Industrial and Commercial Customer Response To Real Time Electricity Prices

Richard Boisvert, Peter Cappers, Bernie Neenan, and Bryan Scott*

December 10, 2004

• R. Boisvert is a Professor in the Department of Applied Economics and Management at Cornell University, and an Academic Associate with Neenan Associates. P. Cappers is a Senior Analyst with Neenan Associates and B. Neenan is President of Neenan Associates. B. Scott is Vice President of BBCI. The authors express their gratitude to CSW, and especially Don Moncrief, for allowing the result of this program to be made public. Any errors or misrepresentations are solely the responsibility of the authors.

Table of Contents

Industrial and Commercial Customer Response To Real Time Electricity Prices

Richard Boisvert, Peter Cappers, Bernie Neenan, and Bryan Scott*

December 10, 2004

Executive Summary

This paper uses hourly data from industrial and commercial customers that volunteered to participate in Central and Southwest Service's (CSW) real-time pricing (RTP) programs to characterize the induced price response. The CSW RTP program, which adopted the two-part, revenue neutral design first introduced by Niagara Mohawk, was introduced to large commercial and industrial customers in Oklahoma 1994, and later in other jurisdictions. This study used summer months' price and usage data for the period 1998-2001. Most of the participants had gained substantial experience managing load under RTP by 1998, and this period was characterized by considerable price volatility.

CSW allowed customers to choose from two RTP designs. Under the RTP program, a two-part, revenue neutral design, nearly identical to that introduced by Niagara Mohawk in 1988, participants received an hourly price schedule by 4:00 the day before it was effective. The RTP prices were used to settle variations in the following day's hourly energy usage from the customer's CBL. The CBL is an hour-by-hour representation of what the customer is deemed to have consumed on the standard CSW tariff otherwise applicable to the customer. Following established practice, CSW

* R. Boisvert is a Professor in the Department of Applied Economics and Management at Cornell University, and an Academic Associate with Neenan Associates. P. Cappers is a Senior Analyst with Neenan Associates and B. Neenan is President of Neenan Associates. B. Scott is Vice President of BBCI.

established the CBL from each customer's historic usage. By construction, under this design participants are fully hedged against hourly RTP price variations at their typical load, but face prices that induce changes that result in economic efficiency improvements.

Alternatively, customers could elect the RTP-LR option, whereby they nominated some of the CBL for additional, short-term hourly price exposure in return for a corresponding reduction in the tariff demand charge. For these participants, the day-ahead prices were provisional. CSW could, within specified limits, adjust their hourly prices upward by $38/kWh with only a single hour's notice, and simultaneously reduce their CBL by the amount of nominated load. Since these customers face greater price exposure, they were expected to be more price responsive.

The price and quantity records of 54 participants were analyzed, of which 39 were enrolled in RTP for all three years, nine were in a program for two of the three years and six participated for a single year. A total of 43 customers enrolled in the RTP program, while 11 customers enrolled in a companion RTP-LR program.

The study makes two important contributions to understanding of RTP participants' response to RTP prices. First, by adopting a demand model that allows price response to vary with the level of prevailing prices, we are able to identify the extent to which a firm's ability to shift electricity between high and low priced periods increases as the peak price increases. CSW's two-part RTP rate design ensures that customers are revenue neutral if they continue to consume at their customer base load (CBL), which provides a second opportunity. By comparing actual usage with that

---

The authors express their gratitude to CSW, and especially Don Moncrief, for allowing the result of this program to be made public. Any errors or misrepresentations are solely the responsibility of the authors.

specified by the CBL, we are also able to provide some preliminary estimates of the effect of these rates on energy conservation relative to shifting load to off-peak periods. To estimate price response, we employ a Generalized Leontief demand relationship that yields substitution elasticities that vary with the level of the peak to off-peak price ratio. This allows for testing whether customers are more responsive at higher prices than at lower prices. The substitution elasticity measures the change in the ratio of daily peak to off-peak usage in response to changes in the off-peak to peak price.

To account for behaviors other than inputs substitution, we specified a separate relationship to distinguish the substitution of off-peak for peak electricity usage from the curtailment of discretionary usage in response to peak prices, a form of conservation or own-price elasticity. A recent study suggests that many customers facing RTP-type pricing respond by reducing discretionary usage (Goldman *et. al.* 2004).

Customer responses were substantial; responses differed substantially in response to price, and under certain conditions. On average, the substitution elasticity for RTP customers was in the range of 0.10 to 0.18, while those of RTP LR participants were almost twice as high, ranging from 0.20 to 0.27. The estimated elasticities of substitution were found to vary directly with the level of the ratio of peak and off-peak prices. This result would seem to rationalize the RTP LR program option; achieving response to real-time system conditions requires a substantial price inducement, as has been observed elsewhere (Neenan Associates and CERTS, 2002).

Price response is highest for high prices of short duration, and falls rather dramatically as the duration of high prices increases. One explanation for this response fatigue is that at some point in time the costs firms incur in curtailing exceed the RTP incentive to curtail, and they restore electricity usage to more typical levels.

About 75% of the responses were found to be due to load shifting from peak to off-peak periods, rather than due to overall energy conservation. This is in contrast to a similar analysis of RTP behavior, where almost 75% of the responses were due to the reduction of discretionary usage (Goldman *et al.*, 2004). Differences in the program populations and in the RTP design may explain this result.

The policy implications of this analysis seem clear. There are marked differences in price responsiveness across customer groups and peak periods. One RTP design will not access the full load management capability. A portfolio of pricing products with different lengths of exposure to high prices maybe the most cost effective way to balance system supply considerations with customers' capabilities to curtail load. Furthermore, since load reduction on peak increases proportionally more when price differences are high**,** rate designs, and particularly those based on critical peak pricing, it may be a more cost effective way to provide curtailments in response to abrupt changes in the availability of operating reserves. The day-ahead RTP program is better suited for controlling price volatility that arises from scheduling imbalances of supply and demand.

Industrial and Commercial Customer Response to Real Time Electricity Prices

Richard Boisvert, Peter Cappers, Bernie Neenan, and Bryan Scott*

December 10, 2004

Introduction

Initiatives to restructure electricity markets have created renewed interest in dynamic pricing. Time-of-use (TOU), real-time pricing (RTP), and interruptible/curtailable rate programs initiatives launched in the 1980's were based only tangentially on economic efficiency arguments. Policy makers saw dynamic pricing as a complement to conservation strategies, and regulated utilities relied on them as a way to defer investments in additional capacity needed to meet established reliability criteria. During that period, many utilities experimented with these kinds of rates, but they were never adopted widely by the industry. Selected real-time pricing programs, such as those by Georgia Power and Duke Power, have persisted. Others, such as the TOU and RTP programs of Central and South West Services (CSW) were in place throughout much of the 1990's (Long, *et al*. 2000). However, the majority of retail electricity customers still buy electricity at fixed prices known well in advance.

Currently, the need for greater direct market participation by retail customers is becoming increasingly clear as wholesale electricity markets move forward. By allowing a larger number of customers to adjust demand in response to high prices, some argue that these retail pricing programs will help discipline the markets, both in terms of

* R. Boisvert is a Professor in the Department of Applied Economics and Management at Cornell University, and an Academic Associate with Neenan Associates. P. Cappers is a Senior Analyst with Neenan Associates and B. Neenan is President of Neenan Associates. B. Scott is Vice President of BBCI..
The authors express their gratitude to CSW, and especially Don Moncrief, for allowing the result of this program to be made public. Any errors or misrepresentations are solely the responsibility of the authors.

lowering average prices and reducing price volatility. Achieving this goal requires the deployment of a broad portfolio of retail rate options that exhibit substantial time differentiation in usage rates to accommodate extensive customer and end-use diversity, while providing profit opportunities for retailers. Clearly, designing rates that will meet the needs of diverse customers requires a thorough understanding how customer use and value electricity, and characterization of their ability and inclination to shift usage from high priced to lower priced periods.

This knowledge is critical to determining if the benefits from these retail rates can outweigh the additional administrative, equipment, and metering costs associated with their implementation. These issues are certainly not new, but the distribution of the costs and benefits of such retail programs will change under competition. However, since these retail programs have never been adopted widely, and customer data on existing programs remains largely proprietary; there have been few systematic studies of the ability of industrial and commercial customers to alter electricity usage in response to price changes.[1] This study contributes to the understanding of how customers respond to dynamic pricing of electricity.

Objectives

This paper uses extensive hourly data from industrial and commercial customers that participated in Central and Southwest Services (CSW) real-time pricing programs to estimate demand response to hourly retail electricity prices posted a day ahead. This is accomplished by quantifying participants' ability to shift electricity between high and

[1] The papers by Herriges *et al*. (1993), Mountain and Hsiao. (1986), Ham *et al*. (1997), Patrick et. al. 1997, and Schwarz *et al*., 2002 are examples of the few of the studies available.

low priced periods.[2] The analysis of this data is particularly significant since most participants were on the RTP rates for two or more years in the late 1990s during which high prices resulted from the extremely hot weather. Further, CSW customers were enrolled in one of two programs, the RTP program or RTP Load Reduction (RTP-LR). In the latter, customers were paid an option payment ($/kW-year) for load pledged to the program and the day-ahead hourly RTP prices were provisional. The utility could, under specified system conditions, increase the RTP price by $0.38, increasing the price these customers paid, thereby increasing their incentive to reduce usage.

The study makes two important contributions to understanding of RTP participants' response to RTP prices. First, adopting a demand model that allows price response to vary with the level of prevailing prices, we are able to identify the extent to which a firm's ability to shift electricity between high and low priced periods increases as the peak price increases. CSW's two-part RTP rate design ensures that customers are revenue neutral if they continue to consume at their customer base load (CBL), which provides a second opportunity.[3] By comparing actual usage with that specified by the CBL, we are also able to provide some preliminary estimates of the effect of these rates on the energy conservation relative to shifting load to off-peak periods.

In the remainder of the paper, a description of the RTP program data used is followed by a discussion of several conceptual issues that need to be resolved in using real time pricing data to estimate substitution elasticities between high and low priced

[2] This study reports the results of programs implemented by CSW in the period 1994-2001. CSW subsequently merged with AEP.

[3] The two-part rate assesses to each customer an access charge calculated by applying the otherwise applicable tariff rate to a customer baseline load (CBL) and then in each hour settling, by either increasing (load growth) or crediting (load reduction) that access charge amount, variations from the CBL at the prevailing RTP prices. The CBL is a hour-by-hour characterization of the participant's typical usage, which CSW developed from historical actual customer usage data.

periods. Next, the formal demand models are specified. The empirical results are presented. The final section of the paper includes summary comments and conclusions.

The Real-Time Pricing Program Data

At CSW, the real-time pricing programs marketed during the 1990's provided its customers with virtual access to market-based prices, since the RTP prices used to settle load variations from the CBL were based on CSW's top-of-stack supply cost. RTP program participants received an hourly price schedule a day in advance, as did RTP-LR participants, but the latter prices were subject to revision with one hour's notice. While this pricing scheme exposed participants to supply cost volatility, and at time prices that exceeded tariff rates by a factor of 10 or more, it also resulted in prices that were below the equivalent tariff prices for over 90% of the hours of the year (Long, *et al*., 2000). For customers that could manage around the high prices, this provided an opportunity to increase load at a lower than tariff cost. For those that wanted to reduce their overall cost of electricity, the high prices were opportunities to shift usage to other times and realize bill credits.

The data used in the analysis are for customers in one of the CSW operating company's RTP and Load Reduction programs. During the summer months (June through September), customers enrolled in the RTP program were given each day a set of 24-hourly prices at which they could purchase electricity the next day.

To enroll, each customer was assigned a customer base load (CBL). The CBL represented the customers' expected electricity usage under the tariff rates that otherwise would apply to that customer. The RTP rate was constructed so that if consumption was exactly equal to the CBL, then the customer remained revenue-neutral relative to purchasing electricity under the equivalent tariff purchase price. If consumption rose above the CBL, the total bill increased. If instead consumption fell below the CBL, and the total bill was decreased.

The RTP-LR program differed from the RTP program in three important ways. First, customers were paid an option payment up-front. Second, on a specified number of occasions, the utility could substitute a $0.38/kWh outage cost component for the amount included in the corresponding day-ahead price quote, thereby raising the price these customers faced relative to that faced by RTP participants. Third, RTP-LR participant's CBL was reduced by the amount of load subscribed for 'curtailment' under RTP-LR, the amount to which the RTP-LR credit was applied. Consequently, during such periods RTP-LR customers not only faced higher settlement prices for usage above their CBL than their RTP counterparts, these prices applied to load below their normal CBL, which provided an added in incentive to reduce load. This option was exercised when real-time prices were expected to be high, which amounted to only a few hours a year. This is evident in the hourly average RTP and RTP-LR prices illustrated in Figures 1 and 2.

There are 54 customers for whom data are available; 43 customers enrolled in the RTP program, and 11 customers enrolled in RTP-LR. Price and usage data is available for the period 1998-2000. Of the 54 customers, 39 were enrolled in one of the programs for all three years, and six participated for only one year, including four new enrollees in 2000. One customer was in the program only in 1998, and another was in a program only in 1999. Nine customers were in a program for only two of the three years, five of them for the 1998-99 period and four of them for the 1999-2001 period. As a result, stacking the participation data results in different levels of participation in each year of the study.

Participant data include the hourly customer base load (CBL), actual hourly kWh, and hourly RTP prices, and a characterization of it business activity. A heat index was constructed to account for load increases associated with hot temperatures.

Although there was no customer control group against which to measure load response to price differences, the availability of a customer's base load (CBL) served a similar purpose. Comparing changes in electricity usage relative to the CBL in high and low priced periods, allows us to isolate load shifting from load growth overall. If these customers can shift load, one would expect to see proportionately more load growth when prices are low. This issue is discussed in greater detail below.

The Electricity Demand Model

The electricity demand model adopted for this study is conceptually similar to the consumer demand model discussed by Braithwait (2000) to explain response to a time-of-use pilot. In that paper, Braithwait (as in an early paper by Caves, *et al*., 1984) derives the electricity demand function from the maximization of a three-level indirect utility function, which is assumed separable in electricity consumption.[4] At the first level, weekday electricity usage is assumed to be allocated by the firm between time periods in which electricity prices it faces differ substantially, for example mid-day and all other hours. The second level allocates monthly usage between weekdays and weekends, while

---

[4] For a production function or utility function to be weakly separable in any partition of its arguments, the marginal rate of substitution between any two inputs or goods in a separable subset is independent of all inputs or goods that are not in the subset (Chambers, 1988, pp. 45-46). In other words, any function defined in n variables, $f(x) = F(x_1, \ldots, x_n)$, that is separable in a partition x^1 through x^m, where x^i is a vector representing a subset of the n variables, can be written as $f(x) = F(f^1(x^1), \ldots, f^n(x^n))$. Each of the sub-functions can be treated as an aggregate input or consumption bundle, essentially a production or utility function in and of itself. Therefore, it is legitimate to think of production or consumption occurring in two steps. Inputs in the sub-function are combined to create the aggregate inputs in the first step. In the second step, these aggregate inputs are used to produce the output via the overall, macro production function. The practical implication is that choice of cost minimizing input levels within any sub-function depends only on prices for those inputs in the sub-function, and other input prices are not required. Consequently, input demands and price response elasticities can be derived from the sub-function alone.

the third determines overall electricity expenditures as a proportion of income, reflecting the relative demand for electricity in relation to all other inputs. Empirically, Braithwait focuses exclusively on the first stage, deriving demand functions using both the constant elasticity of substitution (CES) and Generalized Leontief (GL) forms.[5]

In this application, we adopt a similar strategy, and adjusted for the circumstances. Since the focus is on the use and allocation of electricity inputs by industrial and commercial customers, the theoretical economic model is one that minimizes a three-level electricity cost function for a given level of firm output, which is assumed separable in electricity usage. At the first level, weekday electricity usage is allocated between time periods over which electricity prices differ, the value of electricity to the firm differs, or both. The second level allocates monthly usage between weekdays and weekends, while the third determines overall electricity expenditure as a proportion of total costs, reflecting the relative demand for electricity in relation to all other inputs in the production process.

The focus here is primarily on the first stage in the decision process, which involves the allocation of daily electricity usage between high-price hours (peak) and low-price (off-peak) hours, to quantify price elasticity of electricity substitution. In addition, we estimate separately the extent to which the change in electricity consumption due to differences in prices reflects just substitution behavior, or a combination of substitution and conservation behavior. This latter analysis is made possible by the use

[5]In their analysis of five experimental implementations of residential TOU rates in the United States, Caves and Christensen (1984) estimate a demand model that includes all three stages of electricity demand. If not the only study of its kind, their analysis is one of only a handful of studies that consider more than just the within day energy demand.

of the CBL in the RTP design; it provides use with a measure of typical overall electricity usage to which actual usage under RTP and RTP-LR can be compared.

To estimate the model empirically, we specify a Generalized Leontief (GL) model, which offers a flexible form and places no *a priori* restrictions on the nature of the price response (Boisvert, 1982; Chambers, 1988, Diewert, 1974, Tishler and Lipovetsky, 1997). As a result, price response can vary between pairs of inputs as well as with input levels and input prices.[6] This specification is likely to perform better than the more structured self-dual models such as the constant elasticity of substitution (CES) form.[7] To apply the GL model function, it is convenient to assume that the firm is facing a separable indirect GL cost function and specify only the electricity component, thereby eliminating the need to provide prices for other firm inputs.

The Indirect Generalized Leontief Cost Function

To begin, we specify firm-level production function that is separable in electricity inputs as follows:

(1) $Q = F(x_1, x_2, \ldots, x_n, q(k_1, \ldots, k_n))$,

where Q is output, the x_i are inputs other than electricity, and $k_i, \ldots, k_n$ are amounts of electricity used during periods i through n, respectively. Because production is assumed

[6] The term "flexible functional form" was originally defined by Diewert (1974) and is most often used in the context of indirect cost and profit functions, from which output supply functions are derived. In general, the requirement for a function to be "flexible" precludes the simple imposition of global concavity on the cost function. This means that there could be multiple profit maximizing or cost minimizing optima. This is not a serious problem in our case.

[7] For comparison purposes, we also provide results based on a constant elasticity of substitution (CES) production function in Appendix A. The CES form is particularly convenient because the CES production function and its dual indirect cost or profit functions are self-dual. That is, the indirect profit or cost functions have the same algebraic form. Thus, in this case one arrives at exactly the same expression for estimating usage response to changes in price from either directly manipulating the first-order conditions for cost minimization, or as Braithwait does, by deriving the electricity demand functions from the indirect or dual specification. In the appendix, we focus on the direct cost function formulation because it allows on to recover the distribution parameter from the electricity aggregate so that one could actually calculate the electricity aggregate for use in further analysis.

to be separable in electricity inputs, we can specify the function F as above, where the electricity inputs can be combined according to an aggregator function q. This amounts to specifying a sub-function within F. Any combination of inputs $k_1,\ldots, k_n$ that yields the same value for the energy aggregate, q, is equally productive in producing the firm's output, Q. It is the nature of this sub-function that determines the substitutability of electricity among different periods of the day, and the focus of our empirical investigation.

By appealing to duality theory (Shepard, 1970), we can specify an indirect cost function associated with both the production function Q and the sub-function q above. Because of the assumption that the production function is separable in electricity inputs, we are only concerned with the indirect cost function associated with the electricity aggregate's sub-function. From that sub-function, we can derive expressions for the elasticity of substitution among electricity use during different times of the day.

If we assume that the underlying aggregator function for q is linear homogenous in the electricity inputs (k_i) and that the indirect cost function C is a flexible Generalized Leontief (GL) function, then we have for *n* daily time periods (for $i, j = 1,\ldots,n$):[8]

(2) $C = q \{\sum_i \sum_j d_{ij} (p_i\, p_j)^{1/2}\}$;

This function is linear homogenous in all prices, which is a requirement for a well behaved indirect cost function. That is, if all prices are changed in the same proportion, then C changes in the same proportion as well. We also require that $d_{ij} = d_{ji}$, for symmetry.

[8] Diewert (1974) shows that if the generalized Leontief function (or any cost function) can be decomposed is this form, then the underlying aggregator function for q reflects a constant returns to scale technology.

Shepherd (1970) demonstrates that the optimal factor demands (in this case electricity demands, k_i ($i = 1, \ldots, n$)) can be determined by differentiating (2) with respect to each price as follows:

(3) $\partial C/\partial p_i = k_i = q\,[\,\sum_j d_{ij}\,(p_j / p_i)^{1/2}\,]$.

One purpose of flexible cost functions is to facilitate the estimation of the Allen (1938) partial elasticities of input substitution, which, for a cost function (2), are equal to:[9]

(4) $\sigma_{ij} = C\,C_{ij} / [C_i\,C_j]$,

where the subscripts on σ_{ij} refer to the first and second order partial derivatives of C with respect to inputs i and j. Evaluating equation (4) for the GL cost function given in equation (3), we have:

(5) $\sigma_{ij} = ½\,[C\,d_{ij}\,(p_i\,p_j)^{-1/2}\,] / [q\,a_i\,a_j\,]$,

for all i and j, but $i \neq j$ and $a_i = k_i / q$. In contrast eponymous to the Constant Elasticity of Substitution model, the elasticity of substitution for the GL model varies from observation to observation.[10] In this case, the Allen partial elasticity of substitution varies

[9] As discussed originally by Allen (1938, pp. 508-09), the partial elasticity measures the degree to with the demand for factor j changes as the price of factor i changes. For $\sigma_{ij} > 0$, if the price of factor i increases, then the use of factor j increases, replacing in part factor i in production, and the two factors are said to be *competitive*. If, on the other hand, $\sigma_{ij} < 0$, the two factors are *complements*, and as the price of one of them rises, the demand for both falls. Competitiveness between factors is, on the whole more general than complementary. One factor input cannot be complementary with all others. In the two input case, the direct elasticity of substitution (which measures the percentage change in factor intensities as the inverse price ratio changes by one percent) is equal to the Allen partial elasticity of substitution.

[10] It is this particular feature of the GL model that facilitates our testing of the extent to which price response, defined in this way, increases as the ratio of peak to off-peak prices increase. While this is an attractive feature of this flexible for of the indirect cost function, both Diewert (1974) and Berndt (1991) both emphasize the fact that we are guaranteed that this form is a well-behaved indirect cost function only if all $d_{ij} > 0$. It is possible to generalize this function to allow for some negative d_{ij}, but in this case, the resulting cost function need not be non-negative for all factor prices. Thus, since it is not possible to restrict these parameter estimates to be non-negative, it is necessary to ensure that at each data point the estimated cost function is monotonically increasing and strictly quasi-concave in input prices. To do this, we must verify that the fitted values for all the input-output equations are positive and that the nxn matrix of the σ_{ij} substitution elasticities is negative semi-definite at each observation (Berndt, 1991, p. 465). Furthermore, Diewert (1974, pp. 503-505) shows how to calculate the set of factor prices where these

with the nominal value of the input price ratios, the energy aggregate, and the cost minimizing levels of electricity use. Further, for the Allen own partial elasticities of substitution, we have (for all i):

(6) $\sigma_{ii} = -½ [C \sum_{j\neq i} d_{ij} (p_j^{-1/2} p_i^{-3/2})] / [q\, a_i^2]$.

The Estimation Problem

To calculate the partial elasticities of substitution, we need estimates of the parameters for the cost function given by equation (2). Normally, to estimate the parameters of this cost function, one needs only to assume that an additive error structure is associated with the input demand equations (3), and then estimate them as a system of equations imposing across-equation restrictions to insure symmetry of the parameters. This is accomplished most conveniently by dividing each of equation by q (Brendt, 1991). That is, one can simply estimate for all i:

(7) $a_i = k_i / q = [\sum_j d_{ij} (p_j / p_i)^{½}]$.

When j = i, we have $(p_j / p_i) = 1$, and d_{ii} is a constant in the equation for a_i. In this formulation, we can implicitly restrict the coefficients to be symmetric by always writing the subscripts in the same order.

Unfortunately, because q in our case is the electricity aggregate, it cannot be observed directly, and it is impossible to employ this strategy. However, using full information maximum likelihood (FIML) methods within PROC MODEL in SAS, one can estimate the parameters from equations defined in the ratios of the a_i. That is, we can estimate (for all $i \neq m$):

(8) $k_i / k_m = [\sum_j d_{ij} (p_j / p_i)^{½}] / [\sum_j d_{mj} (p_j / p_m)^{½}]$.

---

conditions hold, and goes on to show that beyond these particular price ranges, the price of one factor has

Within the SAS PROC MODEL we can also impose the symmetry restrictions on d_{ij}. As stated above, since it is not possible to restrict these parameters to be non-negative, it is necessary to verify that the fitted values for all the input-output equations are positive and that the nxn matrix of the σ_{ij} substitution elasticities is negative semi-definite at each observation (Berndt, 1991, p. 465).

In this form, the equations are non-linear in the parameters. Taking the logarithm of both sides of the equation provides a linear model that is more convenient for estimation purposes, as follows:

(8') $\ln [k_i / k_m] = \ln \{[\Sigma_j d_{ij} (p_j / p_i)^{½}] / [\Sigma_j d_{mj} (p_j / p_m)^{½}]\}$.

This strategy will not get rid of the non-linearities, but it will convert each equation into the differences between two logarithms within which there are coefficients imbedded. Whether SAS deals with that kind of non-linearity better than these quotients is an empirical question.

Solving for the Elasticities of Substitution

To evaluate the elasticities of substitution at every data point, one first needs estimated (or predicted) values of a_i, and C/q, the cost per unit of the electricity aggregate. We can predict the a_i's directly by substituting the estimated parameters of (8), denoted by d_{ij}*, into equation (7). For convenience label these $(a_i)_{fit}$. Following Berndt (1991, p. 493), one can obtain predicted values for the unit costs (C/q) of the energy aggregate (q) in the following way:

(9) $(C/q)_{fit} = \Sigma_i\ P_i\ (a_i)_{fit}$.

---

risen sufficiently high for that input to become non-essential in production.

These predicted values for each data point are then substituted into equations (5) and (6) to obtain estimates of the partial elasticities of substitution.[11]

Defining the Electricity Commodity

In the literature, it is generally agreed that the conceptual model above is an appropriate representation of a firm's factor demand system. Because of the continuous nature of electric service and usage, however, an appropriate empirical specification depends ultimately on how one defines the electricity commodities. Because price regimes vary among electricity markets, and firms adopt corresponding usage strategies, the definition of what constitutes electricity commodities is generally treated as an empirical question, determined primarily by the prices faced by customers, the circumstances determining how customers use and value electricity, and the availability of data.

For example, to guarantee variability in the price data needed for the econometric estimation, studies of price response to time-of-use (TOU) rates typically utilize pooled data for customers participating in different TOU rates, or data is pooled across several treatments, where prices or the definition of the peak period varies by the rate and/or the experimental design (e.g., Caves *et al*., 1984; Patrick, 1990; Braithwait, 2000). [12] To

[11] In preliminary analysis, we also experimented with another flexible form, the translog (TL) model (Boisvert, 1982 and Chambers, 1988). The empirical results using this model were not encouraging, most likely for the same reasons that Caves and Christensen (1980 a, b) argue that the TL model does not perform well when substitution elasticities are likely to be small. Another reason for this poor performance is due to the fact that the TL model relies on cost shares in the estimating equations. The performance can be problematic if there are small shares or large relative differences among shares.

[12] Because of the nature of the residential TOU rate, Braithwait (2000) was able to examine price responsiveness of customers in two different ways. The first was to estimate substitution elasticities between peak and off-peak periods. Here, he was assuming that for any given day, there were only two electricity commodities. His second approach was to estimate substitution elasticities among three separate electricity commodities—usage during peak, shoulder and off-peak periods. As one might expect, the substitution elasticities between peak and shoulder periods and shoulder and off-peak were lower than for peak to off-peak periods. These peak to off-peak elasticities were also slightly higher than the overall substitution elasticities when the shoulder period was considered as being part of the off-peak period.

establish the definition of distinct electricity commodities, Caves *et al*. (1987) identified six separate commodities for customers facing a six-hour peak-pricing period of 9 A.M – 12 noon and 1 P.M. – 4 P.M. These peak hours are then further divided into two separate commodities—one two-hour commodity (11 A.M – 12 Noon and 1 P.M – 2 P.M) and one four-hour commodity (9 A.M. – 11 A.M. and 2 P.M. – 4 P.M). Other hours in the day are aggregated into four separate commodities; all are priced the same. The authors asserted that this sub-aggregation of the peak is needed to characterize needle peaking, a load shape characterize by a very distinct peak of very short duration. Extending this commodity definition to RTP-type programs that rely on hourly prices would lead to 24 separate electricity commodities, one for each hour of the day (e.g., Herriges *et al*., 1993 and Schwarz *et al*., 2002).

This 24-hour specification would perhaps be warranted if industrial and commercial customers were able to adjust usage to changing hourly prices on a contemporaneous and ongoing basis. However, there is compelling evidence that firms implicitly characterize the day as being comprised of a peak and off-peak period (Neenan Associates *et al*., 2002a and Goldman, *et al*., 2004). While the exact specification of what comprises the peak hours is firm specific, common business practices, driven in large part by traditional rate structures, support utilizing a single specification to capture most of the variation in usage.[13]

Therefore, we analyze the price response behavior for CSW's customers on RTP rates by estimating the model for several alternative specifications of the peak period that

[13] For example, if customers have operated under a rate that imposes demand charges and peak energy charges during weekdays between noon and 8:00 p.m. for several years, we would expect that all would take this price regime as given in making both short term decisions about when to use electricity and long-term equipment purchase and plant process organization decisions..

differ in length. In so doing, we gain insights into which hourly aggregate firms view as distinct commodities, and select a specification that best represents the data. This strategy has the added advantage of determining the extent to which the periods in which customers are the most price responsive correspond to those in which load reduction is of most value to the system.

The Peak and Off-Peak GL Model

The GL model for two periods (peak and off-peak) is given by the following:

(10) $C = q\{d_{pp} p_p^{1/2} p_0^{1/2} + d_{p0} p_p^{1/2} p_0^{1/2} + d_{0p} p_0^{1/2} p_p^{1/2} + d_{00} p_0^{1/2} p_0^{1/2}\}$

The resulting two equations for average electricity use are given by:

(11) $a_p = k_p / q = d_{pp} + d_{p0} (p_0 / p_p)^{½}$

(12) $a_0 = k_0 / q = d_{00} + d_{p0} (p_p / p_0)^{½}$.

The ratio of these two equations to be estimated in SAS PROC MODEL is:

(13) $[a_p / a_0] = [k_p / k_0] = [d_{pp} + d_{p0} (p_0 / p_p)^{½}] / [d_{00} + d_{p0} (p_p / p_0)^{½}]$,

and for estimation, we have the following logarithmic specification:

(13a) $\ln [a_p / a_0] = \ln [k_p / k_0] = \ln\{[d_{pp} + d_{p0} (p_0 / p_p)^{½}] / [d_{00} + d_{p0} (p_p / p_0)^{½}]\}$.

Finally, we can substitute the estimated parameters denoted d_{ij}*, into equation (7) to calculate $(a_i)_{fit}$ at each data point. In turn, these are substituted into equation (9) to obtain estimates of $(C/q)_{fit}$. Finally, these estimated expressions are substituted into equations (5) and (6) to obtain for each data point:

(14) $\sigma_{p0} = ½ [(C/q)_{fit} d_{p0}^{*} (p_p p_0)^{-1/2}] / [(a_p)_{fit} (a_0)_{fit}]$,

(15) $\sigma_{pp} = -½ [(C/q)_{fit} d_{p0}^{*} (p_0^{1/2} p_p^{-3/2})] / [(a_p)_{fit}^{2}]$, and

(16) $\sigma_{00} = -½ [(C/q)_{fit} d_{0p}^{*} (p_p^{1/2} p_0^{-3/2})] / [(a_0)_{fit}^{2}]$.

In this two-input case, the cross Allen partial elasticity of substitution (14) is equivalent to the direct elasticity of substitution which measures the proportional change in the ratio of peak to off-peak electricity use due to a one percent change in the inverse price ratio (Ferguson, 1969). [14] For this production function to be well behaved, Ferguson (1969) shows that $0 < \sigma < \infty$.[15] The higher σ is, the more responsive energy use is to changing in relative prices between peak and off-peak periods. For example, if $\sigma < 1$, then as the price ratio changes by one percent, the ratio of peak to off-peak energy use changes by less that one percent. For $\sigma > 1$, the ratio on energy use changes by more than one percent as the inverse price ratio changes by one percent.

The Empirical Specification of the GL Model

To estimate the GL model, we must define exactly how the variables used in the empirical regression analysis are calculated from the data. From equation (13), we need to have the ratio of peak to off-peak electricity use. Since RTP rates are designed to be revenue neutral relative to the flat rate background tariff if a firm's energy use replicates the CBL, it is reasonable to expect that changes in electricity use should be measured relative to the CBL.[16] For this reason, we define for each weekday, t, and firm or group of firms, m:

k_{ptm} = peak actual kWh / peak CBL;

k_{0tm} = off-peak actual kWh / off-peak CBL;

[14] This relationship shows that σ is the proportional change in the use of electricity in the peak period relative to the off-peak period (holding output, in this case the electricity aggregate, constant), as the inverse price ratio increases by one percent Fergurson (1969, pp. 103-04).
[15] Since it is not possible to restrict these parameters to be non-negative, it is necessary to verify that the fitted values for all the input-output equations are positive and that the nxn matrix of the σ_{ij} substitution elasticities is negative semi-definite at each observation (Berndt, 1991, p. 465).
[16] Because the CBL data come from electricity use in the year prior to a firm's enrollment in the program, we assume that for any month and day of the week, the CBL is the average for those days in the respective months for which we have CBL data.

p_{ptm} = average hourly peak price / kWh; and

p_{0tm} = average hourly off-peak price/kWh.

Other variables included in the model specification to improve the explanatory power of the model. Binary (0,1) 'dummy' variables are included to account for differences by firm in peak relative to off-peak energy use because of alternative types of production processes, differences in production shifts, differences in business hours, etc.[17] These variables are defined for the m firms (m = 1,…,M) as:

D_m = 1 if the observation is for firm m, and = 0 otherwise.

We also capture the fact that RTP LR customers may respond to prices differently from other firms, since they are subject to a different price regime, by defining a dummy variable as follows:

D_{LR} = 1 if the observation is for a RTP-LR customer, and = 0 otherwise.[18]

Finally, the effect of daily weather on the ratio of peak to off-peak electricity use is captured through a weather index specified as:

w_{tm} = weather index for day t derived from the weather station nearest to firm m.[19]

Given this set of variables, the equation to be estimated is:

$$\ln [\, k_{ptm} / k_{otm} \,] = \textstyle\sum_m (F_m)\, D_m + (w)\, W_{tm} + \{ \ln [\, (L_R)\, D_{LR} + (D_{pp}) + (D_{po})\, [\, p_{0tm} / p_{ptm} \,]^{1/2} \} - \{ \ln [\, (L_0)\, D_{LR} + (D_{oo}) + (D_{op})\, [\, p_{ptm} / p_{0tm} \,]^{1/2} \}, \tag{17}$$

[17] If we had other firm characteristics, the model could be designed to account for differences in these factors directly. Unfortunately, these additional data were not available for this study.

[18] Firms that have faced fixed tariffs for electricity for many years must learn how to respond to price differences between peak and off-peak periods. The efficiency should be higher the longer a firm has faced price variation. In early experimentation with the data, we tried to capture this effect by defining a variable: T_{tm} = the number of months that firm m has been the RTP program on day t. Unfortunately, this variable was unable to capture this effect, and it was dropped from the analysis at an early stage.

where the terms in parentheses are coefficients to be estimated. In the estimation, we impose $L_R=L_0$, and $D_{po} = D_{op}$ to ensure the required symmetry among estimated coefficients. We also require that $D_{oo} + D_{pp} + D_{op} + D_{po} =1$, which normalizes the coefficients to reflect a unit isoquant for the energy aggregate.[20]

Electricity Conservation vs. Shifting

In addition to measuring the elasticity of substitution between peak and off-peak electricity consumption, the availability of data on the CBL by customer by hour enables one to study the extent to which the rate structure promotes energy conservation as opposed to the substitution of usage between peak and off-peak periods. This can be particularly important in accounting for customer actions involving foregoing discretionary usages, such as turning off AC units and reducing lighting and other plug loads, that may not result in a reduction of firm output, and that generally do not require recovery usage at a later time.[21]

This behavior can be analyzed by estimating a model similar to that proposed by Patrick (1990). He distinguished between pure load shifting from responses that involve foregoing consumption, which he referred to as conservation, by estimating the following regression equation:

(18) $\%\Delta Q_T = a + \sum_m (F_m) D_m + \beta_q \%\Delta q_p + u,$

where $\%\Delta Q_T =$ % change in daily usage relative to the daily CBL, $\%\Delta q_p =$ % change in daily peak period usage relative to the CBL during the peak; and u is an error term. We

---

[19] The weather index is based on heating and cooling degree days constructed from mean daily temperature and dew point values for five weather stations located in the utility's service territory. The construction of the index is in Appendix B.

[20] Although the parameter estimates changed if this adding up condition were set to a number different from 1, the estimates of the elasticities of substitution were invariant to the specification.

[21] A more detailed graphic discussion of this situation is found in Appendix C.

estimate this model separately for the group of LR customers and for the group of non-LR customers. For each group, the appropriate dummy variables are included to control for firm-level differences.

To interpret the coefficients of this model, it is important to remember that as the price of on-peak electricity rises (*ceteris paribus*), electricity becomes a more expensive input for customers, and there is an inducement for customers to forego usage based on its cost. At the same time, customers also have an incentive to shift some load from peak to off-peak periods while maintaining firm output.

The parameter, β_q, associated with the variable $\%\Delta q_p$ in equation (9) can be interpreted in a way that isolates these two effects. β_q is the proportion of the reduction in peak demand that is due to overall daily energy conservation. Consequently, only that proportion of peak load reduction equal in percentage terms to the percentage downward shift in total daily load due to the higher cost of electricity is counted as conservation. In doing so, foregone energy consumption is defined as daily conservation only if it disappears from both peak and off-peak periods. This is as it should be because electricity conserved on a particular day involves foregoing consumption proportionally in both the peak and the off-peak period.[22] While the coefficient β_q accounts for the proportion of load reduction on peak that is equal to the overall downward shift in daily load, $(1 - \beta_q)$ is the proportion of peak load shifted to off-peak periods. It captures the non-parallel change in the peak to off-peak load shape that is due to the fact that peak price is higher relative to off-peak price, which leads customers to substitute on-peak

[22] We still do not know from this analysis whether the electricity conserved on the day is never consumed or is consumed on another day. Such an analysis is beyond the scope of this research.

electricity for off-peak electricity. These measures are exact, provided there is no output effect.

Given this interpretation, one would expect that $0 < \beta_q \leq 1$. If β_q were to take on an extreme value of zero, then as peak demand changes relative to the CBL in response to peak to off-peak price differentials, the entire change may be due to shifting usage from peak to off-peak periods. Conversely, if $\beta_q = 1$, then the entire change may be due to conservation. For this latter case to obtain, there would have to also be a corresponding reduction in off-peak demand.

The Empirical Results

To use these data from CSW's RTP customers to examine the load response of customers to price differentials between high-and low-price periods of the day, we defined alternative "peak" and "off-peak" periods that included a different number or sequence of hours of the day. In doing so, we sought to ascertain a balance between a customer's need to treat electricity during blocks of hours as perfect complements and to define electricity commodities where hourly prices within the blocks are similar.

The estimated coefficients for models representing eight peak period definitions are given in Table 1. These peak periods are defined primarily to examine the effects of the length of the peak on the elasticities of substitution, and to a lesser extent, the effect of the time of day during which the peak period is defined.[23] The alternative peak periods

[23] Although the actual elasticity estimates from the models for other peak period definitions do differ, they are similar qualitatively, and the implications for differences by length of peak, relative price differences, etc. are very similar. Thus, these results add little to our understanding and are not reported here. The only result worth comment is that as the end of the peak period was assumed to occur early than 6:00pm, the explanatory power of the models falls significantly.

ranged from three to six hours in length. Four of the peak periods ended at 6:00 p.m., while the other four ended at 7:00 p.m. [24]

Although one might argue that the statistical performance of all these models is reasonable, some definite patterns are evident. First, as the length of the peak period is extended, the explanatory power of the models, as measured by the R^2, falls.[25] In contrast, the considerable precision with which the coefficients are measured, as evidenced by the consistently high t-ratios, *does not* deteriorate in a similar fashion. This is encouraging because it is these coefficients that form the basis for the estimates of the elasticities of substitution.

Although the coefficient on the weather variable in each model is relatively high compared with the standard errors, the effect of the weather index on peak to off-peak electricity usage on the substitution elasticity is modest. In contrast, the dummy variable included in the logarithmic portions of the estimating equation (17) to test for differences in the elasticities of substitution across load reduction and non-load reduction customers was significant; in all models, the t-ratios on the coefficients associated with these variables are large. As is seen below, the inclusion of this dummy variable has a pronounced effect on the estimates of the elasticities of substitution.

As is clear from the algebra above, the coefficients of the estimated models reported in Table 1 are difficult to interpret individually. However, through equations (14), (15), and (16), they can be used to calculate the elasticities of substitution between

[24] The summer system peak for this utility is from 5:00 to 7:00 p.m.
[25] Each of these models contained dummy variables for each firm to control for individual firm differences, but the coefficients are of little use in our analysis, and thus, they are not reported.

peak and off-peak electricity usage, where the elasticities can be different for each price point.[26]

Estimates of the Elasticities of Substitution

From the data in Table 2 and Figure 3, it is evident that RTP LR customers are more responsive to peak-off-peak price differentials than are the other customers. For both groups of customers, however, the elasticities of substitution are generally highest for the three-hour peaks. They range from just under 0.20 to almost 0.27 on average for RTP LR customers, whereas for RTP customers, the averages range between 0.10 to just over 0.18.

For both groups, the elasticities of substitution eventually fall with the length of the peak periods. This is consistent with the findings in the analyses of most TOU rates for residential customers. For our analysis, the substitution elasticities fall throughout the length of peak for the LR customers, while for the other group, the elasticity of substitution actually rises slightly when the peak period is lengthened from three to four hours (if the peak ends at 6:00 p.m.), and from three to five hours (if the peak ends at 7:00 p.m.).

For the most part, these results are as expected. Shifting load to avoid high prices requires a greater effort as the peak period becomes longer, particularly when the peak includes the entire afternoon and shifting load to off-peak hours in effect requires rearranging the entire day's activities. In contrast, shorter peaks leave more room to maneuver, since the early or late afternoon hours are still available (off-peak) to make up for the peak load reduction. Since RTP LR customers are substantially more responsive

[26] As with all flexible forms, there is no guarantee that the GL cost model is well behaved globally. Since all the coefficients are positive, this model is globally well behaved (Chambers, 1988). However, we did

throughout, it is understandable that as the peak period is extended, this "fatigue" factor sets in sooner for them.

The marked differences in price responsiveness across customer groups and peak periods have important implications for the design of programs. Since these programs must be able to balance system supply considerations with customers' curtailment capabilities, it appears that a portfolio of pricing products with different lengths of exposure to high prices might be the most cost effective approach.

Understanding the Variation in the Elasticities of Substitution

The advantages of a portfolio of pricing programs are also underscored by the fact that the elasticities of substitution vary across days.[27] While the ranges in the elasticities of substitution for each customer group and each peak period appear relatively small, in some cases there is more than a 40% difference between the high and the low values (Table 2). One possible hypothesis to explain this variability is that customers are more price responsive as electricity prices rise (Schwarz *et al*., 2002).

To shed some light on this issue, and test this hypothesis statistically, the elasticities of substitution were regressed on the peak to off-peak price ratio (Table 3). In all peak periods, the elasticities of substitution rise with the peak to off-peak price ratio. As the price ratio rises by one unit, the elasticities of substitution increase by as much as 0.013, with most increases ranging from 0.007 to 0.009, providing strong support for this hypothesis; they represent significant changes in the rate of price response relative to the corresponding minimum values of the elasticities of substitution reported in Table 2. For

---

evaluate the relevant Hessian matrix at each data point to verify this fact (Berndt, 1991).

[27] For purposes of comparison, we also estimated a CES demand model. The average elasticities reported in Figure 3 and Table 2 are quite similar to those produced by this CES model reported in the appendix (Table A1 and Figure 1A).

example, the elasticity of substitution for load relief customers during the 4:00pm through 6:00pm peak increases by 0.013 for every unit increase in the peak to off-peak price ratio. From the regression results, one would estimate that it if the peak price were seven times that of the off-peak price, the elasticity of substitution for load relief customers would be nearly identical to the maximum of the range (0.342) reported in Table 2.[28]

In contrast, for small differences in peak to off-peak prices, one would expect the price response for these customers to be closer to the lower end of the range (0.253). This evidence that load reduction on peak will increase proportionally more when price differences are high clearly has important implications for policy and rate designs, particularly for those based on critical peak pricing, because load reduction is also of most value to the overall bulk power system and electricity markets when peak to off-peak price differentials are high.

Conservation vs. Load Shifting

The final issue addressed in this empirical analysis is the extent to which the differentials in peak to off-peak prices foster energy conservation or the shifting of load from peak to off-peak periods. We investigated this issue by regressing the percentage change in daily usage relative to the daily CBL on the percentage change in daily peak period usage relative to the CBL during the peak (equation 18). Although the results of

[28] This estimated increase in the elasticity of substitution is calculated by adding seven times the corresponding slope coefficient from Table 3 to the minimum elasticity of substitution for the load relief customers during the 4:00pm through 6:00pm peak period (e.g. 0.253 + 7*0.013 = 0.344).

these regressions are sensible, the statistical performance of the models is less than one would have hoped (Table 4).[29]

With the exception of the two outliers mentioned in the footnote, the β_q coefficients are all less than 0.25, and therefore, (1 - β_q) is above 0.75. The interpretation of this result is clear. On balance, over 75% of the change in electricity usage relative to the CBL is due to shifting load from peak to off-peak periods. Less than a quarter of the change is due to energy conservation. Further, the share of the change due to conservation is somewhat higher for the non-LR firms. There are a couple of possible explanations for these results. During the study period, the economy of the Southwest was strong, and many firms were looking to expand output. To facilitate such expansion, some of these firms signed up for the RTP rates to avoid demand charges as their load expanded. Thus, one would expect to see less conservation when these firms are faced with these high prices. What is encouraging is that even for firms with overall load growth, it appears the growth was proportionately greater in off-peak periods.

Concluding Remarks and Policy Implications

In this paper, we have used hourly load and price data for over 50 industrial and commercial customers to estimate their electricity demand response to price differences between high and low priced hours. In contrast to a handful of other studies of RTP customer demand response, we define peak and off-peak periods rather than treat electricity in every hour (or in groups of hours with the same price) as a distinct commodity. This modeling simplification assumes that firms cannot load-follow on an

[29] This is true, except for the one coefficient that is greater than unity for non-load reduction customers in the 2:00 p.m. through 6:00 p.m. peak period. According to the model in equation (18) no coefficient should be greater unity or less than zero. The coefficient of 0.952 on the % change in peak load variable for non-LR customers in the 2 through 5 peak is also an outlier that is difficult to explain.

hourly basis, and is consistent with some empirical evidence that firms view electricity consumption in certain blocks of hours as distinct commodities. The "demand inducing" price is the average price in the block of hours. By interpreting the data in this way, we are able to study in a consistent fashion the demand response behavior of firms between peak and off-peak times where the peak periods differ in length and time of day. This type of analysis is essential if markets or individual utilities are to offer a portfolio of price-responsive load products that offer benefits both to customers and to the market.

On balance, we find that customers differ substantially in their demand response to price, but under certain conditions, the response can be substantial. Most of the response was due to load shifting from peak to off-peak periods, rather than due to overall energy conservation. This is not unexpected because the firms were only exposed to high prices for about 10% of the year. During the rest of the year, prices were well below previous tariff levels. Moreover, even for those firms with substantial load growth during the economic expansion of the late 1990s, there was also substantial evidence the price differentials led to proportionately higher load growth during off-peak hours.

Further, the response rates appear to be highest for short peak periods, and fall rather dramatically as the length of the peak increases. One explanation for this reduction in response is that a "fatigue" factor sets in and at some point in time the outage costs become so large that firms must restore electricity usage to more normal levels. However, there is also evidence that the elasticities of substitution vary directly with the differential between peak and off-peak prices. Therefore, since the longer peak periods examined have somewhat lower average price differentials, some of this apparent "fatigue" factor may be a normal reaction to smaller price differences.

The policy implications of this analysis seem clear. Since there are marked differences in price responsiveness across customer groups and peak periods, it would appear that a portfolio of pricing products with different lengths of exposure to high prices might be the most cost effective way to balance system supply considerations with customers' capabilities to curtail load. Furthermore, since load reduction on peak increases proportionally more when price differences are high, rate designs, particularly those based on critical peak pricing, can effectively promote load reduction when it is of most value to the overall bulk power system and electricity markets.

References

Allen, R. G. D. *Mathematical Analysis for Economists*. New York, St. Martin's Press, Inc., 1938, reprinted in 1964.

Berndt, Ernst R. *The Practice of Econometrics: Classic and Contemporary*. Reading MA: Addison-Wesley Publishing Company, 1991.

Boisvert, R. N. "The Translog Production Function: Its Properties, Its Several Interpretations and Estimation Problems," Department of Agricultural Economics, A.E. Res. 82-28, Cornell University, September 1982.

Braithwait, S "Residential TOU Price Response in the Presence of Interactive Communication Equipment," in A. Faruqui and K. Eakin. (eds.) *Pricing in Competitive Electricity Markets*. Boston: Kluwer Academic Publishers, 2000.

Caves, D. and L. Christensen. "Global Properties of Flexible Functional Forms", *American Economic Review* 70(1980a): 422-432.

Caves, D. and L. Christensen. "Residential Substitution of Off-Peak for Peak Electricity Usage under Time of Use Prices", *Energy Journal*, 1(1980b):85-142.

Caves, D. and L. Christensen. "Econometric Analysis of Residential Time-of-Use Pricing Experiments", *Journal of Econometrics*, 14(1980c): 287-306.

Caves, D., L. Christensen, and J. Herriges. "Consistency of Residential Response in Time of Use Pricing Experiments", *Journal of Econometrics* 26(1984): 179-203.

Caves, D., L. Christensen, and J. Herriges. "The Neoclassical Model of Consumer Demand with Identically Priced Commodities: An Application to Time-of-Use Electricity Pricing", *The Rand Journal of Economics* 18(1987): 564-580.

Chambers, Robert G. *Applied Production Analysis: A Dual Approach*. Cambridge: Cambridge University Press, 1988.

Christensen, L. and R. Caves. "Global Properties of Flexible Functional Forms", *The American Economic Review*, 70(June 1980): 422-432.

Diewert, W. E. "Applications of Duality Theory," in M. D. Intriligator and D. A. Kendrick (eds.) Frontiers of Quantitative Economics, vol. 2 (Amsterdam: North-Holland, 1974.

Ferguson, C. E. *The Neoclassical Theory of Production and Distribution*. Cambridge: Cambridge University Press, 1969.

Goldman, C., Hopper, N. Sezgen, O., Moezzi, M., Bharvirkar, R., Neenan, B., Boisvert, R., Cappers, P., and Pratt, D. *Customer Response to Day-ahead Wholesale Market Electricity Prices: Case Study of RTP Program Experience in New York*. Report to the California Energy Commission. LBNL-54761. Berkeley, CA.: Lawrence Berkeley National Laboratory, 2004.

Greene, William H. *Econometric Analysis*. New York: MacMillan Publishing Company, 1990.

Ham, J., D. Mountain, and M. Chan. "Time-of-Use prices and electricity demand: allowing for selection bias in experimental data", *RAND Journal of Economics*, 28(1997): s113-s140.

Herriges, J., S. Baladi, D. Caves, and B. Neenan. "The Response of Industrial Customers to Electric Rates Based Upon Dynamic Marginal Costs," *The Review of Economics and Statistics*, (1993): 446-54.

Long, J., B. Scott, and B. Neenan. "Electricity Marketing: Is the Product the Price?", in A. Faruqui and K. Eakin. (eds.) *Pricing in Competitive Electricity Markets*. Boston: Kluwer Academic Publishers, 2000.

Patrick, R. "Rate Structure Effects and Regression Parameter Instability Across Time-of-Use Electricity Pricing Experiments", *Resources and Energy* 12(1990): 180-195.

Miller, M. N. Whittlesey, and T. Barr. "The Constant Elasticity of Substitution Production Function and Its Application in Research," Technical Bulletin Agricultural Research Center, Washington State University, August 1975.

Moroney, J. R. *The Structure of Production in American Manufacturing*. Chapel Hill, NC: The University of North Carolina Press, 1972.

Mountain, D. and C. Hsiao. "Peak and Off-Peak Demand for Electricity: The Hopkinson Rate in Ontario, Canada", *The Energy Journal*, 7(1986): 149-68.

Neenan, B., D. Pratt, P. Capers, R. Boisvert, and K. Deal. NYISO Price-Responsive Load Program Evaluation Final Report, Prepared by Neenan Associates, LLC for New York Independent System Operator, Albany, NY, 2002a.

Neenan, B., R. Boisvert, and P. Cappers. "What Makes a Customer Price Responsive?" *The Electricity Journal* 15(2002b)(3): 52-59.

Neenan, B., D. Pratt, P. Cappers, J. Doane, J. Anderson, R. Boisvert, C. Goldman, O. Sezgen, G. Barbose, R. Bharvirkar, M. Kintner-Meyer, S. Shankle and D. Bates. "How and Why Customers Respond to Electricity Price Variability: A Study of NYISO and NYSERDA 2002 PRL Program Performance" report to the New

York Independent System Operator (NYISO) and New York State Energy Research and Development Agency (NYSERDA), January 2003.

Neenan Associates. Evaluation of NYISO's Demand Response Program Performance (2003). Prepared by Neenan Associates, LLC for New York Independent System Operator, Albany, NY .Available from www.NYISO.com.

Schwarz, P. M., T. N. Taylor, M. Birmingham and S. L. Dardan (2002) "Industrial Response to Electricity Real-Time Prices: Short Run and Long Run" *Economic Inquiry* 40(2002)(4): 597-610.

SAS Institute, Inc. *SAS User's Guide: Statistics, Version 5 Edition*, Gary, NC, 1985.

Tables and Figures

Figure 1. Hourly Electricity Prices for CSW's Non-Load Relief Customers, June-September, 1998-2000

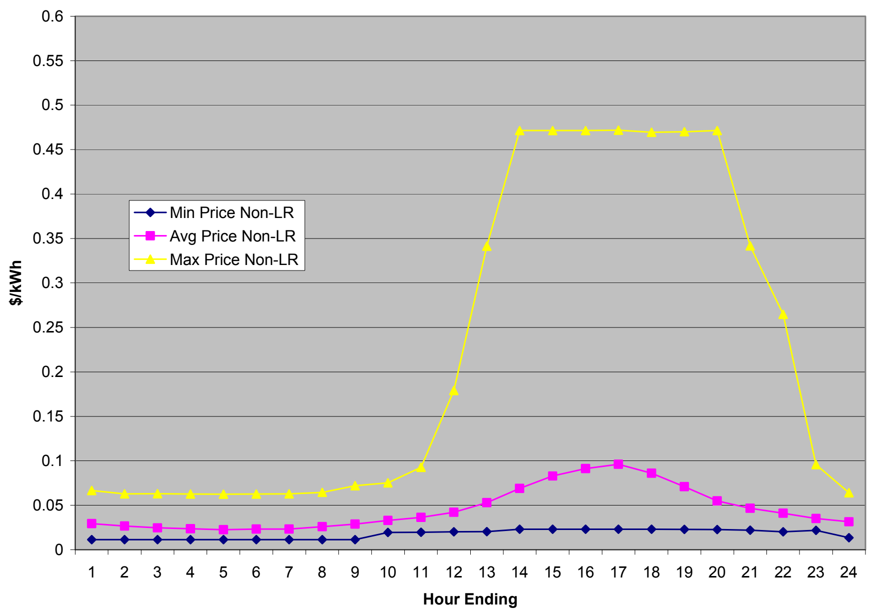


Figure 2. Hourly Electricity Prices for CSW's Load Relief Customers, June-September, 1998-2000

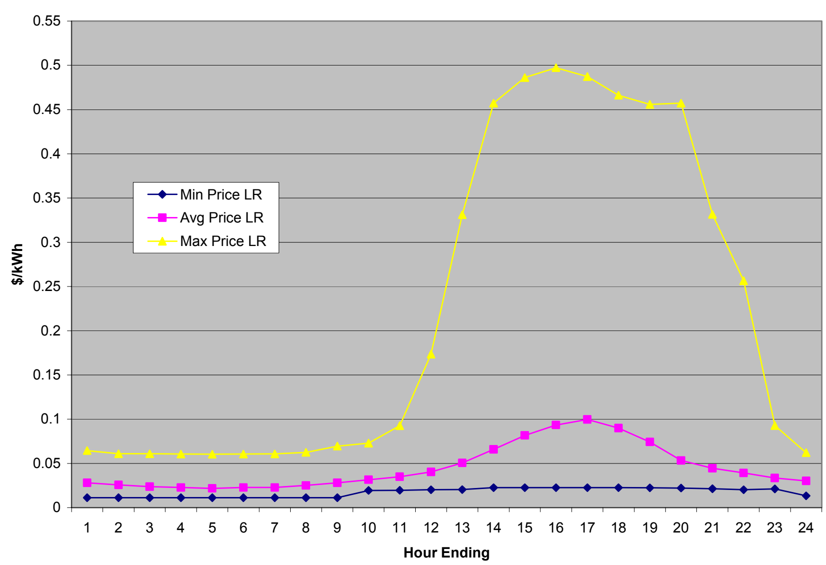


Figure 3. Generalized Leontief Model: Estimated Substitution Elasticities (Average)

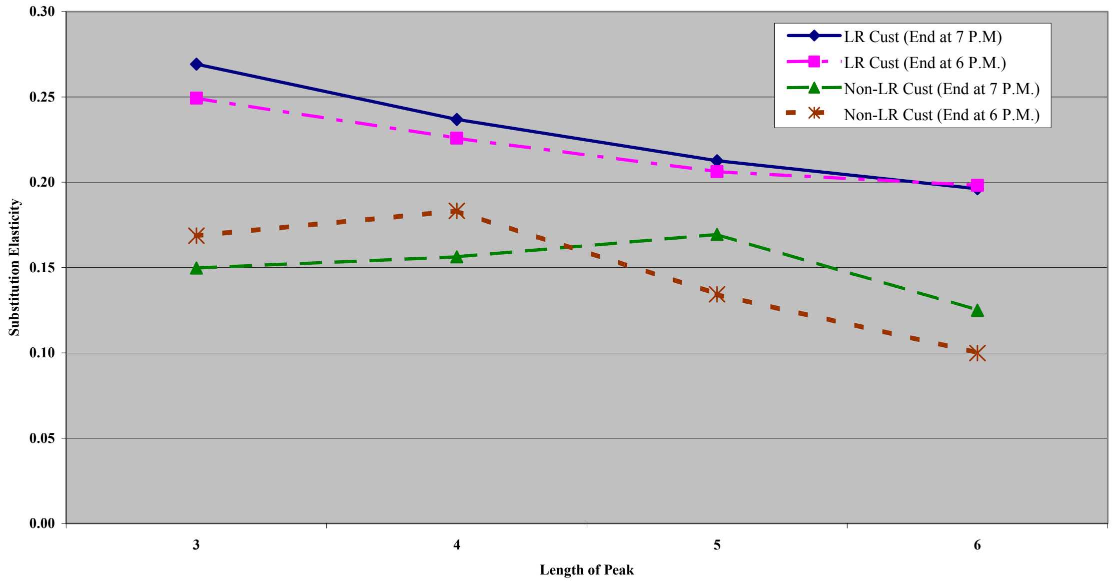


Table 1. PSO Generalized Leontief Model Results

	Parameter Estimates (T-Statistic)*							
Peak Hours (p.m.)#	W	D_{pp}	D_{po}	D_{op}	D_{oo}	L_p	L_o	R^2
4 through 6	0.001 (4.19)	0.547 (40.19)	0.066 (6.21)	0.066 (6.21)	0.321 (22.71)	-0.196 (-4.74)	-0.196 (-4.74)	0.479
3 through 6	0.001 (4.15)	0.481 (42.34)	0.073 (7.72)	0.073 (7.72)	0.374 (29.12)	-0.166 (-3.45)	-0.166 (-3.45)	0.494
2 through 6	0.001 (3.93)	0.433 (45.84)	0.080 (9.85)	0.080 (9.85)	0.406 (32.72)	-0.102 (-1.94)	-0.102 (-1.94)	0.511
1 through 6	0.001 (3.38)	0.461 (66.43)	0.058 (8.89)	0.058 (8.89)	0.423 (40.39)	-0.183 (-4.2)	-0.183 (-4.2)	0.411
3 through 5	0.001 (4.3)	0.464 (36.49)	0.079 (7.99)	0.079 (7.99)	0.377 (26.42)	-0.161 (-3.21)	-0.161 (-3.21)	0.451
2 through 5	0.001 (4)	0.450 (39.22)	0.086 (10.01)	0.086 (10.66)	0.378 (251.03)	-0.094 (-1.76)	-0.094 (-1.76)	0.482
1 through 5	0.001 (2.78)	0.446 (57.41)	0.064 (8.52)	0.064 (8.52)	0.427 (36.14)	-0.177 (-3.72)	-0.177 (-3.72)	0.364
Noon through 5	0.000 (1.64)	0.454 (54.42)	0.047 (6.07)	0.047 (6.07)	0.452 (40.67)	-0.252 (-5.43)	-0.252 (-5.43)	0.245

* The parameters in this model are estimated from equation (17) using PROC MODEL in SAS. The coefficients are defined in equation (17). The W is the coefficient on the weather variable. The L's must be the same for symmetry. It is difficult to give a precise interpretation to these coefficients, but they are used in calculating the elasticities of substitution from equations (14), (15) and (16).

The peak period is for the hour beginning. For example, the peak 4 through 6 begins at 4:00pm goes through the 6:00pm hour and ends at 7:00pm.

Table 2: Range in Substitution Elasticities for Various Peak Periods, Based on the Estimated GL Model

	Load Relief RTP Customers						
	Substitution Elasticities				Peak to Off-Peak Price Ratio		
Peak Hours#	Min.	Avg.	Max.	% Max. of Min.	Min	Avg.	Max
4 through 6	0.253	0.269	0.342	136%	1.04	1.97	10.42
3 through 6	0.224	0.237	0.313	140%	1.04	2.20	14.14
2 through 6	0.203	0.213	0.278	138%	1.04	2.37	13.76
1 through 6	0.185	0.196	0.256	139%	1.04	2.44	11.80
3 through 5	0.238	0.249	0.304	128%	1.04	2.15	9.60
2 through 5	0.215	0.226	0.305	142%	1.04	2.26	16.55
1 through 5	0.197	0.206	0.272	138%	1.04	2.26	13.61
Noon through 5	0.190	0.198	0.255	134%	1.04	2.18	11.56
	Non-Load Relief RTP Customers						
	Substitution Elasticities				Peak to Off-Peak Price Ratio		
Peak Hours#	Min.	Avg.	Max.	% Max. of Min.	Min	Avg.	Max
4 through 6	0.140	0.150	0.180	129%	1.04	1.85	4.69
3 through 6	0.147	0.156	0.186	127%	1.04	2.03	5.44
2 through 6	0.161	0.169	0.198	123%	1.04	2.18	5.88
1 through 6	0.117	0.125	0.154	132%	1.04	2.27	6.56
3 through 5	0.160	0.169	0.195	122%	1.04	2.04	4.92
2 through 5	0.174	0.183	0.211	121%	1.04	2.13	5.33
1 through 5	0.127	0.134	0.161	127%	1.04	2.15	5.99
Noon through 5	0.095	0.100	0.124	131%	1.04	2.10	6.38

The peak period is for the hour beginning. For example, the peak 4 through 6 begins at 4:00pm goes through the 6:00pm hour and ends at 7:00pm.

Appendix A

The CES Specification

The Model

Similar to the analysis in the text, we define a firm's production function that is separable in electricity inputs as:

$$Q = F(x_1, x_2,\ldots, x_n, q(k_p,\ldots, k_o)),$$

where Q is output, x_i are inputs other than electricity and k_p and k_o are electricity used in peak and off-peak periods, respectively. Since electricity is assumed separable from other inputs, and it is of the CES form, we can write the electricity sub-function as:

(1a) $$q = [\, \delta\, (k_p)^{-\rho} + (1-\delta)\, (k_o)^{-\rho}\,]^{-1/\rho}$$

In this function, q is an aggregate electricity input that exhibits constant returns to scale (Moroney, 1972; and Ferguson, 1969). The parameter δ reflects the natural peak kWh intensity of production. The parameter ρ is a transformation of the elasticity of substitution between peak and off-peak electricity use, $\sigma = 1/(1 + \rho)$.[30] This elasticity of substitution is constant regardless of the levels of energy use or levels of output.

To identify the price responsiveness of electricity demand between peak and off-peak periods, it can be shown that the ratio on input use is a function of the inverse of the price ratio for the inputs and the parameters of the δ and σ. This relationship is derived from a model to minimize the electricity cost, as follows:

(2a) $$C = P_p\, K_p + P_0\, K_0,$$

to produce a given level of the electricity aggregate from equation (1a). By manipulating the first-order conditions for this constrained minimization problem, the marginal technical rate of substitution, MTRS, (the ratios of the marginal products of inputs) is set equal to the price ratio. The marginal products for peak and off-peak electricity are (see Miller *et al*., (1975) for the most transparent derivation):

$\partial q / \partial k_p = \delta \, (q / k_p)^{1/\sigma}$ and

$\partial q / \partial k_0 = (1 - \delta) \, (q / k_0)^{1/\sigma}$.

The ratio of these two equations is the marginal technical rate of substitution of k_0 for k_p :

$$\text{MTRS} = [\, \delta /(1 - \delta)\,]\, (k_0 / k_p)^{1/\sigma} .$$

The necessary conditions for cost minimization require that MTRS be set equal to the ratio of input prices:

$$[\, \delta /(1 - \delta)\,]\, (k_0 / k_p)^{1/\sigma} = p_p / p_0$$

where p_p and p_0 are peak and off-peak prices, respectively. Solving this relationship for the relative intensity of electricity use between peak and off-peak periods, we have:

(3a) $k_p / k_0 = \{[\, \delta /(1 - \delta)\,]\ \ [p_0 / p_p\,]\}^{\sigma}$.

A Strategy for Estimation

By multiplying the right-hand-side of equation (3a) by an appropriate error term (ε), and take the logarithms of both sides, we can obtain an unbiased, minimum-variance estimate of σ using ordinary least squares (OLS):

(4a) $\ln [\, k_p / k_o\,] = \sigma \ln [\, \delta /(1 - \delta)\,] + \sigma \ln [p_0 / p_p\,] + \ln \varepsilon$.

The parameter σ measures the proportional change in the ratio of electricity use in peak and off-peak periods due to a percentage change in the inverse price ratio. For this

[30] The algebra needed to derive this relationship, along with the derivation of the elasticity of substitution, is found in Fergurson (1969, pp. 103-04) and is not repeated here.

production function to be well behaved, Ferguson (1969) shows that $0 < \sigma < \infty$.[31] The higher σ is, the more responsive energy use is to changing in relative prices between peak and off-peak periods. For example, if $\sigma < 1$, then as the price ratio changes by one percent, the ratio of peak to off-peak energy use changes by less that one percent. For $\sigma > 1$, the ratio on energy use changes by more than one percent as the inverse price ratio changes by one percent.

The estimated constant term from equation (4a) is, $a = \sigma \ln [\delta /(1 - \delta)]$. To recover δ for a given estimate of σ we know that $a / \sigma = \ln [\delta /(1 - \delta)]$. Alternatively,

(5a) $[\delta /(1 - \delta)] = e^{a/\sigma}$,

(6a) $\delta = (1 - \delta) e^{a/\sigma}$,

(7a) $\delta = e^{a/\sigma} - \delta e^{a/\sigma}$, and

(8a) $\delta = (e^{a/\sigma}) / (1 + e^{a/\sigma})$.

Thus, with OLS, we are able to identify all the parameters of the CES function, with the exception of γ. Recovering a value for δ is of little use for analyzing the RTP data provided to us by CSW, but it may be critical in the simulations of firm behavior on which to design price-responsive rate programs.

One potential disadvantage of the CES specification is that the elasticity of substitution is constant—invariant with respect to initial peak relative to off-peak electricity usage or to the initial relative prices. This is inconsistent with the view by some that when peak usage is high relative to non-peak usage, a customer may find it

[31] This relationship shows that σ is the proportional change in the use of electricity in the peak period relative to the off-peak period (holding output, in this case the electricity aggregate, constant), as the inverse price ratio increases or decreases by one percent Fergurson (1969, pp. 103-04).

easier to shift load in response to a change in relative prices, but that this ability declines as more and more load is shifted.

The Empirical Specification

As in the text above, we must for the CES case also define exactly how the variables used in the empirical regression analysis are calculated from the data. From equation (4a), we need to have the ratio of peak to off-peak electricity use. Since RTP rates are designed to be revenue neutral if a firm's energy use replicates the firm's CBL, it is reasonable to expect that changes in electricity use should be measured relative to the CBL. For this reason, we define for each weekday, t, and firm or group of firms, m:

k_{ptm} = peak kWh / peak CBL;

k_{0tm} = off-peak kWh / off-peak CBL;

p_{ptm} = average hourly peak price / kWh; and

p_{0tm} = average hourly off-peak price/kWh.

There are also several other variables that must be included in the model; they must be defined specifically. One set contains 0-1 or 'dummy' variables for each firm or group of firms. These variables are to account for differences by firm in peak relative to off-peak energy use because of alternative types of production processes, differences in production shifts, differences in business hours, etc. These variables are defined for the m firms (m = 1,…,M):

D_m = 1 if the observation is for firm m, and = 0 otherwise.

We also try to capture the fact that LR customers may respond to prices differently by defining a dummy variable for them:

D_{LR} = 1 if the observation is for a LR customer, and = 0 otherwise.

Finally, the effect of daily weather on the ratio of peak to off-peak electricity use is captured through a weather index:

w_{tm} = weather index for day t from the weather station nearest to firm m for which there are data (see Appendix B).

The full model can now be specified as (for all observations across time t):

$$\text{(9a)} \quad \ln (k_{ptm} / k_{0tm}) = a + \sum a_m D_{tm} + b\, T_m + c\, w_{tm} + d \ln (p_{0tm} / p_{ptm})$$

$$+ f\, D_{LR} [\ln (p_{0tm} / p_{ptm})] + g\, w_{tm} [\ln (p_{0tm} / p_{ptm})] + \ln e_{tm}$$

In this most general form, both the distribution parameter, δ, which is embodied in the parameter 'a' of equation (9a) differs by firm, time in the program and weather. That is for $D_{tm} = 1$ we have:

$$\text{(10a)} \quad a_{tm} = a + a_m + c\, w_{tm} .$$

In this specification, there will be a separate intercept for each firm and each value of the weather index. These variables affect the relative level of usage between peak and off-peak periods, but not the rate at with usage responds to price.

More important, the price response, σ, also depends on some of these other variables. For $D_{LR} = 1$, we have the relevant logarithmic partial derivative given by:

$$\text{(11a)} \quad \partial [\ln (k_{ptm} / k_{0tm})] / \partial [\ln (p_{0tm} / p_{ptm})] = \sigma_{tm} = d + f + g\, w_{tm} .$$

This specification implies that the price response differs by LR firm and weather.[32] Normally a_{tm} and σ_{tm} would be evaluated at the means of w_{tm}. They could also be evaluated at monthly means, etc. During the summer peak months—June, July, August, and September—one would expect extremely hot weather to reduce a firm's ability to

[32] This is a model in which the elasticity of substitution is affected by production processes, weather, or other factors specific to the firm, Z_i. It is a simple extension of the CES model, and as Caves and Christensen (1980b) demonstrate algebraically that the modification is accommodated in the conceptual model by replacing ρ in equation (1a) with $\rho + \sum_i \gamma_i Z_i$.

substitute electricity between peak and off-peak periods. Thus, we would expect g to be negative. We would also expect f to be positive due to the LR program's rate design.

The Empirical Results

The estimated CES models correspond to the same eight peak periods as are estimated with the GL specification discussed above in the text. This CES model is a simpler specification than the GL, and the results are presented here primarily for comparative purposes. In part because the model is simpler, it is also possible in the CES model to recover estimates of the electricity aggregate, q, from equation (1a). This is done by solving for the remaining parameter, δ, using equations (8a) and (10a). As is seen in equation (10a), it is a relatively complex exercise to calculate the individual δ's when there are dummy variables for the individual firms. Therefore, since we have no real use for these parameters in this particular paper, they are not provided. Having consistent estimates of q, however, may be important in simulating customer behavior and in rate design. Thus, this capacity to calculate q may be one reason to use the CES model if the elasticity estimates are not widely different from those derived from a GL or other flexible cost function.

The performance of each of these CES models is similar to that of the corresponding GL model. The explanatory power falls for those models representing the longer peak periods, but the t-ratios on the coefficients remain high (Table A1). Hot weather also increases the peak to off-peak energy usage, but the index did not affect the size of the elasticities of substitution.

In this model, the elasticities of substitution are constant, and they are equal to the coefficient on the logarithm of the inverse price ratio for the non-LR customers. Those elasticities for the LR customers are derived by adding to that coefficient the coefficient for the interaction term between the logarithm of the inverse price ratio and the LR dummy variable (e.g., equation (11a)).

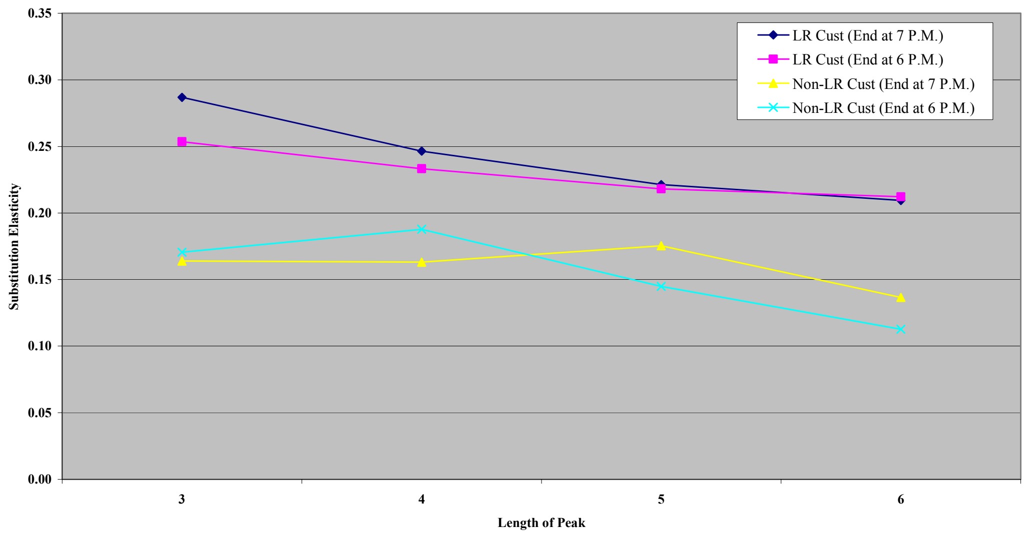
As can be seen in Figure A1, the patterns of elasticities of substitution for the CES model are very similar to those for the mean values of the GL model discussed above. One useful way to compare the two sets of elasticities is to determine if the range in the GL elasticities encompasses the CES values. In all cases for both LR and non-LR customers, the estimated CES elasticities of substitution are within the range of the GL estimates, and are also just slightly larger than the average of the GL elasticities. This pattern clearly derives from the inherent inflexibility of the CES model, but it is encouraging to know that the CES model seems to consistently overestimate the average response derived from a more flexible, but much less computationally tractable, model.

Table A1. PSO CES Model Results w/ Weather Index

	Parameter Estimates (T-Statistic)*				
Peak Hours (p.m.)	Intercept	Log PCES	Log WIDX	LR PCES	R^2
4 through 6	-0.372 (-0.57)	0.164 (8.23)	0.207 (1.42)	0.123 (3.53)	0.479
3 through 6	-0.630 (-0.99)	0.163 (9.53)	0.207 (1.45)	0.083 (2.79)	0.494
2 through 6	-0.784 (-1.26)	0.175 (11.44)	0.204 (1.46)	0.046 (1.69)	0.510
1 through 6	-0.801 (-1.46)	0.136 (10.63)	0.212 (1.73)	0.073 (3.12)	0.412
3 through 5	-0.729 (-1.09)	0.171 (9.34)	0.220 (1.47)	0.083 (2.53)	0.412
2 through 5	-0.900 (-1.4)	0.188 (11.35)	0.250 (1.74)	0.046 (1.52)	0.451
1 through 5	-0.916 (-1.62)	0.145 (10.11)	0.227 (1.79)	0.073 (2.75)	0.482
Noon through 5	-1.084 (-2.08)	0.113 (8.23)	0.254 (2.16)	0.099 (3.87)	0.364

*This equation is (4a) and it is estimated in SAS by OLS. The variable LOG PCES is the loarithim of the off-peak to peak price ratio. The variable LR PCES is the product of this price ratio and a zero-one variable, where the one refers to a load relief customer. The variable LOG WIDX is the logarithm of the heat index. The dependent variable is the logarithm of peak to off-peak load, both normalized by CBL.
The peak period is for the hour beginning. For example, the peak 4 through 6 begins at 4:00pm goes through the 6:00pm hour and ends at 7:00pm.
Note: If there is a line under (over) the coefficient on LOG PCES, then the CES elasticity of substitution for non-LR customers is below (above) its range for the GL model.
Note: If there is a line under (over) the coefficient on LR PCES, then the CES elasticity of substitution for LR customers (sum of coefficients on LOG PCES and LR PCES) is below (above) its range for the GL model.

Figure A1. CES Model Estimated Substitution Elasticities



Appendix B

Development of Weather Variables

We initially obtained historical weather data for five weather stations in CSW's service territory from the National Climatological Data Center (NCDC) Internet site. These weather stations were selected in consultation with CSW staff and correspond to the service territories of the four relevant CSW operating companies. The sites and corresponding operating companies are: Abilene, TX (WTU),Tulsa, OK (PSCO),Corpus Christi, TX (CP&L),Shreveport, LA (SWEPCO - Louisiana/Texas),Fort Smith, and AR (SWEPCO – Arkansas).

The data set encompassed the period June of 1998 through August of 2000 and contained daily mean temperature and dew point values. These were used to calculate heating and cooling degree-days and heat indices on a daily basis.

Variable Construction

The following formulae summarize the calculation of the variables employed in the regression models. These are based on statistics developed by the National Weather Service. Note that the derivation of the Heat Index required several intermediate steps: a) conversion of the temperature and dew point values to Celsius; b) calculation of actual and saturation vapor pressure; and c) calculation of relative humidity.

This was necessary since relative humidity (RH) was not available in the NCDC data for the analysis period; and the RH is required to calculate the heat index.

Calculation of Relative Humidity

Tf= Mean temperature (FE)

Tdf= Mean dew point temperature (FE)

Mean temperature (CE) = $Tc = 5/9 * (Tf - 32)$

Mean dew point temperature (CE) = $Tdc = 5/9 * (Tdf - 32)$

Actual Vapor Pressure = $E = 6.11*10.0^{(7.5*Tc / (237.7+Tc))}$

Saturation Vapor Pressure = $Es = 6.11*10.0^{(7.5*Tdc / (237.7+Tdc))}$

Relative Humidity (%) = $RH = (E/Es) * 100$

Calculation of Degree-Day Indices

Heating degree-days (Base 65)(HDD65)

if Tf >=65, HDD65 = 0

if Tf <65, HDD65 = 65 - Tf

Cooling Degree-Days (Base 65) (CDD65)

if Tf <=65, CDD65 = 0

if Tf >65, CDD65 = Tf - 65

Heat Index (HI70)

if Tf <=70, HI70 = 0

if Tf >70, HI70 = - 42.379

$$
\begin{aligned}
HI70 = &- 42.379 \\
&+ 0.04901523*Tf \\
&+ 10.14333127*RH \\
&- 0.22475541*Tf*RH \\
&- (6.83783*10^{-3})*(Tf^{2}) \\
&- (5.481717*10^{-2})*(RH) \\
&+ (1.22874*10^{-3})*(Tf^{2})*(RH) \\
&+ (8.5282*10^{-4})*Tf*(RH^{2}) \\
&- (1.99*10^{-6})*(Tf^{2})*(RH^{2}).
\end{aligned}
$$

Appendix C

Characterizing Energy Conservation vs. Shifting

To motivate this discussion of how customers respond to high prices, it is important to remember that as the price of on-peak electricity rises (*ceteris paribus*), electricity becomes a more expensive input for customers. Therefore, in addition to the predicted substitution effect, the fact that production costs are now higher, may on a particular day, lead to an output effect as well. There are three, and possibly four, cases to be distinguished; they are characterized in Figures C1 and C2.[33] In Figure C1, the curve E_0 represents combinations of the inputs K_p and K_o that produce an energy aggregate that support the firm's desired (and constant) output. At the expected price ratio of P_p/P_o, the firm would minimize the cost of producing E_0 by using K_{p1} and K_{o1} of peak and off-peak electricity respectively. If there is an increase in the peak period price, to $P_p^* > P_p$, the price line gets steeper and if the firm is to continue to produce E_0, the minimum cost way of doing so is by using more electricity off-peak and less on-peak (e.g. $K_{p2} < K_{p1}$and $K_{o2} > K_{o1}$), resulting in a decrease in the ratio of peak to off-peak. This change in input intensity is measured by the elasticity of substitution; depending on its value, total electricity use could increase or decrease.

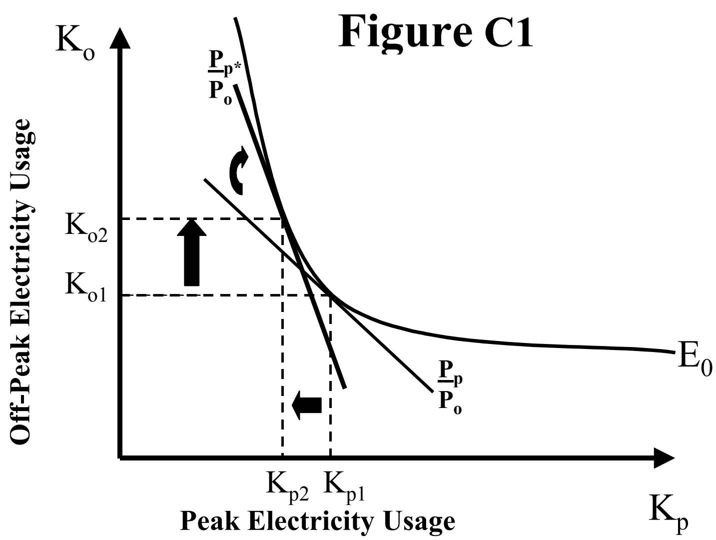
The other situations that can arise from an increase in the peak price of electricity are more complex, because they imply either a decrease in a firm's output, or a reduction in peak usage that might be viewed as discretionary (e.g., a customer elects to reduce employee or occupant comfort by increasing the thermostat setting by 2-4 degrees while still maintaining the firms' output). Figure C2 illustrates a situation in which the firm reduces output by producing a smaller amount of the energy aggregate. Here the initial point of production is on E_0 similar to Figure C1. Now, if the price of peak electricity rises high enough, the firm may need to reduce output by moving to a lower energy aggregate isoquant, E_1. Energy use is now at $K_{p3} < K_{p1}$ and $K_{o3} > K_{o1}$. The ratio of peak to off-peak electricity use has fallen, but without more information, we could not know what happened to total electricity use. If the isoquants in this diagram had a greater curvature (indicating a lower elasticity of substitution), it is possible that the response to this dramatic peak price increase would lead to a lower peak to off-peak usage ratio, but result in lower use in both peak and off-peak periods.[34]

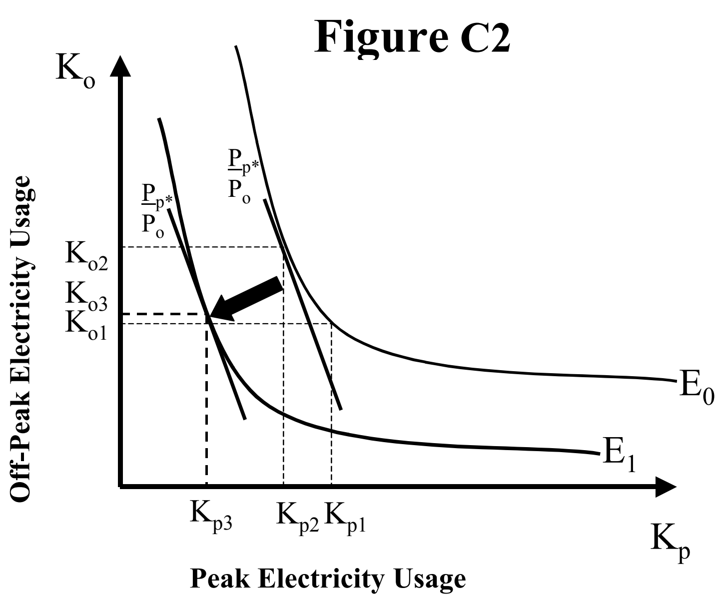
If the increase in peak price is not sufficiently high, then the firm might decide to reduce discretionary peak usage, as suggested above, but maintain output. This could be characterized in Figure C2 by assuming that the portion of the isoquant E_1 below, and to the right of the initial off-peak electricity use of K_{o1}, actually represents an energy aggregate of E_0; output does not change with a reduction in discretionary peak usage, only employee comfort levels do. Thus, the response to an increase in peak price to P_p^* would be for peak use to fall from K_{p1} to K_{p3} and off-peak use would remain at K_{o1}. The

[33] Given that we have assumed electricity is separable in production, any combination of peak and off peak electricity (k_p and k_0) that when substituted into equation (2) yield the same value of the electricity aggregate q is equally productive when used to produce the firm's output Q from equation (1).

[34] The increased curvature in the isoquants could well be due to indivisibilities in electricity use, such as having to shut down equipment or processes for a longer period (e.g., an entire shift that includes peak and some off-peak hours) in their efforts to reduce usage during high-priced peak hours. In the extreme case where peak and off-peak energy must be used in fixed proportions, there is no possibility of substitution, and the isoquants are rectangular (Ferguson, 1969). In this case, the only way to adjust energy consumption in response to high peak prices is to reduce consumption in both periods proportionately. Reduction in off-peak use might also reflect a "good citizen" ethic on the part of the customer. They may reduce peak usage because high prices are often associated with conditions where system reliability is jeopardized and a public appeal may have been issued to customers urging them to lower consumption (conserve). These customers may then turn off devices for hours that extend well beyond the period of

ratio of peak to off-peak electricity use falls. Since electricity use in the off-peak has not changed, total daily usage has declined.





high prices (i.e., the peak period). The consequence of these actions is that the customer's total daily load is reduced proportional to the peak reduction.