

# Capacity withholding equilibrium in wholesale electricity markets

Lester B. Lave and Dmitri Perekhodtsev

July 23, 2001

## **Abstract**

This paper is a quantitative study of the capacity withholding incentives in the deregulated wholesale electricity markets and resulting price spikes. For the analysis we used an N-player Nash equilibrium model based on marginal cost functions of the generating firms assuming completely inelastic industry demand and complete information. The current results show that in the case of continuous marginal costs the withholding incentive always exists and the total withholding increases with the increase of the curvature of marginal cost functions and the extent of heterogeneity of the generating firms. The discrete form of real marginal cost functions imposes certain restrictions on the withholding. The analysis shows that there exists a threshold level of market demand below which no withholding occurs and above which the withholding becomes beneficial. The curvature and heterogeneity of marginal cost functions also affect the level of this threshold. The model is applied to the power generation in California ISO area.

# 1 Introduction

In summer 2000 the wholesale electricity prices in California went up about four times above their usual summer levels. This price increase was very continuous; unlike previous price spikes it lasted from June to September. Although, partially this price increase may be attributed to the increase in natural gas prices and unusually high temperatures that boosted the energy consumption and caused water shortage in the North-West, researches shows that all these factors together do not explain such a price increase [Kahn]. The remaining gap is likely to be due to the imperfections in the structure of the deregulation of the electricity industry in California.

In California, the power system was restructured in 1998 with the intent to separate the generation component of the industry from the transmission and distribution segments. The reason for such deregulation was to make the generation sector competitive. There was an evidence that the wholesale price of electricity might be lowered if the generation were competitive as opposed to being a natural monopoly. Technology improvements in power generation, transmission and dispatch made such deregulation seemingly beneficial for both generators and consumers.

The generating capacity in California is represented mainly by nuclear power plants, hydro-electric plants, fossil fuel plants (natural gas, oil, coal). Nuclear and hydro plants accounts for about one third of in-state generation. Biomass, solar, wind and geothermal generation takes about 10% of the total generation and over a half of the total in-state capacity is due to the gas-fired steam and peaking units including cogeneration (Source: California ISO).

The marginal cost of electricity production on nuclear, hydro and other renewable power plants is very close to zero, while marginal cost of gas turbines is mainly determined by the price of natural gas and emission permits and may differ a lot across plants as the heat and emission rates are different. This makes the state generation marginal cost schedule very steep at the capacity margins.

The organization of the restructured power market is the following. Each day a specific day ahead trading exchange (the Power Exchange, PX) collects hourly supply and demand bids from electricity generating companies and utility distribution companies for the elec-

tricity scheduled to be produced next day. The generators and utilities are allowed to bid differently for different units of energy. The resulting bid schedules are monotonic piecewise "staircase" functions. The Power Exchange intersects the resulting supply and demand schedules and arrives at the market-clearing price for each hour. This price is equal to the bid of the last energy unit purchased on the exchange.

To ensure that the real demand equals to the supply and that the reliability criteria are met a real-time hourly market is conducted by the operator of the electricity grid (the Independent System Operator, ISO). Both the day-ahead market at the Power Exchange and ISO real-time market are designed as uniform price auctions, that is auctions where all winning bids are purchased for the same market clearing price.

The reason for choosing the uniform price auction for the wholesale electricity trading was its similarity to the second price auction scheme which truth-revealing properties were proven by the theory Milgrom and Weber (1982). However, the generating companies are allowed to trade several units of energy at a time. Such multi-unit auction turns out to be very different from one-unit second-price auctions and truth revealing does not work that well in multi-unit auctions.

If each auctioneer trades two units Engelbrecht-Wiggans and Kahn (1998) showed for the buying auction that only one of the two bids is equal to the personal valuation, while the second bid is lower than the personal valuation. Translated into the selling auction this would mean that only one unit is sold for the marginal cost, while the other unit is bid higher than the marginal cost in equilibrium.

When the auctioneer bids for many units, it is convenient to represent his bids as functions. To see what happens when players compete with bid functions an approach called "supply-function equilibrium" was developed by Klemperer and Meyer (1989). In selling auction framework their result is that in the presence of uncertainty about the total demand the resulting supply functions are steeper than the marginal cost curves. However, in the absence of this uncertainty there exists an enormous multiplicity of equilibria in supply functions.

The "supply function equilibrium" was applied to model the wholesale electricity markets by Green and Newbery (1992) and Green (1996) in application to British electricity market approach and Baldick, Grant and Kahn (2000) in application to US markets. However, since the approach is technically involved, these applications were limited to linear supply function

equilibria.

Another approach to model wholesale electricity markets called "bid function equilibria" was developed by Crespo(2000). Bid function equilibria are Nash equilibria of an oligopoly model of multi-unit auction under complete information. The approach predicts asymmetric bidding by producers: a single firm (the price setter) bids strategically while other firms bid their cost. In 2001 Crawford, Crespo and Tauchen (2001) applied the bid function equilibria approach to model British spot market for electricity.

All of the existing models of uniform price auctions give the evidence of the possibility of exercising market power by the agents. There is a number of empirical works aimed to measuring the market power in wholesale power markets. Wolfram (1998) and (1999) measures the market power in the British electricity spot market, Puller (2001) finds the empirical evidence of exercising market power on California deregulated electricity market. Wolak (2001) analyses the positive effect of hedge contracts on the market power on Australian wholesale electricity market.

Works on the forecast of the agents' behavior on the electricity markets were done in USA long before the California energy crisis occurred. Borenstein and Bushnell (1998) made a detailed forecast of the possibility and severity of market power in California using Cournot oligopoly model. Finally, Joskow and Kahn (2001) carried out a detailed comparison of prices and marginal costs during summer 2000 in California concluding that the existing increase in marginal cost of electricity production does not explain the observed price increase. The authors claim that the remaining gap is due to capacity withholding behavior of the power generators provided by the auction design. Indeed, in 1998 Ausubel and Cramton (1998) proved that in multi-unit uniform price auctions there exists an incentive to reduce the demand (if this is a buying auction) or supply (if the auction is selling). Joskow and Kahn assumed that the increase in marginal cost could boost the withholding which led to the sharp price increase.

In this paper we elaborate on the Joskow and Kahn's claim about capacity withholding incentives designing a model of capacity withholding equilibrium. In its structure and simplicity it resembles Cournot model. The model assumes perfect information, it can be applied to a class of nonlinear marginal cost functions and results in a unique equilibrium. All this makes the model particularly useful in modeling withholding behavior and resulting

market power on wholesale electricity markets.

In Section 2 we describe the baseline model of the withholding equilibrium where marginal cost functions of all generators are considered the same linear increasing functions. In Section 2.5 we relax the linearity assumption and in Section 4 symmetry assumption is relaxed. In Section 3 the extended model is applied to California ISO area generation and finally, in Section 4 the results are summarized and the ways of possible improvement of California situation in the framework of the existing auction scheme are discussed.

## 2 Continuous linear symmetric model

### 2.1 Withholding incentives

The design of the uniform price auction is supposed to be truth-revealing, that is, since the market-clearing price does not depend on agent's decisions directly, agents should not have incentives to bid above their cost. However, this works only for the auctions where each seller sells only one unit of good. In the electricity auction where sellers operate with several units this does not hold.

*Figure 1. is about here*

On the Figure 1. the light-grey bars represent the bids of all the auction participants but one, the dark-grey bars are the marginal costs of the units of the last generating company, the vertical line represents the demand and the horizontal line determines the correspondent market-clearing price. The shaded areas between the marginal cost of the last generator and the market-clearing price represent the profit of the generator. If other things kept constant, the generator can improve on his situation and get a higher profit by means of withholding some units.

*Figure 2. is about here*

On the Figure 2. the most expensive unit in terms of marginal cost is withheld. The effect of this is a shift of the demand line to the right by the amount of withheld capacity and an increase in market-clearing price. The generator derives some extra profit from the remaining inframarginal units because of the price increase and loses the profit from the

unit withheld. If this extra profit is bigger than the lost, such a withdrawing is profitable for the generator. If the generator withheld a next expensive unit instead of most expensive inframarginal unit as indicated on Figure 3. he would be worse off other things (the size of both units) being equal since the sacrificed profit now is higher.

*Figure 3. is about here*

Finally, it seems that the withholding is the only profit-increasing strategy when the bids of other parties and the demand are held constant. Bidding higher than the marginal cost for the inframarginal units does not make any difference until you bid below market price. If one bids for a unit above the market price than you can think of such a unit as of a withheld unit.

To summarize, one can think of the following heuristic rules of withholding:

1. the more inframarginal units you have, the more incentive to withhold you have
2. the steeper the bids of the other parties go up, the more likely you withhold
3. choosing between two equal sized units to withhold you would prefer to withhold the most expensive

## **2.2 Continuous symmetric model**

To understand the nature of the withholding behaviour in the whole group of electricity generators in order to understand whether this process blows up or stabilizes somewhere an attempt to find a Nash equilibrium of the withholding is made. However, it is extremely difficult to operate with bids in the form of discrete units of completely different generators of difference size and marginal costs. So we make a couple of simplifying assumptions which allow me to find the equilibrium in close form. These assumptions are:

1. continuous marginal cost
2. symmetry

With continuous marginal costs instead of stepwise functions one can deal with functions instead of multiple units. This is justified if the average size of the units is negligible compared to the total supply of the generator. However, later we will show that even in continuous

model the size of units plays important role. The symmetry helps a lot in constructing a Nash equilibrium for  $N$  players and allows us not to account for the capacity constraints of each generator.

*Figure 4. is about here*

Figure 4. shows the symmetric case (no withholding by far). I need to take increasing concave functions for marginal cost of each generator so that the industry marginal cost resemble the real one. If the marginal cost of each generator is  $p = mc(q)$ , where  $q$  is the quantity he produces, then, the industry marginal cost will be  $MC(Q) = mc(Q/N)$ , where  $N$  is the number of generators. If the industry demand is  $D$ , then the market-clearing price will be  $p = mc(D/N)$  and each of the generators will face his demand share  $d = D/N$ .

*Figure 5. is about here*

It is important to see what is the effect of withholding some capacity by one of the generators when the rest of generators do not withhold. On the left panel of Figure 5. there is the generator's marginal cost, on the right panel industry marginal cost. If the generator withholds the amount of inframarginal capacity  $q$ , the same amount  $q$  will disappear on the right panel to the left from the demand. To make up for this deficit the demand line will need to shift to the right by the same amount  $q$ . The market-clearing price will go up correspondingly and the individual demand lines will shift to the right by  $q/N$  for all generators. For the generator that withheld, the total demand will change by  $-q + q/N$ .

As we noticed before, a generator should withhold the most expensive inframarginal units in terms of marginal cost. However, there is one more important point about which exactly units to withhold. Suppose that the generator decided that it is profitable to withdraw amount of capacity  $\Delta q$  - the most expensive inframarginal capacity, as indicated on Figure 6. After he does that (assuming that everybody else are still holding to their bids) his personal demand line shifts to the right by  $\Delta q/N$  and he finds himself selling  $\Delta q/N$  of previously ultramarginal capacity, which is much more expensive than some of the capacity he withheld. So, foreseeing this effect the generator, once decided to withhold  $\Delta q$ , would prefer to do that so that  $\Delta q/N$  of the capacity immediately to the right from the demand line and  $\Delta q(N - 1)/N$  of the most expensive inframarginal capacity is withheld.

*Figure 6. is about here*

### 2.3 Nash Equilibrium

In this section the equilibrium conditions for the case of linear marginal cost function will be derived. Precisely, marginal cost of each generator is assumed to be  $mc(q) = q$ . The system demand is assumed to be perfectly inelastic. Unlike most auction models we assume complete information. This assumption is reasonable here since the industry is pretty transparent. Marginal cost of generation of the rivals as well as the system demand can be estimated with high degree of accuracy.

Before considering the Nash equilibrium we find the amount of capacity  $\Delta q$  that is optimal for a generator to withhold when the industry demand is  $D$ .

Suppose the generator withholds  $\Delta q$  exactly in the way described above. The resulting price increase will be

$$\Delta p = \Delta q / N$$

*Figure 7. is about here*

The extra profit on the inframarginal capacity will be (Figure 7.):

$$R(\Delta q) = \frac{D - \Delta q(N - 1)}{N} \frac{\Delta q}{N}$$

The sacrificed profit is:

$$C(\Delta q) = \frac{\Delta q^2(N - 1)^2}{2N^2}$$

The total change in profit will be given by

$$\Delta \pi = D \frac{\Delta q}{N^2} - \Delta q^2 \frac{N - 1}{N} - \frac{1}{2} \Delta q^2 \frac{(N - 1)^2}{N^2}$$

This is a quadratic function with the maximum in

$$\Delta q = \frac{D}{N^2 - 1}$$

Suppose now that there exists an equilibrium  $\Delta q^*$  such that if all of the agents withhold  $\Delta q^*$ , nobody wants to deviate. To find such  $\Delta q^*$  suppose that  $N - 1$  agents already withheld  $\Delta q^*$ . Then find the optimal capacity the Nth agent would choose to withhold. Since  $N - 1$



agents withheld  $\Delta q^*$  each, the industry demand line is now at  $D' = D + \Delta q^*(N - 1)$  and the best withholding amount for the Nth agent will be

$$\Delta q = \frac{D + \Delta q^*(N - 1)}{N^2 - 1}$$

For the equilibrium we need this  $\Delta q$  be equal to  $\Delta q^*$ . Solving the resulting equation in terms of  $\Delta q^*$  we get

$$\Delta q^* = \frac{D}{N(N - 1)}$$

This is the amount of capacity that each generator withholds in equilibrium, making the total withholding

$$\Delta Q^* = \frac{D}{N - 1}$$

## 2.4 Supply

Now when the equilibrium withholding of each generator is obtained the effective industry supply function can be obtained. Whenever the industry faces the demand  $D$ , the total withholding is  $D/(N - 1)$ . This is equivalent to shifting the industry demand to the level  $D' = \frac{ND}{N-1}$ . That is, the industry inverse supply function will now be

$$p = \frac{Q}{N - 1}$$

instead of the competitive case where supply is equal marginal cost  $p = \frac{Q}{N}$ .

That means that whatever the demand is, the resulting price will always be by  $N/(N - 1)$  higher in the case of linear marginal cost (Figure 8.)

*Figure 8. is about here*

If we now depart from the continuity assumption, that is, remember that the generators bid for generating units of certain size, this will have an interesting implication.

Suppose, the average size of a unit is  $\Delta$ , then, if the optimal withholding  $\Delta q^*$  is less than  $\Delta$  no withholding should take place.

This makes the industry supply with non-zero generating units consist of two parts:

supply equal to the marginal cost  $p = Q/N$  when  $Q < \Delta N(N - 1)$

and supply  $p = Q/(N - 1)$  when  $Q > \Delta N(N - 1)$  (Figure 8.)

## 2.5 Nonlinear marginal costs

The assumption of linear marginal cost functions is non-realistic. The real industry marginal cost function suggests that in the symmetrical case marginal costs should be convex functions. However, it is hard to obtain the equilibrium withholding in closed form for marginal cost function more complex than quadratic. Therefore, to deal with nonlinear cost functions one could use the trick of linearization: linearize the marginal cost function at the individual generator demand (Figure 9.). One can use linearization if the resulting equilibrium withholding  $\Delta q$  is small enough to think that the real marginal cost curve does not differ much from its tangent line over that interval.

*Figure 9. is about here*

In the linearized case the optimal withholding does not differ from the linear case  $mc(q) = q$ . That is, optimal withholding is not affected by the slope of marginal cost. There is a straightforward explanation to that: the extra profit of withholding is proportional to the slope of the marginal cost  $\alpha$

$$\Delta\pi = \alpha(R(\Delta q) - C(\Delta q))$$

and the value  $\Delta q$  maximizing the extra profit does not depend on the slope  $\alpha$ .

So, as in the linear case, in the linearized case the optimal withholding is

$$\Delta q^* = \frac{D}{N(N-1)}$$

However, if the marginal cost functions are very convex, the linearization will give inaccurate results. In fact, the curvier is the marginal cost function, the higher is the total withholding.

In general, if the number of firms is  $N$  and the industry demand is  $D$ , the optimal withholding of a firm  $\Delta q$  given the other  $N - 1$  firms bid their marginal costs is given by the formula (see Appendix) :

$$mc' \left( \frac{D + \Delta q}{N} \right) \left( \frac{D - (N-1)\Delta q}{N} \right) = (N-1) \left( mc \left( \frac{D + \Delta q}{N} \right) - mc \left( \frac{D - (N-1)\Delta q}{N} \right) \right)$$

With the substitution

$$\begin{aligned} y &= \frac{D + \Delta q}{N} \\ z &= \Delta q \end{aligned}$$

the formula reduces to

$$mc'(y)(y - z) = (N - 1)(mc(y) - mc(y - z))$$

It turns out that for constant relative risk aversion marginal cost functions of the form  $mc(q) = Aq^\epsilon + B$  where  $A, B$  and  $\epsilon$  are constants, the optimal individual withholding and, therefore, equilibrium individual and industry withholdings are proportional to the industry demand. This makes the above functional form very useful in approximating real marginal costs.

The relative risk aversion (RRA) parameter  $\epsilon$  together with the number of firms  $N$  determines the coefficient  $\alpha$  through a transcendent equation

$$\alpha\epsilon - (N - 1)(1 - \alpha^\epsilon) = 0$$

If all but one firms bid their marginal costs, the optimal withholding for the remaining firm is

$$\Delta q = \frac{(1 - \alpha)D}{N - (1 - \alpha)}$$

Thus, the equilibrium individual withholding will be given by

$$\Delta q^* = \frac{(1 - \alpha)(D + \Delta q^*(N - 1))}{N - (1 - \alpha)}$$

and that will be

$$\Delta q^* = \frac{D(1 - \alpha)}{N\alpha}$$

The total industry withholding is

$$\Delta Q^* = \frac{D(1 - \alpha)}{\alpha}$$

*Figure 11. is about here*

Figure 11. shows how the multiple  $\frac{1-\alpha}{\alpha}$  depends on the parameter of relative risk aversion  $\epsilon$  and the number of generators  $N$ . One can see that the total withholding increases as the number of firms decreases and as the RRA parameter  $\epsilon$  increases. With  $\epsilon$  reasonably below 10 and number of firms about 10 the industry withholding is about 20 – 30% of the demand.

As for the supply function, the convexity of marginal cost has substantial effect it even in the linearized approximation. For example, if the generators' marginal cost functions

were  $mc(q) = q^2$  and the industry marginal cost were  $MC(Q) = \left(\frac{Q}{N}\right)^2$ , then in linearized withholding equilibrium the effective inverse supply would be  $p = \left(\frac{Q}{N-1}\right)^2$ . That is, the resulting price is higher than the marginal cost by  $\left(\frac{N}{N-1}\right)^2$  (Figure 10.).

*Figure 10. is about here*

If marginal cost function can be approximated by the constant relative risk aversion function  $mc(q) = Aq^\epsilon + B$ , the effective inverse supply function will be

$$p = mc\left(\frac{D + \Delta Q^*}{N}\right) = mc\left(\frac{D + \frac{D(1-\alpha)}{\alpha}}{N}\right) = mc\left(\frac{D}{\alpha N}\right)$$

Since  $\alpha$  decreases as coefficient of relative risk aversion  $\epsilon$  increases, the resulting price is affected by the curvature of marginal cost to high extent.

## 2.6 Linear asymmetric case

The symmetry assumption made in Section 2.3 is also far from reality, the capacity of generating companies varies in the wide range from 7000MW to less than 100MW. The complete departure from symmetry assumption is computationally rather difficult. However, it is possible to make an asymmetric estimation assuming all the generator's marginal cost functions being of the same functional form. We first restrict the analysis of asymmetric behavior to linear marginal cost functions. Suppose the marginal cost of generator  $i$  is

$$mc_i(q) = A\frac{q}{\beta_i} + B$$

Where  $A$ ,  $B$  and  $\epsilon$  are constants;  $\beta_i$  represents the proportional capacity of generator  $i$  and  $\sum \beta_i = 1$ .

Similarly to the analysis of Section 2 the set of optimal withholding strategies  $\{\Delta q_i^*\}$  is given by a system of linear equations:

$$\Delta q_i^* = \frac{\beta_i^2 (D + \sum_{j \neq i} \Delta q_j^*)}{\sum_{j \neq i} \beta_j (\sum \beta + \beta_i)} = \frac{\beta_i^2 (D + \sum_{j \neq i} \Delta q_j^*)}{1 - \beta_i^2}, \forall i$$

Or,

$$\begin{bmatrix} \left(1 - \frac{1}{\beta_1^2}\right) & 1 & 1 & \dots & 1 \\ 1 & \left(1 - \frac{1}{\beta_2^2}\right) & 1 & \dots & 1 \\ & \dots & & & \\ 1 & 1 & 1 & \dots & \left(1 - \frac{1}{\beta_N^2}\right) \end{bmatrix} \begin{bmatrix} \Delta q_1^* \\ \Delta q_2^* \\ \vdots \\ \Delta q_N^* \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix} D$$

It is interesting to see how the total industry withholding  $\Delta Q = \sum \Delta q_i^*$  depends on the degree of heterogeneity. Denote  $H = N \sum \beta_i^2$  as a measure of degree of heterogeneity.  $H$  equals to one when all generators are identical and increases as they become heterogeneous. The dependence of the ratio of the total industry withholding  $\Delta Q$  to the industry demand  $D$  on  $H$  is shown on Figure 12 for  $N = 10$ .

*Figure 12. is about here*

Now the withholdings are different across generating companies and this has important implications when one seeks to account for the size of energy units. The optimal withholding  $\Delta q_i^*$  positively depends on the agent's market share  $\beta_i$ . Agents withhold only if their  $\Delta q_i^*$  are greater than their unit size  $\Delta$ . Since optimal individual withholding is proportional to the industry demand  $D$ , big generators start withholding under low values of industry demand while small generators find beneficial to withhold either under very high levels of industry demand or never (Figure 13).

*Figure 13. is about here*

The threshold industry demand beyond which the generators start massive withholding is now diffuse compared to symmetric case. Small generators that do not withhold under reasonable levels of demand can be considered price takers. They bid marginal cost and for simplicity they can be stacked together and considered as a competitive fringe.

Technically this can be done in the following way. Suppose agents  $1, \dots, I$  are big enough to start withholding under reasonable levels of demand and generators  $I + 1, \dots, N$  are small and can be thought of as a competitive fringe.

Then, to find the optimal withholding of the big generators solve the reduced system:

$$\begin{bmatrix} \left(1 - \frac{1}{\beta_1^2}\right) & 1 & \dots & 1 \\ & \dots & & \\ 1 & \dots & 1 & \left(1 - \frac{1}{\beta_I^2}\right) \end{bmatrix} \begin{bmatrix} \Delta q_1^* \\ \vdots \\ \Delta q_I^* \end{bmatrix} = \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix} D$$

## 2.7 Asymmetric nonlinear marginal costs

To account for both nonlinearity and asymmetry of marginal cost functions it is convenient to think of the industry marginal cost in the form

$$MC(Q) = AQ^\epsilon + B$$

so that marginal cost of generator  $i$  be

$$mc_i(q) = A \left( \frac{q}{\beta_i} \right)^\epsilon + B,$$

where

$$\sum \beta_i = 1 \quad ,$$

$A$ ,  $B$  and  $\epsilon$  are constants. Similarly to the calculations in section 2.5 the optimal withholding of generator  $i$   $\Delta q_i$  given that other generators do not withhold is

$$\Delta q_i = \frac{D(1 - \alpha_i)\beta_i}{1 - \beta_i(1 - \alpha_i)},$$

where  $\alpha_i$  is found from the transcendent equation

$$\alpha_i^\epsilon - \frac{1 - \beta_i}{\beta_i}(1 - \alpha_i^\epsilon) = 0$$

(See the calculations in the Appendix).

To find the vector of equilibrium withholdings the following system of equations should be solved:

$$\Delta q_i^* = \frac{(1 - \alpha_i)\beta_i(D + \sum_{j \neq i} \Delta q_j^*)}{1 - (1 - \alpha_i)\beta_i}, \forall i$$

or

$$\begin{bmatrix} \left(1 - \frac{1}{\gamma_1}\right) & 1 & 1 & \dots & 1 \\ 1 & \left(1 - \frac{1}{\gamma_2}\right) & 1 & \dots & 1 \\ & \dots & & & \\ 1 & 1 & 1 & \dots & \left(1 - \frac{1}{\gamma_N}\right) \end{bmatrix} \begin{bmatrix} \Delta q_1^* \\ \Delta q_2^* \\ \vdots \\ \Delta q_N^* \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix} D, \text{ where } \gamma_i = (1 - \alpha_i)\beta_i$$

Similarly to section 2.6 if generators  $I + 1$  to  $N$  are thought to form a competitive fringe a reduced system is solved:

$$\begin{bmatrix} \left(1 - \frac{1}{\gamma_1}\right) & 1 & \dots & 1 \\ & \dots & & \\ 1 & \dots & 1 & \left(1 - \frac{1}{\gamma_I}\right) \end{bmatrix} \begin{bmatrix} \Delta q_1^* \\ \vdots \\ \Delta q_I^* \end{bmatrix} = \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix} D$$

### 3 Empirical study

In this section the model described above will be applied to the behavior of California deregulated electricity market. According to the California Independent System Operator Summer Assessment about 47,000MW of California generating capacity is distributed across the generation types as shown on Table 1.

*Table 1. is about here*

In our study we will focus on the gas-fired electricity generation. There is a series of reasons for that.

First, gas-fired generation has the highest marginal cost and most often the marginal supply resource that clears supply and demand belongs to this part of generation capacity. So, the gas-fired power plants conveniently constitute the "top" of CAISO's area generation marginal cost curve.

Second, the rest of generation capacity is not likely to participate in the withholding behavior. Nuclear plants are a continuous technology which is very expensive to interrupt. Wind, solar, geothermal, biomass generation and cogeneration usually belong to small firms often consisting of one generating unit which makes the withholding not worthy for them. Hydro generation usually is owned by federal, state or municipal agencies and profit-maximization is not the purpose number one for these agencies. In addition, the operators of hydro power are to high extent restricted by weather conditions, which decreases the withholding opportunities. The same can be said about wind and solar generation.

Third, there are five big firms owning thermal and peaker gas-fired capacity who has a real opportunity and incentives for withholding. Table 2. shows the list of these firms.

*Table 2. is about here*

To find the system withholding, first, the marginal cost curve of the thermal and peaking generation was found. For that purpose EIA-767 database was used. This database provides the information on the electricity produced and the amounts of fuel purchased on the unit basis. Since the fuel cost constitutes the major part of the operating costs this data is used to construct the system marginal cost curve. In fact, the price of the gas does not matter much

for our analysis as long as we consider the gas cost as the only operating cost. However, to be specific Figure 14. shows the marginal cost of gas units under the gas price of \$6/Mcf.

*Figure 14. is about here*

To find the coefficient of relative risk aversion  $\epsilon$  of such marginal cost curve we first approximate it with a polynomial of high order. Empirically, the order of the polynomial was chosen 8. This order gives a good approximation of the general shape not paying attention to local nonlinearities. After that the coefficient of relative risk aversion  $\epsilon$  is calculated along all the curve (Figure 15.)

*Figure 15. is about here*

Now we apply the model discussed above using the resulting RRA coefficient for  $\epsilon$  and the percentages from Table 2. for the market shares of the five major generating firms. A simplifying assumption used here is that all the generators may not have the same RRA coefficient though the model suggests that. The rationale for thinking that they do is that for the withholding decision what matters most is not the self marginal cost curve but the marginal cost of all other generating firms. The assumption that all the residual marginal cost curves resemble the system marginal cost is more or less reasonable.

The schedule of the industry withholding as a function of the demand faced by the gas-fired units is shown on Figure 16.

*Figure 16. is about here*

The size of units of the gas-fired generation vary in the interval from 20MW to 813MW according to EIA Form 767 and the average unit size could be taken for the discrete withholding analysis. However, there is a clear dependence between the size and the marginal cost of units. Figure 17. shows the 20-units moving average of the unit size as a function of the demand faced by gas units. These values of unit sizes were taken for correspondent demand levels.

*Figure 17. is about here*



After the unit size is accounted for through an iterative rationing procedure (Appendix), the total withholding becomes as shown on Figure 18. The vertical line represents the demand level beyond which the capacity constraints become binding for the withholding behavior. This level corresponds to the system demand of 43,000MW. The system demand level at which the withholding starts according to the model is about 38,000MW. However, because of the forced outages in the low-cost part of the system marginal cost schedule, these numbers may be lower. For example, if about 3,000MW of hydro power was curtailed due to low water levels, the withholding would start at 35,000MW. The withholding schedule has a sharp increase in the high-demand part. This occurs because of the high curvature of the marginal cost at high levels of demand.

*Figure 18. is about here*

From the information about the withholding the resulting supply function can be obtained. With the demand below 38,000MW the system supply matches with the system marginal cost, whereas when the demand is above that point the industry supply becomes much steeper than the marginal cost. At the demand levels beyond 43,000MW the capacity constraints become binding and the pricing is no longer determined by the marginal cost. Whatever is the highest bid, it is taken, so the market clearing price may be as high as the price cap.

According to the FERC order from June 19, 2001 the price cap is set at the marginal cost of the most expensive peaker. This price cap does not eliminate the withholding incentives. Withholding will still be profitable under low levels of demand. As the demand increases beyond the levels where the capacity constraints become binding the withholding will have to diminish in order to avoid blackouts, but bidding at the price cap will still occur much more often than the supply at the marginal cost would suggest.

*Figure 19. is about here*

The study of the historical data gives some support of the theory. The California ISO has the data on planned and unplanned non-operational generation on the daily basis. One could think of the total peak system demand as of a sum of the peak load and the planned non-operational generation and consider the unplanned non-operational generation withheld

for the strategic purposes. The dependence of the withheld capacity on the total demand can be represented on a dotplot on Figure 20. The regression line of the ascending part of the dotplot is

$$UNPLANNED = -744 + 0.201 \quad TOTAL$$

$$(0.821) \quad (0.057)$$

And for the descending part:

$$UNPLANNED = 16891 - 0.301 \quad TOTAL$$

$$(0.000) \quad (0.000)$$

*Figure 20. is about here*

The shape of this dependence resembles the withholding graph obtained from the model on Figure 18. However, the increasing part of the real withholding schedule is situated on the 30,000 – 35,000MW level of the total demand, while the predicted increasing part is in the range of 38,000 - 42,000MW of the total demand. The reason for this shift might partly be the fact that each period of time roughly 60% of hydro generation is available (Joskow & Kahn (2001)) which can explain a demand shift of about 5,000MW. The real peak unplanned outage is also higher than predicted. This may be due to either collusive behavior of the generating firms or actual unplanned maintenance operations.

## 4 Conclusion

In this project we arrived at the following results

- The withholding behavior has an equilibrium
- The total withholding decreases as the number of generators increases
- Heterogeneity of capacity across the generators adversely affects the total withholding
- The equilibrium quantity withheld does not depend on the slope of the marginal cost function. It rather depends on how convex the marginal cost function is.

- Convexity of marginal cost functions adversely affects the total withholding
- Due to the non-zero size of energy units there exists a specific level of industry demand for each generator below which it is not worthy for the generator to withhold. These threshold levels depend on the average size of units, the number of generators and their market shares.
- For very small generators the generator-specific threshold demand levels may be too high to occur in practice. Such generators are never likely to withhold and can be thought of as a competitive fringe. The presence of a competitive fringe in the market tends to reduce the equilibrium withholding as well as the resulting price.
- When the capacity constraints become binding, the withholding is determined by the remaining capacity. Thus, there is a trend of withholding reduction at extreme levels of industry demand.

The source of problems in the California ISO area is in the tight capacity. The steep and convex part of the marginal cost on the capacity margin tends to boost the price not only under extraordinary high system demand but under much lower demand levels because of the withholding incentives. According to the above analysis, the following might help to reduce these incentives:

Further disinvestment: special attention should be paid to the five major owners of gas-fired generation. The best would be to make them or some of them competitive fringe. However, even if the disinvested firms still find it profitable to withhold, the total withholding will be smaller if the number of the firms increases. The resulting disinvested firms should be of equal capacity to reduce the possible withholding.

Restrict the number of bids per generating firm. Fix the assignment of energy units to the correspondent bidding units once and forever. This could make the withholding start at higher demand levels. Ideal decision would be to allow the generators to make only one bid for all their generation. The resulting spending of the purchasing parties may be higher but the benefit from the eliminated withholding can offset this loss.

Adding even small capacities can help a lot since this would make the demand fall on the flatter part of the marginal cost more often.

## References

Ausubel, Lawrence M., Cramton, Peter "Demand Reduction and Inefficiency in Multi-Unit Auctions" mimeo, University of Maryland (1998)

Baldick, Ross, Grant, Ryan, Kahn, Edward "Linear Supply Function Equilibrium: Generalizations, Application, and Limitations" POWER working paper, PWP-078 (Aug., 2000)

Borenstein, Severin, Bushnell, James "An Empirical Analysis of the Potential for Market Power in the California's Electricity Industry" NBER Working Paper 6463 (1998)

Crawford, Gregory S., Crespo, Joseph, Tauchen, Helen "Bidding Asymmetries in Multi-Unit Auctions: Implications of Bid Function Equilibria in the British Spot Market for Electricity" mimeo (2000)

Crespo, Joseph G. "Theories of Oligopoly in Multi-Unit Auctions" mimeo, University of North Carolina at Chapel Hill (2000)

Engelbrecht-Wiggans, Richard and Kahn, Charles M. "Multi-Unit Auctions With Uniform Prices" *Economic Theory* 12 (1998): 227-258

Green, R. "Increasing Competition in the British Electricity Spot Market" *Journal of Industrial Economics* 44(2) (1996): 205-216

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

Joskow, Paul, Kahn, Edward "A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market During Summer 2000" mimeo, 2000

Klemperer, Paul D. and Meyer, Margaret A. "Supply Function Equilibria in Oligopoly under

Uncertainty" *Econometrica* 51 (Nov., 1989): 1243-1277

McAdams David "Collusive-Seeming Equilibria in the Uniform-Price Auction" mimeo, 2000

Milgrome, Paul R., Weber, Robert J. "A Theory of Auctions and Competitive Bidding" *Econometrica* 50(5) (Sept., 1982): 1089-1122

Puller, Steven L. "Pricing and Firm Conduct in California's Deregulated Electricity Market" mimeo, University of California in Berkeley (2001)

Wolak, Frank A. "An Empirical Analysis of the Impact of the Hedge Contracts on Bidding Behavior in a Competitive Electricity Market" NBER working paper 8212 (April 2001)

Wolfram, Catherine D. "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) (1998): 703-725

Wolfram, Catherine D. "Measuring Duopoly Power in the British Electricity Spot Market" *American Economic Review* (1999)

## Appendix

### Optimal unilateral withholding, general case

Suppose the industry marginal cost is  $mc(q)$ , then the generator  $i$ 's marginal cost, whose market share is  $\beta_i$  is  $mc\left(\frac{q}{\beta_i}\right)$ . Generator  $i$  faces the demand  $D\beta_i$ . When he withholds capacity  $\Delta q$ ,  $\Delta q\beta_i$  of this quantity is to the right of the individual demand line and  $(1 - \beta_i)\Delta q$  to the left. The extra revenue earned is

$$\Delta R = (D\beta_i - \Delta q(1 - \beta_i))(mc(D + \Delta q) - mc(D))$$

And the sacrificed profit is

$$\Delta C = (1 - \beta_i)mc(D)\Delta q - \int_{D\beta_i - \Delta q(1 - \beta_i)}^{D\beta_i} mc\left(\frac{d}{\beta_i}\right) dd = (1 - \beta_i)mc(D)\Delta q - \beta_i \int_{D - \frac{1 - \beta_i}{\beta_i} \Delta q}^D mc(t) dt$$

Marginal revenue and marginal cost are

$$\begin{aligned} \Delta R' &= -(1 - \beta_i)(mc(D + \Delta q) - mc(D)) + mc'(D + \Delta q)(D\beta_i - \Delta q(1 - \beta_i)) \\ \Delta C' &= (1 - \beta_i)mc(D) - (1 - \beta_i)mc\left(D - \frac{1 - \beta_i}{\beta_i} \Delta q\right) \end{aligned}$$

Thus, the first order condition is:

$$\left(D - \frac{1 - \beta_i}{\beta_i} \Delta q\right) mc'(D + \Delta q) - \frac{1 - \beta_i}{\beta_i} \left(mc(D + \Delta q) - mc\left(D - \frac{1 - \beta_i}{\beta_i} \Delta q\right)\right) = 0$$

### Constant relative risk aversion marginal costs

In the last formula of the previous section make the following substitution:

$$\begin{aligned} y &= D + \Delta q \\ z &= \frac{\Delta q}{\beta_i} \end{aligned}$$

The result of this substitution is

$$mc'(y)(y - z) = \frac{1 - \beta_i}{\beta_i} (mc(y) - mc(y - z))$$

Now with the RRA marginal cost  $mc(Q) = AQ^\epsilon + B$  if one assumes  $z$  proportional to  $y$  as

$$z = (1 - \alpha_i)y,$$

the formula boils down to the condition for  $\alpha_i$

$$\epsilon\alpha_i = \frac{1 - \beta_i}{\beta_i}(1 - \alpha_i^\epsilon)$$

### Iterative algorithm to account for the unit size

Instead of a unique threshold demand level as in the symmetric case, in the asymmetric case we need to construct an algorithm in order to understand at what level of demand each firm starts withholding. For that purpose, the following iterative algorithm is run at every demand level. The average unit size at the demand level  $D$  is  $\Delta(D)$ .

Step 0.  $J_0 = 1, \dots, N$ , where  $N$  is the number of firms.

Step 1. Vector of withholdings  $\Delta q_0$  is found from the linear equation  $A_{J_0}\Delta q_0 = B_{J_0}$  (see section 2.7)

Step 2. If  $\min \Delta q_0 < \Delta(D)$

$$\Rightarrow J_1 = J_0 \operatorname{argmin} \Delta q_0$$

$$\Delta q(\operatorname{argmin} \Delta q_0) = 0$$

Step 3. Continue to Step 1. Matrix  $A_{J_1}$  will be a submatrix of matrix  $A_{J_0}$  where only columns and rows with the indices  $J_1$  are taken from initial matrix  $A$ . Stop when either  $J_k = J_{k-1}$  or when  $k = N$ .

Figure 1. Marginal cost bidding

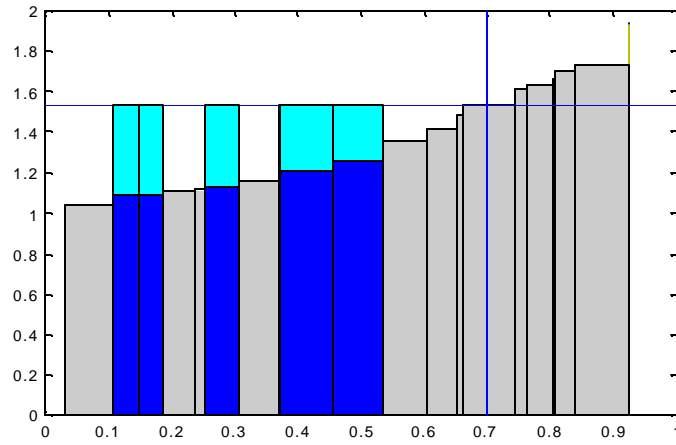


Figure 2. Result of withholding of the most expensive inframarginal unit

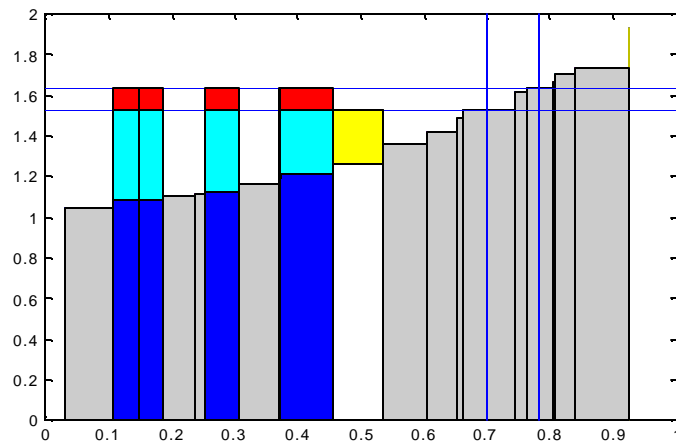


Figure 3. Withholding less expensive unit

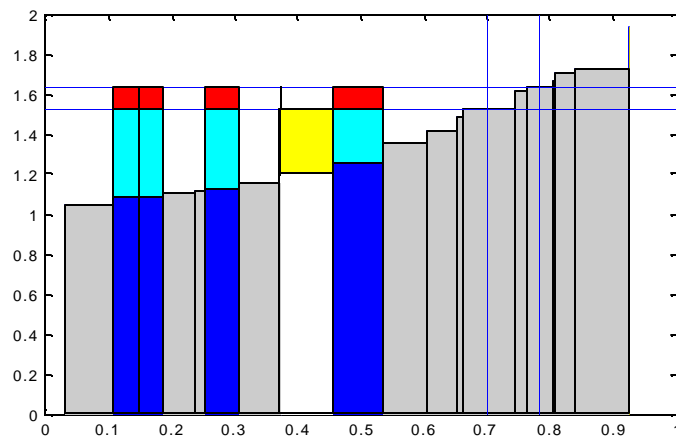




Figure 4. Symmetry

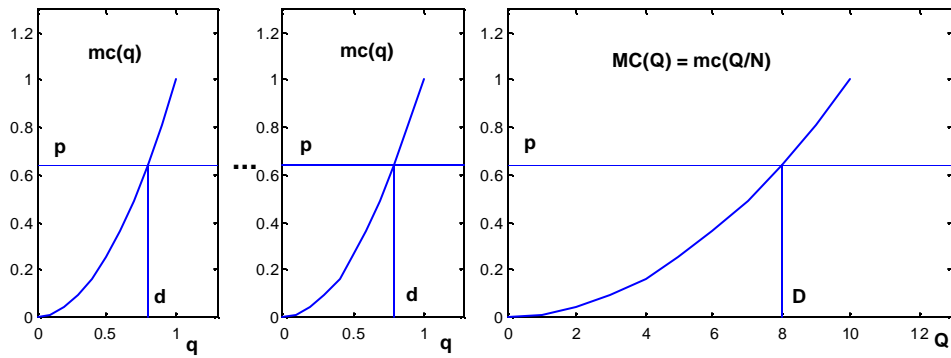


Figure 5. Effect of unilateral withholding

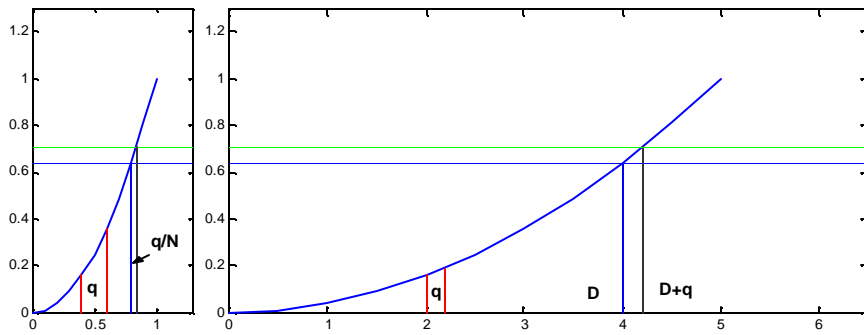


Figure 6. Optimal location of the withheld units

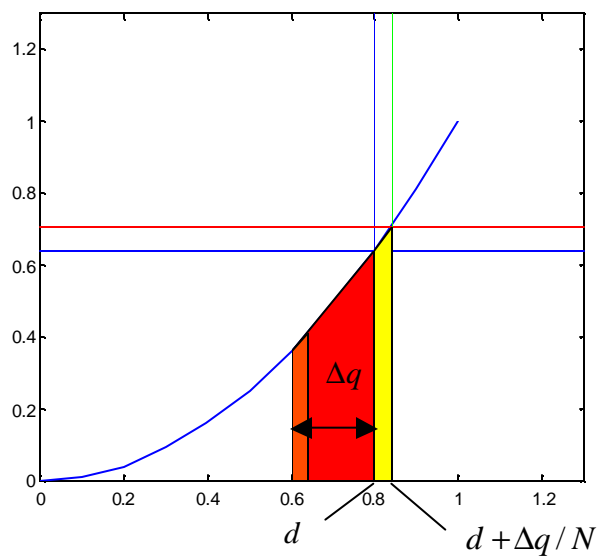


Figure 7. Optimal withholding in the case of linear marginal cost

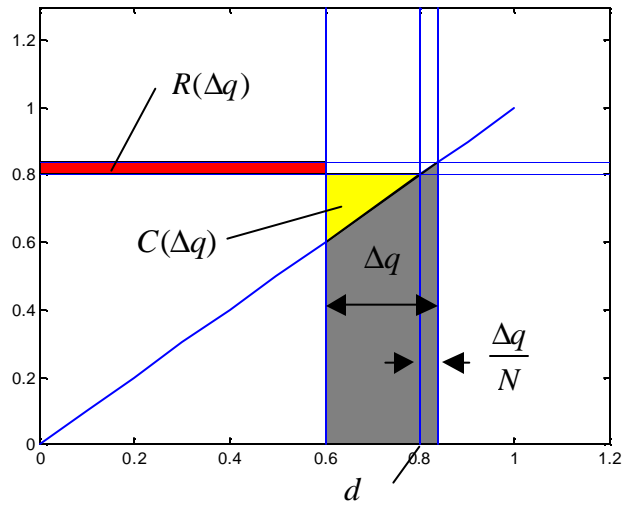


Figure 8. Effective supply functions (linear case,  $N = 6$ )

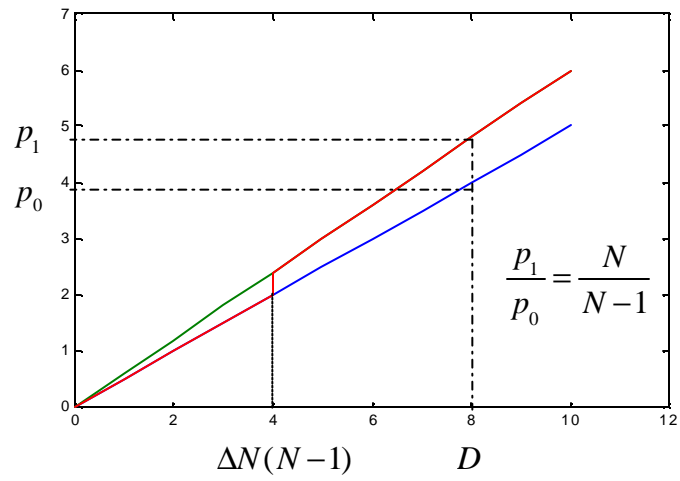


Figure 9. Nonlinear marginal cost function

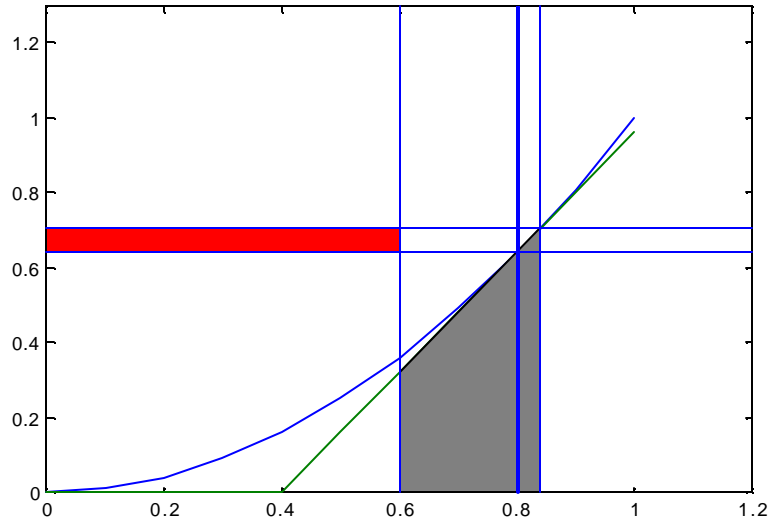


Figure 10. Effective supply (quadratic marginal cost,  $N = 6$ )

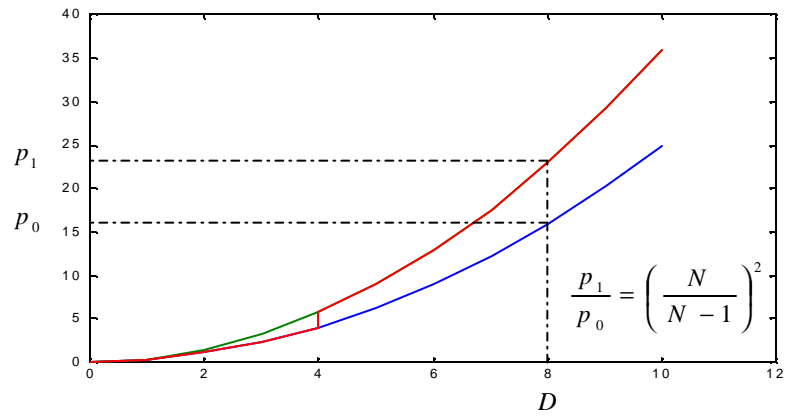


Figure 11. Industry withholding depending on the elasticity of marginal cost and  $N$

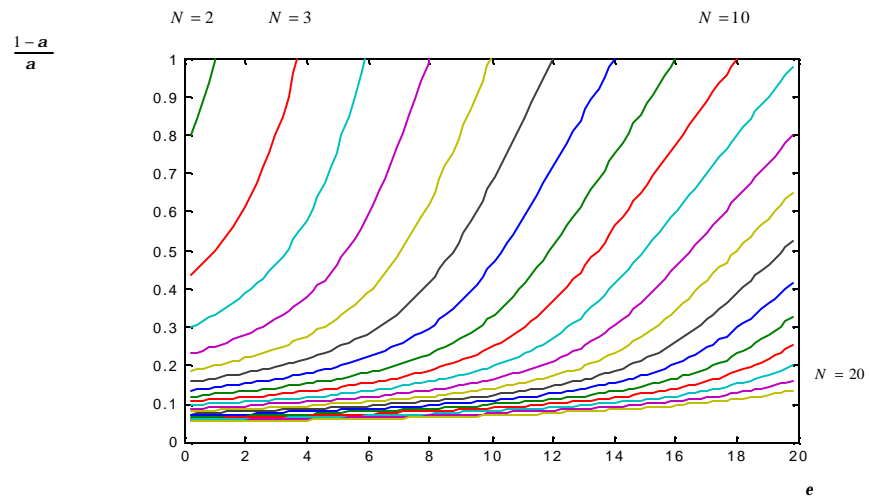


Figure 12. Dependence of withholding to demand ratio on marginal cost heterogeneity

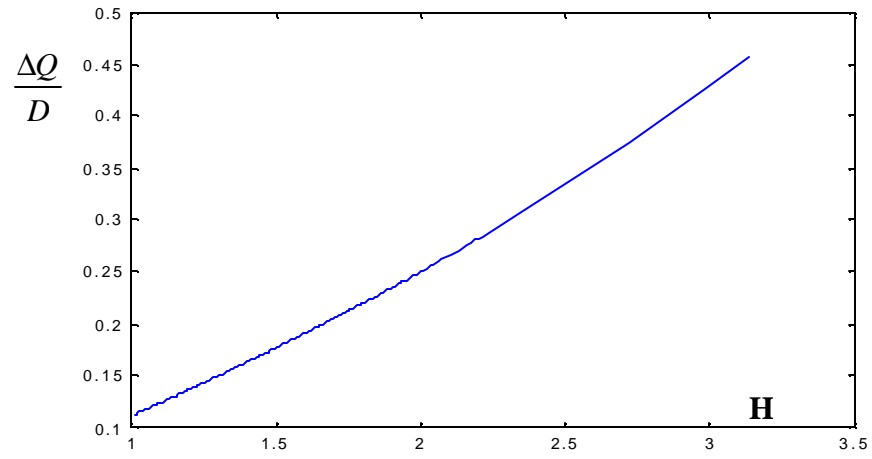


Figure 13. Effective supply with heterogeneous agents

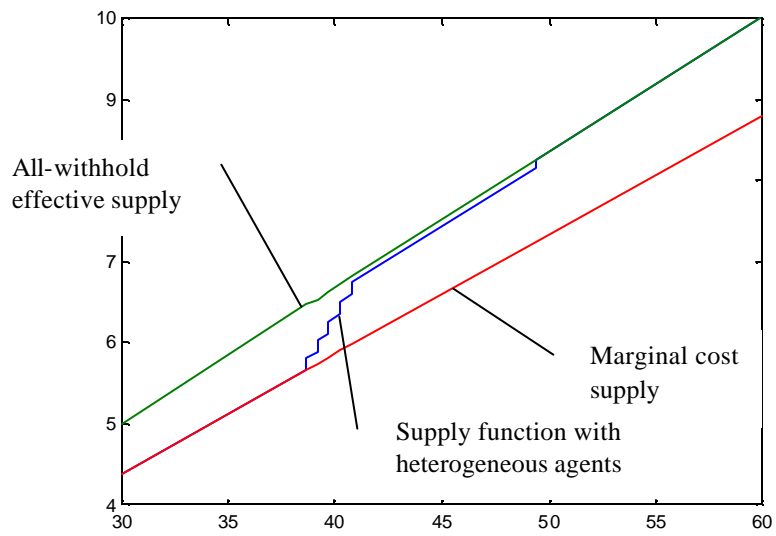


Table 1. CAISO Generation capacity

California ISO	MW	%
Coal	37	0.08%
Biomass	1,041	2.23%
Cogeneration	5,892	12.64%
Geothermal	1,293	2.77%
Hydro	11,801	25.32%
Nuclear	4,358	9.35%
Peaker	2,239	4.80%
Solar	379	0.81%
Thermal	17,690	37.95%
Wind	1,883	4.04%
<b>Total</b>	<b>46613</b>	

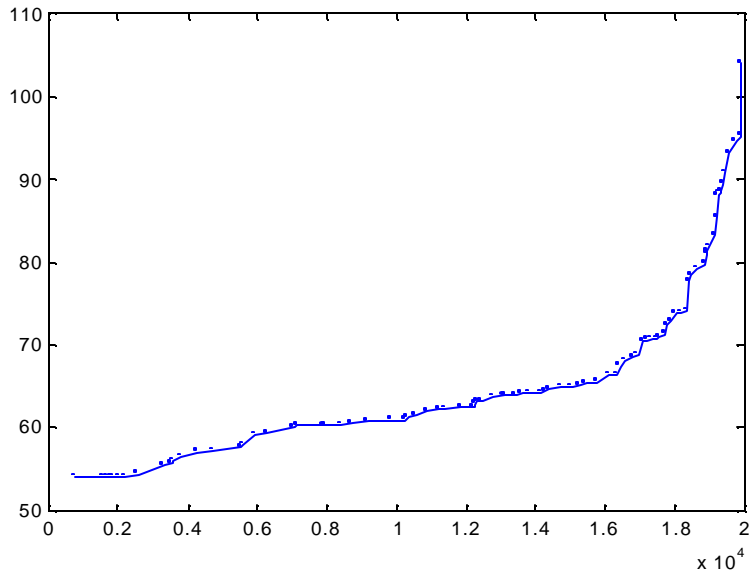
(Source: CAISO 2001 Summer Assessment)

Table 2. Major generation-owning firms

Firm	Capacity	%
AES	3,828.9	19.21%
DUKE	2,257.0	11.33%
DYNERGY	1,998.5	10.03%
RELIANT	3,530.0	17.71%
SOUTHERN	2,997.9	15.04%
<b>Total</b>	<b>19,929.0</b>	

(Source: California Power Exchange capacity database <http://www.calpx.com/>)

Figure 14. Marginal cost of gas units (\$/Mcf)



(Source: EIA Form-767)

Figure 15. Coefficient of relative risk aversion of the marginal cost curve

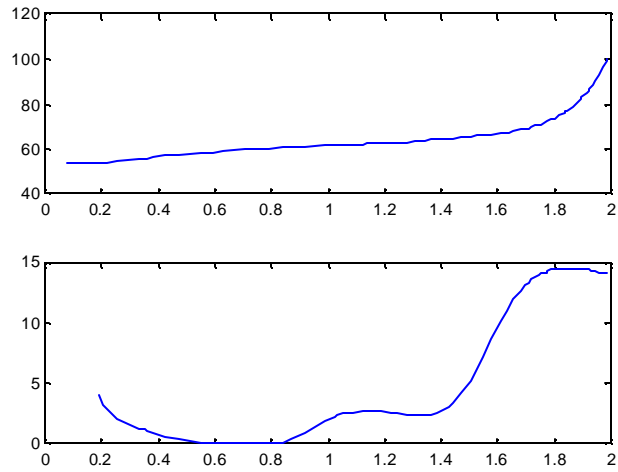


Figure 16. Total withholding as a function of demand faced by gas units (continuous case)

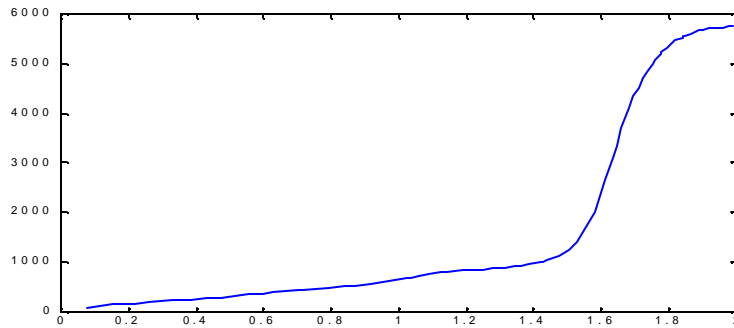
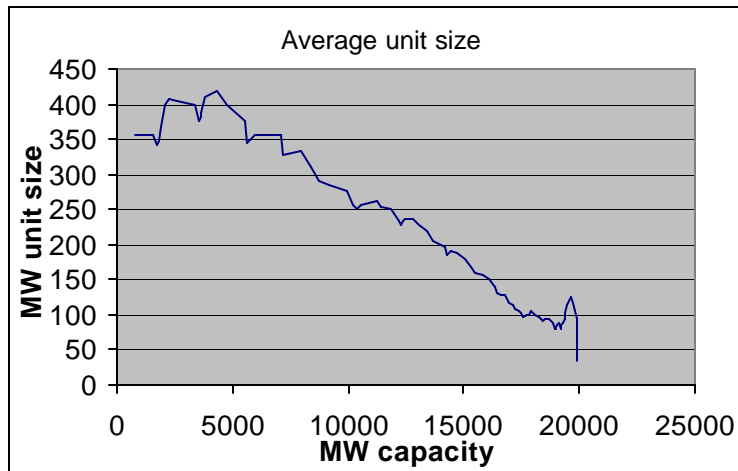


Figure 17. The average unit size as a function of demand faced by gas units



(Source: EIA Form-767)

Figure 18. Withholding schedule accounted for the unit size

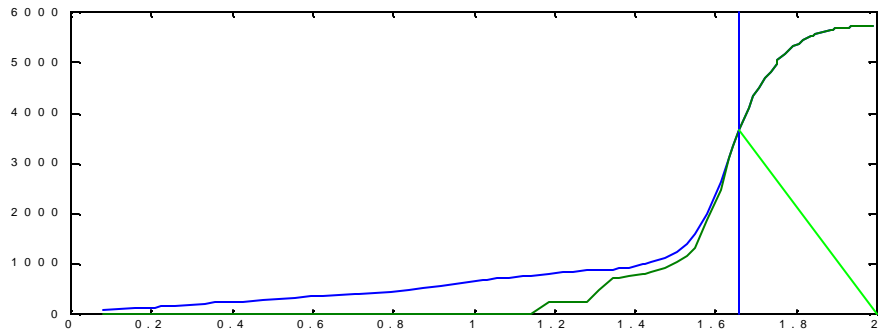


Figure 19. Supply under the June 19 FERC Order

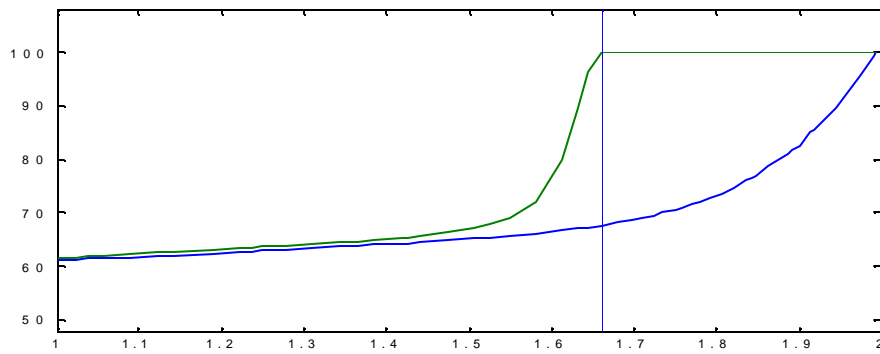
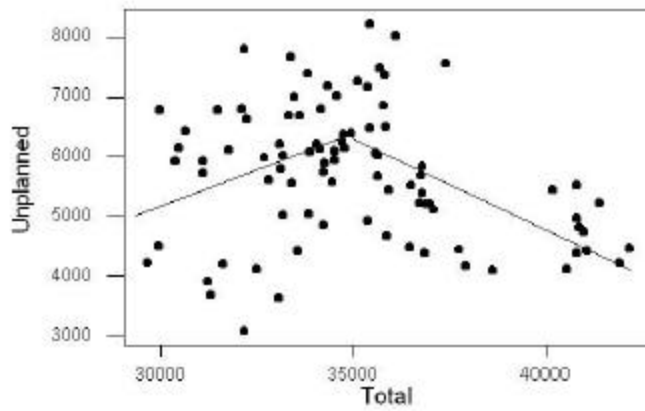


Figure 20. Unplanned non-operational generation 01/01/01 – 06/05/01



(Source: California ISO [www.caiso.com](http://www.caiso.com))