	76
Pasadena City of	2	2
Placer County Water Agency	242	242
Redding City of	1	1
Sacramento Municipal Util Dist	698	691
San Francisco City & County of	385	385
Santa Clara City of	29	29
Sierra Pacific Power Co	3	3
Southern California Edison Co	1,156	1,156
Turlock Irrigation District	218	154
Ukiah City of	4	4
Utica Power Authority	5	5
Yuba County Water Agency	363	363
Solano Irrigation District	12	12
Northern California Power Agny	249	249
South Sutter Water District	0	7
Total	**13,541**	**13,541**

Source: 1999 EIA Form 860A.

Appendix 2. Calculating Ancillary Services Revenues for SP15 Generators

We know from the ISO zonal data how much of each ancillary service is provided in each zone in each hour. The challenge is to allocate these quantities to each of the generators who can provide the particular service. From this allocation, we estimate generators' revenues from each service using the ISO price data.

Inputs

The amount of any given ancillary service a particular generator can supply depends on a number of its characteristics. These include such variables as the unit's scheduled level of output, ramp rates, AGC minimums, maximum capacities, minimum stable generation, and quick-start ability. There are various publicly available data sources that provide these variables. The reliability-must-run (RMR) contracts between the generators and the ISO provide the particular information on ramp rates and AGC minimums. Henwood Energy provided maximum capacities. The value for minimum stable generation (MSG) is estimated from the EPA data. Whether a unit could start up quickly (in less than 10 minutes) was also estimated. For all steam units, we assumed they could not start quickly. For all GTs and CTs, we assumed they could achieve full output in the ten-minute timeframe. For the hydro units in the region, the amount of generation that they provided in June of 2000 is taken from the EIA Form 759 data, which reports how much each generating unit produces in each month.

The ISO does not publish the scheduled level of output for any of the generators. However, the EPA provides estimates of the level of output for generators for which it is monitoring emissions in its CEMS database. We assume the CEMS output levels are the scheduled quantities for each of the generators. This, in fact, is not the case since the CEMS data reflect real time operations that include being called from capacity allocated to ancillary services. At best, the real time data are only approximations for the generation schedules.

Methodology

In allocating the ancillary services revenues across units, there are two steps. First it is necessary to calculate the how much each generating unit has to provide for each of the services. This is done based on generating unit characteristics. Griffes (1999) illustrates how supply curves for these services can be constructed from generating unit characteristics. Second, demand must be allocated to each of the units. In awarding ancillary services contracts, the ISO examines and ranks the bids for the service from each of the generators from lowest to highest. The ISO sets the prices at the level of generator that just meets the quantity it is purchasing. While it is possible to emulate this process, the methodology would require information about each generator's bidding strategy. Rather than hypothesize about the bidding strategies and estimate these bids, we have simply allocated capacity in a pro-rata fashion based on generator's available capacity to provide the particular service.

In calculating the quantity available to provide ancillary services, the units must be available and/or on line. We have assumed that positive output in the CEMS data means the generator is on line. Further, the scheduling of subsequent reserve services must account for how much capacity has already been scheduled for the generator. The ISO clears each of the ancillary services markets sequentially, starting with Regulation Up and ending with Replacement. When determining capacity available for spin, we must take into account for how much Regulation Up has already been scheduled for the unit.

The calculation is slightly different for hydro units. We estimate the capacity factor for the month and apply that percent to the generator's capacity in determining how much it has available to sell into the market. For example, let us assume Big Creek 1 has a capacity factor of 96% and maximum capacity of 300 MW. Because hydro units can ramp instantaneously, we assume it has 12 MW available for Regulation Up and 288 MW for Regulation Down.

Once the total capacity available to provide a particular ancillary service is determined, it is necessary to allocate zonal procurement to each of the generators that have capacity available to provide the service. This allocation is done on a pro rata basis with generators' allocated portion being directly related to the capacity that they have available for that service.

For example, based on generators' characteristics, we have estimated the amount of capacity each generator in SP15 has to provide to the Regulation Up market. Table A2.1 shows our work for the first hour of June 21, 2000. The load and available capacity for Regulation Up are estimated. The table shows that the amount allocated to each unit is in proportion to the total available capacity to provide Regulation.

Table A2.1 Supply for June 21 Hour 1 (MW)

Unit	Load	Capacity	Capacity Available for Reg Up	Allocated Capacity for Reg Up
SCE Hydro	959	1,156	197	22.3
Alamitos 1	12	175	105	11.9
Alamitos 2	12	175	105	11.9
Alamitos 3	26	320	192	21.7
Alamitos 4	112	320	192	21.7
Alamitos 5	72	480	144	16.3
Alamitos 6	71	480	144	16.3
Huntington Beach 1	21	215	129	14.6
Huntington Beach 2	20	215	129	14.6
Redondo Beach 6	10	175	105	11.9
Redondo Beach 7	132	480	105	11.9
Coolwater 42	74	121	46.5	5.3
Encina 4	27	300	189	21.4
Encina 5	27	330	192	21.7
Etiwanda 3	70	320	192	21.7
Mandalay 2	24	215	129	14.6
Mountainview Power C 2	64	63	0	0
El Segundo 4	187	335	148	16.8
Ormond Beach 1	398	750	352	39.9
Ormond Beach 2	62	750	480	54.3
South Bay 1	37	147	84	9.5
South Bay 2	32	150	84	9.5
South Bay 3	32	171	84	9.5
Total	**2,481**	**7,843**	**3,528**	**399**

Table A2.2 shows the results of applying this method for the units and days in Table 6, service by service. We exclude replacement reserve because of difficulty estimating start-up times for all relevant generators.

Table A2.2 AS Revenues by Service ($)

Service	Jun-17		Jun-21			Jun-23
	Alamitos 2	Huntington Beach 2	Alamitos 1	Alamitos 2	Redondo Beach 6	Alamitos 1
Reg Down	1,294	312	1,899	2,524	2,271	1,443
Reg Up	13,674	18,413	10,020	8,731	6,082	24,694
Spin	11	14	21,624	21,143	11,681	3,088
Non-Spin	159	196	12,666	10,818	3,798	70,756
Total	**15,139**	**18,935**	**46,209**	**43,216**	**23,832**	**99,982**

Appendix 3.
The Alamitos 2 Example: A Comparison of Normal and Observed Heat Rates

Harvey & Hogan assert that start-up and minimum-load costs may explain the withholding behavior we identify. They provide indirect evidence for this assertion by documenting one instance in which a unit actually operated unprofitably over the course of an entire day, Alamitos 2 on June 17. In this appendix, we shed further light on the Alamitos 2 anomaly.

The CEMS database is our best available source for the actual heat rate performance of the California steam units in a given hour, but it masks important information about the units' long-term operating history or their expected capabilities. A comparison with Klein (1998) is helpful. The Klein estimates are based on a formula fit to test data and seem generally to under-predict actual heat rates. The Klein formula is useful, however, as a benchmark for heat rate degradation that might reflect operating problems at a particular unit.

We begin with an example of a unit which shows a close correlation between CEMS and Klein. Table A3.1 shows heat rates for Alamitos 4, a 320 MW steam unit. On three Sundays in June 2000, the unit ran at an average of about 50 MW, which is roughly 15% of its capacity. By contrast, on the other 27 days, it ran at 207 MW or 65% of capacity. As we would expect, the average heat rate on the low-load-factor days was 29-35% higher than on the high-load-factor days. Both CEMS and Klein show this predictable increase.

Table A3.1 Alamitos 4 Heat Rate and Average Loading

Number of Days	CEMS Heat Rate	Klein Heat Rate	CEMS HR/Klein HR	Average MW
3	16765	17827	0.9404	50.14
27	13033	13161	0.9903	206.93

The unit that Harvey & Hogan select, Alamitos 2, is a 175 MW unit. Table A3.2 shows its heat rates on June 17 (Harvey & Hogan's test day) and June 21 (the other day in June with only one high-price hour).

Table A3.2 Alamitos 2 Heat Rate and Average Loading

Date	CEMS Heat Rate	Klein Heat Rate	CEMS HR/Klein HR	Average MW
Jun-17	21422	17346	1.2350	45.04
Jun-21	17842	16745	1.0655	61.92

Table A3.2 shows that Alamitos 2 ran more on the 21st than on the 17th (62 MW vs. 45 MW) and therefore its average heat rate was lower. However, whereas there is fairly close agreement (7%) between CEMS and Klein on the 21st, there is a much larger deviation (24%) on the 17th. Examination of the performance of Alamitos 2 shows that the 5-day period between June 15 and 19 was less efficient than the norm during the rest of the month. Table A3.3 illustrates this.

Table A3.3 Alamitos 2 Heat Rate and Average Loading

Number of Days	CEMS Heat Rate	Klein Heat Rate	CEMS HR/Klein HR	Average MW
25	17761	16573	1.0717	66.73
5	21841	17963	1.2159	41.97

Since Harvey & Hogan are looking for a case where the average heat rate is high, the period from June 15-19 for Alamitos 2 satisfies this criterion. We do not know why these heat rate variations occur. Like so much else in the analysis of generator behavior, these variations could be endogenous.