<table>
<thead>
<tr>
<th>Title</th>
<th>Empirical analysis of the spot market implications of price-elastic demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors(s)</td>
<td>Siddiqui, Afzal S.; Bartholomew, Emily S.; Marnay, Chris</td>
</tr>
<tr>
<td>Publication date</td>
<td>2004</td>
</tr>
<tr>
<td>Series</td>
<td>Centre for Financial Markets working paper series; WP-04-03</td>
</tr>
<tr>
<td>Publisher</td>
<td>University College Dublin. School of Business. Centre for Financial Markets</td>
</tr>
<tr>
<td>Link to online version</td>
<td><a href="http://www.ucd.ie/bankingfinance/docs/wp/SIDDIQUI1.pdf">http://www.ucd.ie/bankingfinance/docs/wp/SIDDIQUI1.pdf</a></td>
</tr>
<tr>
<td>Item record/more information</td>
<td><a href="http://hdl.handle.net/10197/1157">http://hdl.handle.net/10197/1157</a></td>
</tr>
</tbody>
</table>
Empirical Analysis of the Spot Market Implications of Price-Elastic Demand

Afzal S Siddiqui*    Emily S Bartholomew    Chris Marnay
University College Dublin    Berkeley Lab    Berkeley Lab

Abstract

Regardless of the form of restructuring, deregulated electricity industries share one common feature: the absence of any significant, rapid demand-side response to the wholesale (or, spot market) price. For a variety of reasons, electricity industries continue to charge most consumers an average cost based on regulated retail tariff from the era of vertical integration, even as the retailers themselves are forced to purchase electricity at volatile wholesale prices set in open markets. This results in considerable price risk for retailers, who are sometimes forbidden by regulators from signing hedging contracts. More importantly, because end-users do not perceive real-time (or even hourly or daily) fluctuations in the wholesale price of electricity, they have no incentive to adjust their consumption in response to price signals. Consequently, demand for electricity is highly inelastic, and electricity generation resources can be stretched to the point where system stability is threatened. This, then, facilitates many other problems associated with electricity markets, such as market power and price volatility. Indeed, economic theory suggests that even modestly price-responsive demand can remove the stress on generation resources and decrease spot prices. To test this theory, we use actual generator bid data from the New York control area to construct supply stacks, and intersect them with demand curves of various slopes to approximate different levels of demand elasticity. We then estimate the potential impact of real-time pricing on the equilibrium spot price and quantity. These results indicate the immediate benefits that could be derived from a more price-elastic demand. Such analysis can provide policymakers with a measure of how effective price-elastic demand can potentially reduce prices and maintain consumption within the capability of generation resources.

1. Introduction

In order to facilitate deregulation, regulatory agencies worldwide have begun to introduce competition into areas of the electricity industry that are technologically amenable to it. This generally implies that the generation and retailing components of the formerly vertically integrated investor-owned utilities (IOUs) are to be provided competitively, while the transmission and distribution sectors continue to be regulated due to their “natural monopoly” characteristics. Most of the reforms, however, target the supply side by attempting to design market rules and structures to induce efficient economic dispatch of generation and allocation of transmission access. Comparatively little deregulation effort has been directed towards ensuring that the demand side is able to respond to market signals. Indeed, this hallmark of most truly competitive commodities markets is missing from markets for electricity. In most cases, this is an artefact of the era of central planning under which consumers were exposed to virtually constant retail rates determined by tariffs that were linked to the costs of the relevant IOU. Such a price-inelastic demand is not viable in tandem with a competitive supply side because it stretches generation resources to the point where system stability is threatened. This rigidity also exacerbates ongoing problems with deregulated electricity industries, such as excessive price volatility and the exercise of market power.

Theoretically, if end-use consumers are exposed to real-time electricity prices, then they can adjust their consumption to reflect market conditions. This then not only reduces the demand for electricity during peak hours, but also lowers the electricity spot price and the need to build more power plants in the long run. For regulators, making end-use consumers price responsive has costs, such as the installation of real-time metering, as well as the aforementioned benefits. Therefore, a policymaking problem is to try to quantify the empirical effects of price-elastic demand on electricity consumption. If exposing relatively few end-use consumers to the spot price can capture most of the benefits from real-time pricing, then the costs of instituting such a programme would be outweighed by its benefits. However, how would price-elastic demand affect electricity consumption in an actual deregulated electricity industry?

*College Lecturer, Department of Banking and Finance, Michael Smurfit Graduate School of Business, University College Dublin, Blackrock, Co. Dublin, Republic of Ireland.
Tel: +353.1.716.8091
Fax: +353.1.283.5482
E-mail: afzal.siddiqui@ucd.ie
We seek to address this question in this paper by undertaking an analysis of the deregulated industry under control of the New York Independent System Operator (NYISO). In particular, we use auction data from the NYISO’s day-ahead electricity markets to construct supply stacks for various zones. We then use a simple linear demand function to approximate price-elastic end-use consumers in order to determine the effect that a certain level of price elasticity has on the equilibrium spot market price and consumption. Our objective is to quantify the benefits from such a pricing protocol and to determine whether most of the objectives, such as lower prices and consumption, can be achieved at modest levels of price elasticity. Towards that end, our paper is organised as follows:

- Section 2 introduces the theory and implications of price-elastic demand
- Section 3 provides an overview of the NYISO control area and insight into the construction of the supply stacks used in our empirical analysis
- Section 4 summarises the methodological framework and the main results
- Section 5 concludes and offers directions for future research in this area

2. Theory of Price-Elastic Demand

From economic theory, the price elasticity of a demand (or supply) curve is the percentage change in the quantity demanded due to a unit percentage increase in the price. Mathematically, if the inverse demand curve as a function of the price is $Q_d(p)$, then the elasticity at some point $(p^*, q^*)$ is $\eta(p^*, q^*) = \left| \frac{p^*}{q^*} \frac{\partial Q_d}{\partial p} \right|$. Specifically, with a linear inverse-demand specification, i.e., if $Q_d(p) = a - bp$, where $a > 0$ and $b > 0$, then $\eta(p, q) = \frac{p}{q}$. From Figure 1, it is possible to determine the impact on the equilibrium price and quantity that changes in elasticity will have. Due to the shape of electricity supply curves, even a modest slope to the demand curve will have a profound impact on the price. For example, a substantial decrease in price can be achieved (from $p_0^*$ to $p_1^*$) if the perfectly price-inelastic inverse demand, $Q_{d0}(p)$, is given a slight slope to rotate it to $Q_{d1}(p)$. Note that further increases in slope achieve only mild decreases in the price (compare $p_1^*$ with $p_2^*$).

![Figure 1. Price-Elastic Linear Demand](image)
In general, the comparative statics associated with changes in \( b \) are not easily computable due to non-linearities in the market equilibrium. For instance, even if the inverse-supply curve, \( Q_s(p) \), is linear\(^1\), then the impact of a change in \( b \) on \( p^*, q^* \), and \( \eta(p^*, q^*) \) is not linear. After some calculation, it is possible to state that with both linear supply and demand curves, the following relations hold:

1. \( \frac{\partial p}{\partial b} < 0 \)
2. \( \frac{\partial^2 p}{\partial b^2} > 0 \)
3. \( \frac{\partial q}{\partial b} < 0 \)
4. \( \frac{\partial^2 q}{\partial b^2} > 0 \)
5. \( \frac{\partial \eta}{\partial b} > 0 \)
6. \( \frac{\partial^2 \eta}{\partial b^2} < 0 \)

The first four state that increasing the slope of the inverse demand curve decreases both the equilibrium price and quantity, albeit at an increasing rate. In other words, the first unit increase in the slope decreases the equilibrium price (or quantity) more than the next unit increase. For the elasticity, the outcome is reversed as increasing the slope increases the resulting price elasticity of demand, but at a decreasing rate.

For electricity industries in particular, the shape of the supply curve is flat over large ranges of quantity before ramping up sharply as the output constraint is reached. This reflects the fact that the marginal cost of supplying electricity increases with production. Indeed, a supply shock or demand surge causes a disproportionate increase in the equilibrium price. In addition, this low price elasticity enables producers to exercise market power more easily. While their ability to undertake such measures can be mitigated by increased forward contracting (see Wolak (2000)), only effective demand-side response can prevent sustained increases in the electricity price.

Greater price responsiveness can be induced through either interruptible load programmes or real-time pricing. In Oren et al. (1987) and Smith (1989), the concept of electricity product differentiation is used to allow utilities to implement a pricing structure in which the probability of outage varies. An analysis of the interruptible load protocol, as implemented in California during 2001, reveals that it was unsuccessful due to the lack of response from consumers as calls for interruption became more frequent (see Marnay et al. (2001)). Instead, real-time pricing directly provides the signals consumers require for adjusting their demand. In Borenstein (2001), a methodology to enable real-time pricing while maintaining stable monthly consumer bills is introduced. Here, hedging is used by the utility to lock in the price, with any gains or losses from its forward position used to adjust the real-time price perceived by end-use consumers accordingly. Therefore, the variability of hourly prices is maintained while removing much of the monthly fluctuations in electricity bills.

In practice, however, the extent to which electricity consumption is price elastic is difficult to measure. An evaluation of the NYISO’s Price-Responsive Load (PRL) programme, in which consumers bid to act as interruptible loads, reveals an average price elasticity of about 0.03 (see Neenan Associates and CERTS (2002)). Surveys indicate that customers were deterred from participation by the severe penalties for non-compliance and by the high degree of perceived risk relative to benefits. A study of the San Diego retail area during 2000 (when retail rates doubled) indicates a price elasticity of approximately 0.06 (see Bushnell and Mansur (2001)). This low value may have been caused by the five-week lag in wholesale price exposure and the promise by politicians to abolish the new pricing regime. Still, because demand is completely price inelastic to begin with and the supply stack curves sharply upward as capacity limits are approached, even a small increase in the responsiveness of consumers may be enough to lower equilibrium prices by a significant amount. The extent of this effect is what we aim to measure empirically in this paper by constructing supply curves for a deregulated electricity industry and then inducing price elasticity in the demand. Before implementing the methodology, we will first discuss the NYISO control area.
3. Overview of the NYISO Control Area

3.1 Market Structure

The New York Independent System Operator (NYISO) manages the electricity grid covering the entire state of New York and runs wholesale electricity markets through which approximately half of the state’s electricity is purchased. The state is divided into eleven load zones (see Figure 2). The price of wholesale electricity at any point in the system is known as the locational-based marginal price (LBMP) and is based on the cost of providing the next increment of load at that point. A LBMP is calculated both for each of the eleven load zones (and becomes the price paid by loads), and for each of over 400 specific generation buses (which is the price paid to generators for producing at that point). When all electricity can be supplied at lowest cost because there is no transmission congestion, the price is almost uniform across the state, varying only because of losses in the grid. Often, different locations have different market-clearing prices because of congestion.

Figure 2. New York ISO Control Area Load Zones

The NYISO runs two financially binding energy markets: the day-ahead market and the real-time market. About 90% of the energy traded in the NYISO wholesale markets is done so in the day-ahead market. Both loads and generators can place bids to buy and sell electricity in the day-ahead market. Generators are allowed to bid either blocks of energy for a given price, or curves defined by price/quantity points, and loads can specify both a fixed bid amount (power they will buy at any price) and a price-capped amount (load they will buy only if the clearing price is at or below a given price). The day-ahead market is a financially binding market.

3.2 Supply and Demand Stacks

The generator offers published with a six-month lag on the NYISO web site are the center of the analysis for this paper. Since these offers are published anonymously, it is difficult to determine the specific bus or even zonal location. Therefore, we consider all offers to be in one large market not separated by congestion. For each hour, we first sort the generator bids by offer price and then add them horizontally by calculating the cumulative quantity offered at each price. We then approximate curve offers with 1 MW–wide block offers before they are added to the stack and sorted. We finally intersect the resulting stacks with the amount of generation bids scheduled, as published by the NYISO in the Day-Ahead Market.
Energy Report on its web site, to determine a clearing price for the entire market at that hour. An example of how these supply stacks were constructed can be seen below in Figure 3.

In creating these offer stacks and identifying the market-clearing price, we make a few assumptions. First, we ignore minimum energy bids, which, for this analysis, is not an unreasonable assumption because we are focusing on the higher quantity end of the offers, not at start-up costs. We also do not take into account minimum run-times, start-up costs, unit-commitment considerations, security constraints, or other factors that may result in dispatching units out of merit order. Finally, we do not address network topology and possible congestion that leads to congestion pricing on the grid.

![Figure 3. Constructing the Supply Stack](image)

4. Empirical Methodology and Results

The objective of this paper is to estimate empirically the impact of a change in the slope of the demand curve on the equilibrium price, quantity, and elasticity \( \frac{\partial p}{\partial b} \), \( \frac{\partial q}{\partial b} \), and \( \frac{\partial \eta}{\partial b} \), respectively. The former two indicate the immediate benefits of a more price-elastic demand, while the latter determines the longer-term ability of end-users to respond to supply shocks. In order to measure the benefits of any potential real-time pricing programs, the rates of change of the impact of the change in slope can also be calculated, viz., \( \frac{\partial^2 p}{\partial b^2} \), \( \frac{\partial^2 q}{\partial b^2} \), and \( \frac{\partial^2 \eta}{\partial b^2} \). These parameters have policymaking implications for real-time demand-side responsiveness because they can be used to determine the level of price elasticity that is sufficient to guarantee a certain percentage decrease in the equilibrium price or quantity.

The empirical estimation of these quantities for a given hour in a geographical region of NYISO relies on the availability of supply stack and market equilibrium data. Since these data are generally widely available, the estimation procedure can be implemented. As a first step, we construct the (non-linear) supply stack from the given data. Then, we intersect a perfectly price-inelastic, i.e., vertical, demand with the supply stack at the given market equilibrium as in Figure 1. This initial equilibrium, \( (p^*_0, q^*_0) \), is the result of such an intersection. Note that the elasticity associated with this equilibrium, \( \eta(p^*_0, q^*_0) \), is zero, thereby implying that the consequences of any supply shock are incident entirely upon the consumers through a proportionally higher price. In the next step, we give the inverse demand curve a slight slope, \( \frac{1}{b_1} \), so that it now becomes \( Q_{d1}(p) \) and attains a new equilibrium at \( (p^*_1, q^*_1)^2 \). We iterate this process for \( n - 1 \) different linear demand curves after the initial one, each more elastic than the previous one so that the \( (n - 1) \)st sloped curve intersects the supply curve at its lowest step, i.e., \( 0 = b_0 < b_1 < \cdots < b_i < \cdots < b_{n-1} = \min \left\{ b \left| a - bp = \min_p Q_i(p) \right. \right\} \). For each linear inverse demand curve, \( Q_{di}(p) \), we compute the equilibrium price, quantity, and elasticity \( (p^*_i, q^*_i) \), and \( \eta(p^*_i, q^*_i) \), respectively.
In particular, \( b_0 = 0 \) and \( b_j = \min \left\{ b \left| \frac{a - k_{n-j}}{b} = c_{n-j} \right. \right\} \Rightarrow b_j = \frac{a - k_{n-j}}{c_{n-j}}, 1 \leq j \leq n - 1 \), where the supply curve has the following form for \( c_1 \leq c_2 \leq \cdots \leq c_n \):

\[
P_s(q) = \begin{cases} 
  c_1 & \text{if } k_0 \leq q < k_1 \\
  c_2 & \text{if } k_1 \leq q < k_2 \\
  \vdots \\
  c_n & \text{if } k_{n-1} \leq q < k_n 
\end{cases}
\]

Figure 4 summarises the procedure of calculating the equilibrium prices and quantities with varying demand elasticities for a supply step function with three increments, where \( P_{di}(q) \) are the standard linear demands. Here, \( b_0 = 0 \Rightarrow p_0^* = c_3 \) with \( q_0^* = a, b_1 = \frac{a - k_2}{c_2} \Rightarrow p_1^* = c_2 \) with \( q_1^* = k_2 \), and \( b_2 = \frac{a - k_1}{c_1} \Rightarrow p_2^* = c_1 \) with \( q_2^* = k_1 \).

Using the calculated equilibria and assuming that the iteration intervals are small relative to the initial equilibrium quantity, we estimate the impact of changes in the slope on market attributes as follows:

1. \[
\frac{\partial p}{\partial b_{(p_i^*, q_i^*)}} \approx \lim_{\Delta b_i \to 0} \frac{\Delta p_i^*}{\Delta b_i} \approx \frac{p_i^* - p_{i-1}^*}{b_i - b_{i-1}}, i = 0, \ldots, n - 1
\]
2. \[
\frac{\partial q}{\partial b_{(p_i^*, q_i^*)}} \approx \lim_{\Delta b_i \to 0} \frac{\Delta q_i^*}{\Delta b_i} \approx \frac{q_i^* - q_{i-1}^*}{b_i - b_{i-1}}, i = 0, \ldots, n - 1
\]
3. \[
\frac{\partial \eta(p, q)}{\partial b_{(p_i^*, q_i^*)}} \approx \lim_{\Delta b_i \to 0} \frac{\Delta \eta(p_i^*, q_i^*)}{\Delta b_i} \approx \frac{\eta(p_i^*, q_i^*) - \eta(p_{i-1}^*, q_{i-1}^*)}{b_i - b_{i-1}}, i = 0, \ldots, n - 1
\]

The derivatives of these quantities can also be estimated via a similar procedure. Together, such estimations will provide policymakers and analysts with a measure of how effective price-elastic demand can be in reducing prices and maintaining electricity consumption within the capability of generation resources.

With NYISO data for the year 2002, we estimate the effect of the slope on the equilibrium price, quantity, and elasticity. From a policymaking perspective, these quantities can be used to determine how large of a price elasticity is necessary to induce a certain decrease in the market-clearing price or quantity of electricity. Before presenting the summary statistics, we outline the significance of our results through a case study for hour 14 on 08 August 2002. Using the bid data, we first construct the supply stack for the hour (see Figure 5). Notice how the curve slopes upward sharply as the supply capacity is reached at around 28000 MW. Next, we use the fact that the market-clearing price and quantity for this hour are US$55.68/MWh and 20152 MW, respectively, to determine how the price, quantity, and elasticity would change as the linear inverse demand is progressively given a larger slope as indicated in Figure 4. By plotting the resulting market-clearing price and quantity versus the elasticity (see Figure 6 and Figure 7, respectively), we obtain the empirical effect of greater elasticity.

In particular, the data confirm the theoretical results that \( \frac{\partial p}{\partial b} < 0, \frac{\partial^2 p}{\partial b^2} > 0, \frac{\partial q}{\partial b} < 0, \) and \( \frac{\partial^2 q}{\partial b^2} > 0 \). Indeed, as the slope (or, elasticity) increases, the market-clearing price and quantity decrease but at a diminishing rate. Therefore, from Figure 6 and Figure 7, a policymaker can immediately learn the level of demand elasticity that is necessary to reduce the price or quantity by a specific amount. For example, in order to reduce the price by 25% to US$41.76, a price elasticity of 0.179 is required. Consequently, the market-clearing quantity for this level of price responsiveness would be 17087 MW (about 15.2% lower).
With hourly data for the entire year 2002, we then calculate summary statistics that indicate the average price elasticity necessary to reduce the market-clearing price by a given percentage in each hour (see Table 1). The average market-clearing price and quantity during this year were US$50.54/MWh and 15804 MW, respectively. In order to reduce the price by 25% in a given hour, for example, we slope the inverse demand curve until it intersects the supply stack for that hour at the desired price. We then record the level of elasticity that makes this price reduction possible as well as the corresponding market-clearing quantity at that price. If we do this for each hour, then we obtain the summary statistics as in Table 1. Here, an average price elasticity of 0.23 is necessary to reduce the market-clearing price by 25%. At this price, the average market-clearing quantity of electricity is 13004 MW, or an 18% reduction on average from its original value. The calculations for the 50% and 75% reduction in price scenarios follow accordingly.

Although the levels of price elasticity necessary to obtain significant percentage decreases in the market-clearing price seem to be beyond the scope of what were experienced in NYISO and San Diego (0.03 and 0.06, respectively), it should be noted that the programmes implemented in these areas provided only distorted price signals. Indeed, they were not the envisioned real-time pricing protocols. In Borenstein et al. (2002), analysis of a successful real-time pricing protocol managed by the Georgia Power Company reveals price elasticities of 0.20 and 0.28 at moderate and high prices, respectively, for large customers (with loads greater than 5000 kW) facing hour-ahead prices. While most of the customers face day-ahead prices and are relatively inflexible, i.e., with elasticities between 0.02 and 0.06, the presence of large, price-responsive customers might be enough to maintain the market-clearing price and quantity of electricity within manageable ranges. In fact, most of the benefits of price responsiveness may be obtained by introducing even such a limited real-time pricing programme: in Table 1, reduction of the market-clearing price by 50% requires a level of price elasticity (0.53) that is unlikely to exist in most electricity industries. Hence, the policymaking implication from this analysis is that most of the feasible benefits of price-responsive electricity demand can be captured by targeting the larger customers (e.g., industrial or commercial ones) for participation in real-time pricing programmes. Any effort directed towards smaller customers will not make much impact because their price elasticities are too low to reduce the price even by 5%. Our analysis quantifies this conventional wisdom by indicating precisely how much the price can be reduced for each marginal increase in the level of elasticity.
Figure 5. NYISO Supply Stack for 08 August 2002 (1400)

Figure 6. Effect of Demand Elasticity on the Market-Clearing Price for 08 August 2002 (1400)
5. Conclusions

Electricity industries worldwide suffer from a lack of demand response. Although electricity is theoretically an inelastic good in the short-run, the steep slope of the supply stack implies that even modest response by demand could translate into significant price decreases. The price elasticity of demand is, therefore, important in determining the success of any real-time pricing programme. In this paper, we attempt to quantify to what extent a given level of price elasticity in the NYISO control area during 2002 would affect the market-clearing price and quantity. First, we use publicly available data to construct supply stacks for NYISO at each hour of the year. We then use the market-clearing price and quantity during that hour to anchor a perfectly price-inelastic linear demand curve. Next, we induce price elasticity in the demand curve by varying its slope. We calculate the new market-clearing price and quantity at each interval, which are then used to determine the price elasticity of demand at various output levels. Using these estimates, we finally obtain the average level of price elasticity that would be needed to reduce the average market-clearing price during the year by a certain percentage. In particular, we find that an average price elasticity of 0.23 would be necessary for the average price to be reduced by 25%.

The policymaking consequence of this research is that for any desired reduction in the market-clearing price or quantity, the corresponding price elasticity can be ascertained. A proposed real-time pricing programme can next be targeted towards those customers whose degree of price responsiveness is large enough to increase that of the entire load to the desired level. In our case, we can decide exactly beyond which level it is infeasible to extend the scope of the real-time pricing programme in order to pursue any further reduction in the market-clearing electricity price. Indeed, the remaining customers would have
levels of elasticity much lower than those required to translate into subsequent price decreases. Hence, through our analysis, the relative benefits of price elasticity can be quantified as a decision-analysis tool for policymakers.

For future research, we would like to extend our framework to allow for non-linear demand curves. We could also use the results to examine what degree of price responsiveness is necessary to mitigate the impact of any supply shocks. Since interruptible load programmes have been used extensively in both NYISO and California, it would also be insightful to analyse their performances vis à vis that of real-time pricing programmes. Finally, we would like to analyse the forward market implications of price-elastic demand. Although price responsiveness unambiguously reduces the spot market price and quantity, its effect on the forward price is not so clear (see Siddiqui (2004)). In fact, the forward price can be either decreased (due to the resulting lower spot price) or increased (due to the increased covariation of retailers’ revenues with the spot price). Using data from markets that have implemented real-time pricing programmes, we can estimate statistically the conditions under which each effect dominates.

Bibliography


✝ This work was supported by funding from the Business Research Programme at University College Dublin. The authors would like to thank Meredith Fowlie and Celeste Saravia of the University of California Energy Institute and Bernard Lesieutre of Berkeley Lab for their comments and suggestions.

1 This reflects linear marginal costs, or equivalently, quadratic costs, which ignores the economics of electric power generation.

2 The new equilibrium is again determined by intersecting the supply stack with the new demand curve.

3 The curves are truncated for price less than US$-200/MWh in order to preserve the scale.

4 Of course, the actual price elasticity needed in a given hour may be more or less than this average value, as indicated by the standard deviation.