Volume Title: Social Experimentation

Volume Author/Editor: Jerry A. Hausman and David A. Wise, eds.

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-31940-7

Volume URL: http://www.nber.org/books/haus85-1

Publication Date: 1985

Chapter Title: The Residential Electricity Time-of-Use Pricing Experiments: What Have We Learned?

Chapter Author: Dennis Aigner

Chapter URL: http://www.nber.org/chapters/c8372

Chapter pages in book: (p. 11 - 54)

1 The Residential Electricity Time-of-Use Pricing Experiments: What Have We Learned?

Dennis J. Aigner

1.1 Introduction

Over a period of six years, the Department of Energy (DOE) has been engaged in a cooperative program of residential time-of-use (TOU) rate experiments, involving time-of-day or seasonally varying prices. There are fifteen completed or ongoing projects, the first of which began in 1975 in Vermont.

The main goal of this program of experimentation was to determine whether TOU pricing would produce sufficient alterations in the load curves of residential customers to justify implementation of such rates. This "justification" involves three specific effects—the revenue impact on the utility, the amount of capacity reduction implied, and changes in consumer welfare. The experiments were designed, to a lesser or greater degree, to address one or more of these issues.

A number of design considerations have an impact on the ultimate usefulness of the experimental data that have been forthcoming, not the least of which is the amount of variation available in peak, midpeak and off-peak prices. Many of the DOE experiments have but one set of TOU prices, and therefore the inferences available are limited to a single statistical comparison of control-group and experimental households.

Dennis J. Aigner is professor of economics, University of Southern California.

This paper was prepared for presentation at the NBER Conference on Social Experimentation, 5–7 March 1981, Hilton Head, South Carolina. This version has been revised slightly to account for points raised in the general discussion of it as the conference, but not to such extent that the pertinence of the formal discussants' remarks are diminished.

Research on the experimental data discussed herein has continued at a rapid pace. In the interim since this paper was drafted several important pieces of work relevant to the task have appeared that were unable to be included.

While in other respects such an experiment may be well designed, its results are not generalizable to a situation where the TOU prices are different than those used in the experiment. Other design or sampling issues of some importance include the type of stratification used, the choice of a model used for an optimal allocation of observations to cells, whether sample size is adequate to allow for estimation of relevant parameters with sufficient precision, the influence of incentive or compensation payments, the means for handling attrition, the nature of the experimental environment, and so forth.

Design issues relate directly to how the data are developed and used for purposes of analysis. In addition, often latitude exists in the choice of a statistical framework for analysis apart from requirements or limitations implied by the data, its collection, and quality.

The purpose of the present paper is to consider the empirical results available so far from the DOE experiments in light of design and analysis concerns and the goals the experiments were ostensibly designed to serve. This is done by focusing on price elasticities as the important summary parameters of interest, since they feed directly into calculations of welfare and revenue impacts and impinge on the matter of rate design itself.

In the following section, the available elasticity results are summarized and discussed. Section 1.3 presents the basis for a welfare analysis of a move to TOU rates. In section 1.4, the question of implementation is considered. Finally, in the concluding section, we take a retrospective view and return to the initial design issues raised.

1.2 Summary of Elasticity Estimates

As might be expected, the early DOE projects are of highly variable quality from the standpoint of being able to make valid statistical inferences from them. Even some of the more recently established projects are of limited value for the purpose of estimating elasticities by time of use.

Table 1.1 summarizes the status of each project and its potential usefulness for estimating price elasticities. In assessing a project's suitability for price-elasticity estimation, the crucial factor in its experimental design is the degree of independent price variation offered to customers. Without rate-structure variation it is difficult to ascertain anything more than qualitative effects on customer demands resulting from the institution of a TOU rate-structure. Consequently, our primary criterion in judging a project's usefulness is the amount of rate-structure variation employed. Other design characteristics are also important, of course, but we regard these as being of secondary interest.

Detailed descriptions and comprehensive evaluations of all aspects of the experiments have been made by Research Triangle Institute (RTI)

Table 1.1 Status of DOE Rate Demonstration Projects

State	Status	Results/Potential
Arizona	ongoing	available; of interest
Arkansas	completed	available; of limited interest
California (LADWP)	completed	available; of interest
California (SCE)	ongoing	partially available; of interest
Connecticut	completed	available; of limited interest
New Jersey	cancelled	not available; of no interest
New York	completed	not available; of no interest
North Carolina (BREMC)	completed	available; of limited interest
North Carolina (CP&L)	completed	partially available; of interest
Ohio	completed	available; of limited interest
Oklahoma	completed	partially available; of interest
Puerto Rico	ongoing	not available; of no interest
Rhode Island	cancelled	available; of no interest
Vermont	completed	available; of no interest
Wisconsin	completed	available; of interest

(U.S. Department of Energy 1978) on behalf of DOE, and by the University of Michigan's Survey Research Center (Hill et al. 1979) on behalf of the Electric Power Research Institute. Thus only a brief description and evaluation of each project will be given here. In the appendix we discuss each project in greater detail and justify our evaluations of them.

The demonstration projects can be classified into two groups on the basis of when they were begun, since only midway into the program did the DOE issue guidelines on statistical sampling, experimental design, etc., to be used by new projects. These guidelines clearly are minimal when judged by the standards of some of the best projects, yet several of the most recently established projects do not meet them in important respects.

The first group oı experiments to be funded took place in Arizona, Arkansas, California (Los Angeles Department of Water and Power—LADWP), Connecticut, New Jersey, New York, Ohio, Vermont, and Wisconsin. The Vermont experiment has such serious flaws that we doubt the results should be utilized even on a local level. The Arkansas, Connecticut, New York, and Ohio experiments all suffer from a lack of variation in price treatments, with at most only two different TOU rate structures employed. The New York and Ohio experiments have additional serious design flaws, while the Arkansas and Connecticut experiments are otherwise well designed. The Arizona, California (LADWP), and Wisconsin experiments are all well designed with a wide variation in price treatments.

The group of newer experiments includes those in California (Southern California Edison—SCE), North Carolina (Blue Ridge Electric Membership Corporation—BREMC and Carolina Power & Light—CP&L), Oklahoma, Puerto Rico, and Rhode Island. The designs of these experiments for the most part benefit from the DOE guidelines, particularly in the crucial area of price variation. Only the Rhode Island experiment fails in this regard, employing a complex but unvarying rate structure that precludes isolation of TOU price effects. The Puerto Rico experiment, despite its favorable rate-structure design, has so many idiosyncratic features that it should be viewed at best as of local interest only. The California (SCE), North Carolina (CP&L) and Oklahoma experiments are all well designed in other respects in addition to their use of several price treatments.

Thus six experiments in various stages of completion offer the wide price variation desirable for estimation of TOU price elasticities. Data from the Arizona and Wisconsin experiments have been generally available for some time, with the Wisconsin data seeming slightly better in quality; a number of demand studies have been conducted utilizing these sources. We report the results of these studies, as well as some studies of the Connecticut experiment which utilize the data with a single price treatment. Data will soon be fully available from the California (LADWP), California (SCE), Oklahoma, and North Carolina tests, but at the time of this writing only preliminary reports on these experiments are available. To the extent that comparative results from those other projects that do not possess price variation but are otherwise well designed are pertinent, they are also discussed.

While the notion of price elasticity (own price, cross price, compensated, etc.) is certainly well understood, elasticity estimation within the context of a statistical model of electricity demand and/or consumption by time of use presents some unique methodological problems, solutions to which are still evolving. Earlier reports from the Electric Utility Rate Design Study (Electric Utility Rate Design Study 1977a, 1977b) contain reviews of the empirical work available on the average price elasticity of demand and introductions to the topic of TOU-demand modeling and elasticity estimation. Each of the studies cited contains a more detailed exposition of the models and methods used. General references are Aigner and Poirier (1979) and Lifson (1980).

The existing empirical elasticity estimates come in a variety of types, distinguished by whether they are compensated or uncompensated, "partial" or "full." The definitions of these elasticities and the concepts they measure are given in Aigner and Poirier (1979) and also in Hendricks and Koenker (1979). In our presentation of empirical elasticity results from the demonstration projects, the main focus is on the uncompensated partial own-price elasticities, the most commonly reported elasticities. After presenting and analyzing these in some detail, we turn to a discussion of other elasticity types and their usefulness.

Estimates from previous empirical work of uncompensated partial own-price peak-period elasticities are presented in table 1.2. Because of the single TOU rate structure employed in the Connecticut experiment, most researchers have not attempted to estimate price effects, but instead have focused on explaining load patterns as functions of socioeconomic variables, weather, and household demographic characteristics. Examples are the studies by Engle et al. (1979), Hendricks, Koenker, and Poirer (1978), and Hausman, Kinnucan, and McFadden (1979).[1] The Engle and Hendricks papers report on attempts to characterize and model the individual load curve.[2] The Hausman paper is insightful because it computes a welfare-based price index for electricity which suggests that the price of electricity went up for households facing the TOU rate (their welfare therefore went down), and thereby explains the reduction in monthly consumption experienced by them.

Lawrence and Braithwait (1977), however, do obtain estimates of price elasticities from the Connecticut data using the linear expenditure system of demand equations. By imposing a restriction on the consumption requirements of household appliances, they are able to surmount the problem of lack of variation in the price data. They find that peak-period elasticities hover around -0.5 while the midpeak and off-peak elasticities are quite a bit smaller.

Data from the Arizona experiment have received attention from several analysts. Indeed, the range of econometric work on this data set is more extensive than any other. Beginning with the work of Taylor (1977) and Atkinson (1979a, 1979b), further refinements to the econometric methodology and additional results were contributed by Aigner and Hausman (1980), Lau and Lillard (1979), and Hill et al. (1980). In DOE's assessment of the Arizona experiment, Miedema et al. (1978) also estimate TOU price effects.

Taylor's numerical results are not reported in the table because they are so anomalous. He fits both a straightforward linear-regression model and a model employing relative usage (for example, monthly kWh consumption in the peak period relative to total monthly consumption) as its dependent variable. In each instance the price variables perform poorly. There are no statistically significant price coefficients in any of the three equations, and in the peak and midpeak equations most of the own-price terms have positive signs. Income, temperature, and the measure of

1. The White et al. (1978) Research Triangle Institute report on the Connecticut experiment is also in this vein, since they too make no attempt to estimate price effects. They do not try to model the load curve or any aggregate thereof, but merely examine shifts in TOU consumption in response to the implementation of the TOU rate structure. They find that the TOU rates do cause some shifting of consumption out of the peak period.

2. The Hendricks, Koenker, and Poirer (1978) paper uses a methodology for approximating the load curve by a spline function. See Electric Utility Rate Design Study (1977a) for an introduction to these ideas.

Table 1.2 **Uncompensated Partial Own-Price Elasticities of Electricity Demand by Time of Day**

	Connecticut		Arizona		Wisconsin	
	Narrow Peak (4 hrs.)	Broad Peak	Narrow Peak (3 hrs.)	Broad Peak (5–8 hrs.)	Narrow Peak (6 hrs.)	Broad Peak (9–12 hrs.)
Peak Period Summer			−0.41LL −0.18HOTW −0.17AH		(−0.41, −0.66) −0.81	(−0.48, −0.84)CC (−0.81, −0.83)ATK
			−0.64	(−0.69, −0.79)AT		
Winter	(−0.46, −0.66)					
Midpeak Period Summer			−0.26LL −0.24HOTW −0.47AH			
			−0.70	(−0.38, −0.58)AT		
Winter	(−0.24, −0.50)					

Off-Peak Period		−0.46LL			(−0.51, −0.77)	(−0.30, −0.64)CC
Summer		−0.19HOTW	−0.36AH		−0.09	(−0.21, −0.24)ATK
		−0.23		(−0.35, −0.57)AT		
Winter	(−0.29, −0.36)					

Sources: Connecticut: Lawrence and Braithwait (1977, table 6, p. 74). Arizona: Entry AT is from Atkinson (1979b, table 10, p. 92); AH from Aigner and Hausman (1980, table 4, p. 18); LL from Lau and Lillard (1979, table 3, p. 27); HOTW from Hill et al. (1980, table 4, p. 21). Wisconsin: Entry CC is from Caves and Christensen (1980b, tables 6 and 7); ATK is from Atkinson (1979b, table 11, p. 93).

Notes: Connecticut: Ranges for Connecticut are constructed over results for the months of November 1975 and January 1976, and over subperiods of the day (two of which were designated "peak," three "midpeak," one "off-peak"). The elasticity estimates given here are those estimated using the MIN assumption concerning appliance kWh requirements (see Lawrence and Braithwait 1979, 69).

Arizona: Since Atkinson (AT) pools the data over months, the ranged are constructed for the broad peak period only. Aigner and Hausman (AH) attempt to correct for truncation bias in their results. They use the one summer month of August 1976. Moreover, since length of the peak period is an independent variable in their model, a single elasticity is reported, although separate elasticities for individual pricing periods could have been calculated. Lau and Lillard (LL) work only with households that faced the narrow peak period, over the period May–October 1976. They pool the data over months. Hill et al. (HOTW) also attempt to correct for the influence of the incentive scheme in their work.

Wisconsin: Ranges were constructed by Caves and Christensen (CC) over the results for the months of July and August 1977, over alternative definitions of the peak period (in the case of "broad peak"), and over alternative TOU rate differentials from 2:1 to 8:1. Since Atkinson (ATK) pools the data over two available months and over prices, the ranges constructed are for the broad peak period only.

applicance capacity are significant explanatory variables in Taylor's equations, which would lead one to conclude that nonprice variables completely overwhelm price influence in explaining customers' response patterns.

Miedema et al. (1978) obtain similar results using monthly data over the same period as Taylor.[3] They estimate four different models of electricity consumption by time-of-day, employing as regressors prices, income, demographic characteristics, weather, lagged consumption, and allowance of prices to interact with the other variables. In none of the models, estimated individually for each of six months, do they obtain a single significant price-elasticity estimate.[4]

By way of contrast, Atkinson, Aigner and Hausman, Lau and Lillard, and Hill et al. all estimate significant price effects, although their estimates are dissimilar. These researchers take a neoclassical utility-maximization approach to the estimation problem and make a separability assumption about electricity consumption.[5] Their resulting partial elasticity estimates are presented in table 1.2.

Atkinson's (1979a) results are based on a translog model estimated over the same six-month period as Taylor and Miedema, et al. His results show significant own-price elasticities for all periods. When the peak is broadly defined, Atkinson finds the peak-period elasticity to be around -0.7, larger (in absolute value) than those corresponding to the midpeak and off peak. For the narrowly defined peak he finds that midpeak elasticity increases to a level slightly above the peak elasticity. The conclusion from Atkinson's work is that manipulating the on-peak price would result in the largest quantity response, at least when the peak is broadly defined—as opposed to an alternative whereby the on-peak/off-peak differential is changed by altering the off-peak price.[6] Although Atkinson's empirical work can be faulted on several grounds, subsequent attempts by others to improve on the econometric model and methods he employed still present a sharp contrast to the conclusions of Taylor and Miedema et al.

The work of Aigner and Hausman (1980) represents an interesting methodological contribution and contains empirical results that run

3. These results are similarly omitted from table 1.2.

4. In an appendix to their study, Miedema et al. (1978) employ a translog model and obtain for the most part negative and significant elasticity estimates. They dismiss these, however, as being mere artifacts of the assumptions inherent in this model specification.

5. For more on this approach see Aigner and Poirier (1979) or Hendricks and Koenker (1979).

6. Although the TOU experiments focus on manipulation of prices, there are other interrelated, controllable "parameters" of interest, such as the differential itself, length and/or starting time of the peak period, etc. Their corresponding elasticity measures are likewise interrelated. See appendix A of Aigner and Poirier (1979) for details on these matters.

counter to Atkinson's in both the computed levels of uncompensated elasticities and in their patterns. Aigner and Hausman attempt to correct the Arizona data for potential biases due to one aspect of the compensation scheme which protected sample households from paying any more for electricity under experimental rates than they would have paid under the prevailing (control) rate. Their results suggest that for these Arizona households the midpeak own-price elasticity is largest. The econometric specification they use is a powerful one.[7] With it, Atkinson's Arizona results are reconciled to those emanating from the Wisconsin project, which will be discussed shortly.

Similarly, Lau and Lillard (1979) offer a substantial methodological improvement over Atkinson's work—while keeping to his model in most other respects—by specifying a rich "pooling" model to exploit the cross-sectional/time-series nature of experimental data. They find the peak and off-peak elasticities to be about the same magnitude, with a relatively small midpeak elasticity. But there are shortcomings in their work (possibly only minor in nature) which need to be resolved before firm conclusions can be drawn.

A study related to the Aigner and Hausman work is the recent paper by Hill et al. (1980). These authors attempt to account for the effects of the rather complicated incentive structure faced by the experimental households in 1976. While Aigner and Hausman concentrate on the so-called maximum constraint, whereby households knew they would pay no more for electricity under TOU prices than they would have under standard rates, Hill et al. look at the more detailed aspects of the incentive scheme. If a household consumed the same amounts of electricity in each TOU pricing period in any month in 1976 as it did in the corresponding month in 1975, its 1976 monthly bill would have been 85 percent of the 1975 bill. In fact, very few customers bumped up against the maximum constraint during the experiment, so consideration of the exact details of the incentive scheme may demonstrate some further insights into the effects of it on TOU price response.

What Hill et al. find, using a different model than Aigner and Hausman (the same model used by RTI), is a similar pattern of uncompensated own-price elasticities to that reported by Aigner and Hausman, in that the midpeak elasticity is largest, followed by off-peak and on-peak values. The statistically significant cross-price effects are also negative, but the magnitudes of all these elasticities are quite different between the two studies.

Hill et al. also analyzed the 1977 data, where many changes were introduced in the experimental design, including elimination of the incen-

7. Their handling of socioeconomic variables is quite different from Atkinson's, although both use the same translog demand model.

tive scheme (the maximum bill constraint was still in force). Some differences in the estimated regression equations were evident, comparing the 1977 results to 1976. In particular, more of a tendency toward shifting peak-period kWh's into the midpeak and off-peak periods was observed.

Turning to the Wisconsin experiment, Atkinson (1979b) and Caves and Christensen (1980a; 1980b) use only two of the first available months of data when households were on experimental rates.[8] Thus their quantitative results must be viewed as very tentative. Atkinson processes the Wisconsin data through the same model he used to analyze the Arizona data, and with similar results. He finds a substantially higher on-peak own-price elasticity than for the off-peak period, whether the latter is broadly or narrowly defined.

Caves and Christensen approach the data cautiously, employing several demand models consistent with economic theory. In their earlier (1980a) study they evaluate three alternative models, rejecting the translog in favor of the constant elasticity of substitution (CES) and generalized Leontief systems (although they present elasticity estimates only for the latter). Their broad-peak results are similar to those of Atkinson in that the on-peak elasticity is larger (in absolute value) than the off-peak elasticity. Their narrow-peak elasticity results, however, show the on-peak elasticity to be smaller than the off-peak figure, a pattern that, while different from the one found by Atkinson, is in accord with Aigner and Hausman's result.[9] In their later (1980b) study Caves and Christensen employ the CES demand model and extend it to allow for the effects of nonprice variables like consumption level, appliance stocks, and household characteristics. They find that in addition to prices, appliance stocks have significant effects on time-of-day demands. Unfortunately, they present no own-price elasticity estimates, preferring to focus on the substitution possibilities between peak and off-peak consumption.

It is difficult to summarize the empirical results given in table 1.2 since the elasticity estimates frequently conflict with each other. While no consistent overall pattern emerges from the table, the estimates support the following conclusions. First, there is agreement that all the peak-period demands are inelastic when expenditure is held constant. Second, in all cases where the peak period is broadly defined, the researchers find that peak-period elasticity exceeds the midpeak and off-peak elasticities. This pattern, however, does not generally hold for the narrow-peak

8. Caves and Christensen (1980b) also use data on customers in the statistical control group for the same two months of the preceding year to improve the efficiency of their estimates.

9. Although Aigner and Hausman's finding is an overall and not just a narrow-peak-period result.

results or even for the Aigner and Hausman results which do not vary by peak length.

We now turn to a discussion of other elasticity types. Table 1.3 presents compensated partial own-price elasticity estimates from the previously cited empirical studies. These estimates are in all cases smaller than the corresponding uncompensated partial elasticities, a consequence of the fact that electricity has a positive expenditure effect.[10] Since these elasticities hold utility constant when prices change, they are of particular usefulness in analyzing time-of-day pricing policies where the goal is to induce consumers to shift their consumption patterns without causing them any loss in welfare. The compensated elasticity provides a measure of how much consumption would change in response to a price change if consumers' electricity expenditures were simultaneously adjusted (e.g., via a lump-sum payment that could only be used to purchase electricity) to prevent their welfare from changing. Unfortunately, many of the studies do not report compensated elasticities. The estimates that are reported are quite low, especially those obtained by researchers using the Arizona and Wisconsin data.

Cross-price elasticities are also estimated by a number of studies, and they are crucial to understanding how price changes affect the whole pattern of consumption by time of day. Because of the large number of cross-price elasticity estimates that would need to be reported, we refrain from presenting them in detail and opt instead for a more general discussion of the results. As with the own-price elasticities, the cross-price elasticity estimates vary widely across studies. Lawrence and Braithwait (1977) obtain small but positive estimates using the Connecticut data, indicating a slight degree of substitutability between electricity consumption by time of day. Working with the Arizona data, Atkinson (1979b) and Aigner and Hausman (1980) find, however, that the cross-price elasticities are generally negative, implying gross complementarity of time-of-day consumption. Caves and Christensen (1980b) also obtain negative uncompensated cross-price elasticity estimates with the Wisconsin data, although their compensated cross-price elasticity estimates are positive.

Almost all the studies report "partial" rather than "full" elasticities, the former not accounting for the indirect effects that price changes have on consumption through their effect on electricity expenditure. The full elasticities correspond to the common notion of the term "elasticity," and a compelling case is made by Hendricks and Koenker (1979) for them being the appropriate measures for public policy use (see Hendricks and

10. See Aigner and Poirier (1979, 9, equation [2.11]), for the relation between the uncompensated and compensated partial elasticities.

Table 1.3 **Compensated Partial Own-Price Elasticities of Electricity Demand by Time of Day**

	Connecticut		Arizona		Wisconsin	
	Narrow Peak (4 hrs.)	Broad Peak	Narrow Peak (3 hrs.)	Broad Peak (5–8 hrs.)	Narrow Peak (6 hrs.)	Broad Peak (9–12 hrs.)
Peak Period						
Summer			−0.01LL		−0.01	(−0.03, −0.04)ATK
Winter	(−0.15, −0.32)					
Midpeak Period						
Summer			−0.11LL			
Winter	(−0.18, −0.44)					
Off-Peak Period						
Summer			−0.01LL		−0.11	(−0.00, −0.01)ATK
Winter	(−0.19, −0.28)					

Sources: Connecticut: Lawrence and Braithwait (1977, table 7, p. 75). Arizona and Wisconsin: See table 1.2.

Notes: Arizona: Neither Atkinson (1979b) nor Aigner and Hausman (1980) report compensated elasticities. Wisconsin: Caves and Christensen (1980a) do not report compensated elasticities. See table 1.2 for further explanations.

Koenker 1979, 27). The problem lies in converting partial to full elasticities.[11] This procedure requires knowledge of the elasticity of electricity expenditure with respect to the price of electricity, so reliable estimates of the latter must be found. Caves and Christensen (1980a) and Hendricks and Koenker (1979) both compute full elasticities from partial ones, but both must rely on an assumed value of the expenditure elasticity.

Besides the projects reviewed herein, Research Triangle Institute has provided a summary report (Miedema and White 1980) that covers the efforts in Ohio, Rhode Island, and one of the North Carolina experiments (Blue Ridge Electric Membership Corporation). In these experiments, as in the Connecticut test, only one TOU rate was applied. In Ohio, the sample was small and the ultimate sample design so tenuous that the results of the test should probably be discounted entirely.

In the North Carolina BREMC test, again the sample was small (roughly one hundred experimental households) and one rate schedule was considered. In this experiment, which ran for twelve months only, no perceptible alterations from control-group behavior were observed in the test group.

The Rhode Island experiment provided an additional wrinkle to the menu of experimental tariffs by using a time-differentiated demand charge as well as the usual time-varying kWh prices. This test ran for thirteen months. As in the BREMC case, RTI found essentially no statistically significant TOU effects in Rhode Island.

These comparative results, along with their analyses of Arizona, Connecticut, and Wisconsin, prompted RTI to conclude:

> All studies showed some reduction in usage during the peak period under TOU rates. However, reduction in usage during the peak period was not accompanied by statistically significant increases in base-period usage. Total usage seemed either to decline or remain the same in all projects. . . . Peak-day usage shifts and average-day usage shifts appeared to be about the same. (Miedema and White 1980, 4)

Preliminary work on the Oklahoma and North Carolina (CP&L) experiments was recently made available at a DOE-sponsored conference in Denver. While no elasticity estimates were presented, there were summaries of the qualitative findings to date. In Oklahoma, Brown et al. (1980) report no significant TOU or seasonal effects. Flat rates proved to yield some conservation response compared to declining block rates, but otherwise no conservation or differential effects overall or by TOU price ratio were observed.

11. This conversion is not necessary if the demand model is formulated in such a way that the elasticities directly estimated have a "full" rather than a "partial" interpretation. However, as Hendricks and Koenker (1979) point out, these "simple" demand models have deficiencies that may very well more than counteract this benefit.

In the CP&L test, the RTI group (Miedma 1980) conclude that while some evidence of peak period and overall reductions in consumption can be found, for the most part these effects are not statistically significant to those test groups facing a two-part (customer and energy charge) tariff. With demand charges included, many of the TOU response effects are significant.

To date, there is but limited published work relating to the LADWP (Los Angeles) experiment being conducted by Rand, even though the test itself began in the summer of 1976. Some elasticity results are reported in Acton and Mitchell (1980). The primary use of these estimates is to fuel an illustrative cost-benefit calculation that provides the machinery for evaluating the question of TOU rate implementation, not to present an exhaustive collection of values for various rating periods, customer groups, etc.

What Acton and Mitchell do is to use the elasticity estimates in order to compute the welfare and bill changes that would occur to residential customers under an illustrative TOU tariff, organized by consumption level. In this way it is made clear that only for large users ($\geq$ 1100 kWh/month) will the welfare benefits be sufficient to offset the additional costs of metering. For households with swimming pools, the verdict on TOU pricing is favorable for households at lower consumption levels (80 percent of the sample households with average monthly consumption $\leq$ 1100 kWh own pools).

An interesting sidelight of the Acton and Mitchell analysis is a brief discussion of the possible ramifications of going to a voluntary TOU rate. Customers who will experience bill reductions in excess of the cost of metering are likely volunteers, but they need not also be the customers who should be included from the point of view of welfare analysis (economic efficiency). For example, for the set of customers for whom it is beneficial in a welfare sense to be put under a TOU rate (monthly consumption $\geq$ 1100 kWh), most would experience a bill increase at initial consumption levels after incorporating added metering costs. Thus a voluntary program might discourage those customers who are prime candidates for inclusion unless they can be persuaded by the fact that their bills could be lowered by shifting sufficient usage out of the peak period.

That the target customer for implementation of TOU rates in Los Angeles is a large user or a user with a pool may be idiosyncratic of the service territory under consideration to a great extent. So while we may not encourage generalization of the Acton and Mitchell results to New York City, they should be roughly applicable in the southwestern part of the country. Their points about voluntary versus mandatory implementation and their techniques of analysis are worthy of widespread attention.

In a related study, Lillard and Acton (1980) analyze the seasonal-pricing portion of the LADWP experiment. This experiment involves 624 households. Using data covering twenty-three months through September 1977, Lillard and Acton find, in comparing the test customers to the statistical control group, that summer usage is decreased and winter usage is increased in response to the differentially higher summer price, but that the estimated price elasticities associated with these responses are very small—on the order of $-.06$. In an analysis of anticipated bill and welfare changes, again the estimated effects are so small that very little is to be gained on a per customer basis. But even so, over a large population the net gain (in welfare terms) can be consequential.

While a preliminary report on the California (SCE) experiment has been written (Aigner and Lillard 1980), to date no results from this test have been made public.

Finally, at this point not much is known about the behavior of peak demand (coincident or otherwise) as a function of (kWh) price. Moreover, reported results in all the studies apply generally to the "average weekday," not to the relatively few individual days on which the system peaks and similar extremes occur. The Connecticut data have been manipulated satisfactorily in this respect by considering changes in the customer's entire load curve, but owing to the lack of price variation in the experiment the resulting reductions in peak demand are difficult to generalize beyond the particular experimental conditions observed. This is not to say that information on daily peak-demand effects is not generally available. It is in any experiment where fifteen-minute data were collected. But to date almost none of the analysis work has paid it any attention, relying one supposes on the assumption that if peak-period kWh's are reduced, so must be the level of instantaneous demand.

1.3 Welfare Analysis of TOU Pricing[12]

In assessing the potential for gains and losses to customers who move to a TOU pricing scheme, it is not sufficient merely to examine bill effects. Projected bill changes may have a great deal to do with whether or not customers choose to accept a voluntary program of TOU rates but they are only part of the total impact a customer experiences. The other part of the total impact emanates from quantity adjustments per se. Granted that consumption of both peak and off-peak electricity has positive marginal benefit (more consumption implies more satisfaction), reducing peak-period consumption in response to its higher price, for example, must result in less satisfaction. The way these matters are handled in conven-

12. In preparing this section we have borrowed liberally from Acton and Mitchell (1980).

tional economic theory is through the concept of consumer's surplus. The basic idea is easily illustrated.

Consider figure 1.1, which shows a linear demand curve for, say, off-peak electricity. Prior to the implementation of TOU rates, assume the customer paid a flat-rate $\bar{p}$ per kWh and consumed $\bar{x}_1$ kilowatt-hours per month. At the new price, p_1, the customer now consumes quantity x_1. He pays a price p_1 for every unit consumed. But according to the demand schedule, he would be willing to pay a slightly higher price for consumption slightly less than x_1, a slightly higher price still for consumption slightly less than that, and so on back to the initial consumption level $\bar{x}_1$. But he receives these marginal units of consumption at price p_1 instead of at price $\bar{p}$. Thus, a "surplus" value accrues to him from the price change which is represented by the area of the shaded trapezoid $\bar{p}p_1ba$ (the amount G + U). A discriminating monopolist could price his product such a way that each additional unit (moving from $\bar{x}_1$ to x_1) is just that much less expensive so as to extract this surplus value from the consumer.

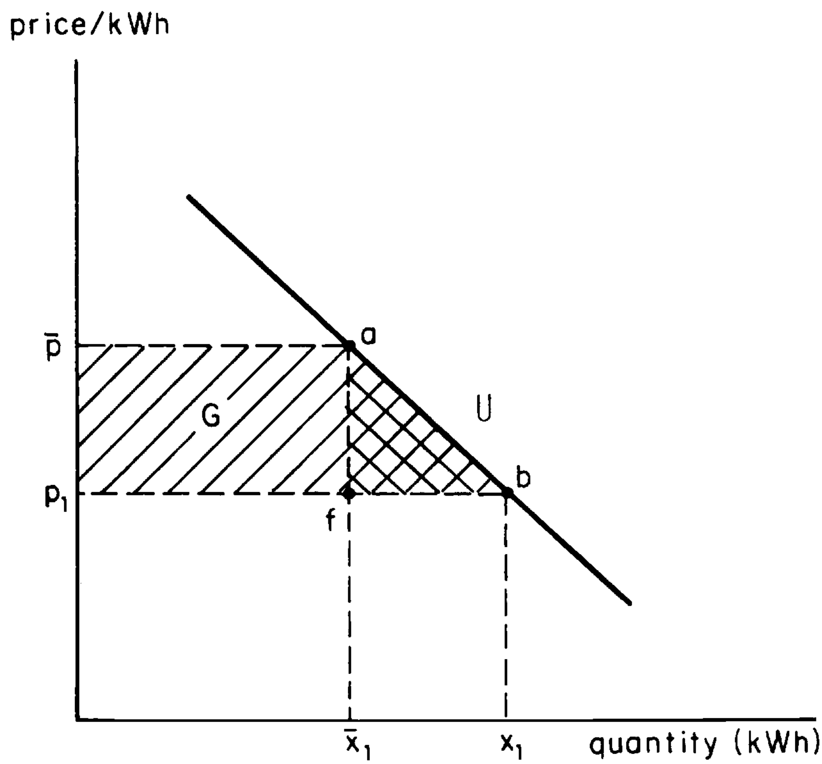

Fig. 1.1 Measurement of change in consumer's surplus from a price decrease from $\bar{p}$ to p_1

Otherwise the shaded area, which is in dollar units, represents a real gain in the customer's "welfare."

In a similar fashion, the increase in peak-period kWh's will result in a loss of this same sort of welfare. Figure 1.2 illustrates the situation, wherein the peak-period price per kWh is increased from a flat rate of $\bar{p}$ to p_2. At this price the customer will consume x_2 kWh's as compared to $\bar{x}_2$ when the price was $\bar{p}$. Using the same heuristic argument as before, the change in the consumer's surplus is a loss, in the amount of the area of the shaded trapezoid $p\bar{p}_2cd$ (the amount L).

Adopting a conventional notation, these areas are, respectively,

$$\Delta p_2\left(\frac{x_2+\bar{x}_2}{2}\right) \text{ and } -\Delta p_1\left(\frac{x_1+\bar{x}_1}{2}\right),$$

where the Greek delta means "change in" or "difference," i.e., $\Delta p_1 = p_1 - \bar{p}$ and $\Delta p_2 = p_2 - \bar{p}$. Then the change in the consumer's surplus from

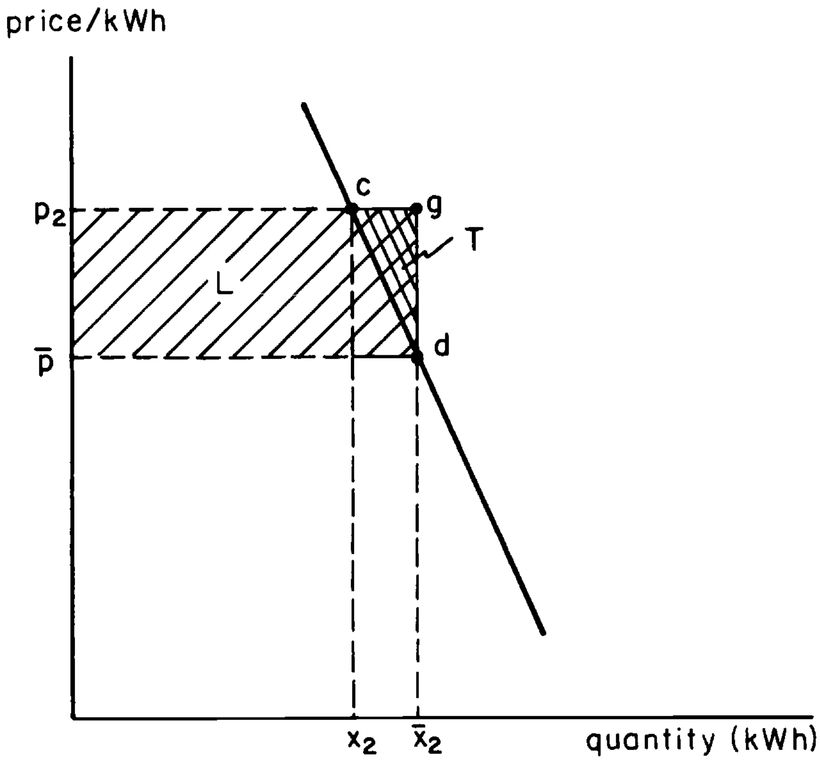

Fig. 1.2 Measurement of change in consumer's surplus from a price increase from $\bar{p}$ to p_2

the combined effects of adopting a higher peak-period price and a lower off-peak-period is given by:

$$(1) \qquad \Delta CS = -\left[\Delta p_1\left(\frac{x_1+\bar{x}_1}{2}\right) + \Delta p_2\left(\frac{x_2+\bar{x}_2}{2}\right)\right].$$

From the producer's viewpoint, calculation of the gains and losses from moving to TOU pricing is simplified by assuming that $\bar{p}$, the flat rate charged in each period, is a quantity-weighted average of the marginal costs of production in each time period and that the customer charge recovers all fixed costs. If mc_1 and mc_2 are the marginal costs in the off-peak and peak periods, respectively, and ϕ_1 and ϕ_2 are the proportions of off-peak and peak consumption relative to total consumption for the class, then

$$(2) \qquad \bar{p} = \phi_1 mc_1 + \phi_2 mc_2\,.$$

Assuming that the TOU prices, p_1 and p_2 are set equal to their respective marginal costs, mc_1 and mc_2, and again that the customer charge (E) is used to recover all fixed costs, then in figure 1.1 the producer will lose revenue by reducing the off-peak price (demand is assumed to be inelastic) in the amount $p_1x_1 - \bar{p}\,\bar{x}_1$ for a change in production costs of $p_1(x_1 - \bar{x}_1)$. This gives a net change (loss) of

$$-\bar{p}\bar{x}_1 + p_1\bar{x}_1 = \Delta p_1\bar{x}_1,$$

a negative quantity. Similarly, in the peak period revenues are increased by changing from $\bar{p}$ to p_2 (again, demand is assumed to be inelastic), and costs are reduced, giving a net effect (gain) of $\Delta p_2\bar{x}_2$. In sum, the change in net revenue to the producer is

$$(3) \qquad \Delta NR = \Delta p_1\bar{x}_1 + \Delta p_2\bar{x}_2\,.$$

If we adopt as our criterion for judging the overall benefits or costs to society of the adoption of the TOU pricing scheme (E, p_1, p_2) the sum $\Delta CS + \Delta NR$, then the change in welfare becomes

$$(4) \qquad \begin{aligned} \Delta W &= \Delta CS + \Delta NR \\ &= -\left[\Delta p_1\left(\frac{x_1+\bar{x}_1}{2}\right) + \Delta p_2\left(\frac{x_2+\bar{x}_2}{2}\right)\right] \\ &\quad + \Delta p_1\bar{x}_1 + \Delta p_2\bar{x}_2 \\ &= -\tfrac{1}{2}\,(\Delta p_1\Delta x_1 + \Delta p_2\Delta x_2)\,. \end{aligned}$$

This sum is equivalent to the sum of the areas U and T in the figures.

This welfare measure is only approximate. First, the demand curves used to develop these formulas are linear. Secondly, the contemplated price changes not only have direct influence on the quantities consumed that are depicted in the figures but also have indirect effects due to the interrelationship of each quantity on *both* prices. An exact measurement for ΔW depends on more adequate measurement of ΔCS and requires knowledge of this set of two interrelated demand equations.[13] This is what the TOU pricing experiments can provide.

For a representative customer, one who at the initial consumption levels $\bar{x}_1$ and $\bar{x}_2$ consumes electricity in the two periods exactly in the proportions ϕ_1 and ϕ_2, the equation (4) also can be interpreted as showing a direct welfare effect and a bill effect. At the initial consumption levels, x_1 and x_2, a movement to prices p_1 and p_2 will create a change in a customer's bill of exactly

$$(5) \qquad \Delta B_0 = \Delta p_1 \bar{x}_1 + \Delta p_2 \bar{x}_2$$

Whether this number is positive or negative depends precisely on how

$$\bar{x}_1/(\bar{x}_1 + \bar{x}_2) \text{ and } \bar{x}_2/(\bar{x}_1 + \bar{x}_2)$$

compare to ϕ_1 and ϕ_2. If they are respectively equal, which is the definition of the representative customer, then $\Delta B_0 = 0$. The shifts in consumption that are observed, Δx_1 and Δx_2, therefore represent unambiguous welfare improvements. Ultimately, the bill changes by an amount

$$(6) \qquad \Delta B = \Delta B_0 + p_1 \Delta x_1 + p_2 \Delta x_2$$

(in which $\Delta B_0 = 0$ for the representative customer).

Tracing through the changes in the bill and in consumer's surplus from the price changes to p_1 and p_2 for the representative customer, we are lead to the same formula for ΔW derived previously in equation (4). In figure 1.1 the representative customer's bill increases by $x_1 \bar{x}_1 bg$, whereas his surplus has increased by the trapezoidal area G + U. The net effect is a gain in welfare, the area U. Similarly, in the peak period his bill falls by the amount $x_2 \bar{x}_2 gc$, but surplus is reduced by the area L. The net effect is a gain, the area T.

For a nonrepresentative customer, one with consumption proportions different from ϕ_1 and ϕ_2, the calculations are the same, but $B_0 \neq 0$. Thus, there will be a shift in revenue either to the customer or to the utility at initial consumption levels in addition to the changes summarized in equation (4). As mentioned previously, calculating the exact welfare changes depends on a knowledge of the full demand structure; the above calculation should be viewed as approximate only.

The entire objective of this exposition lies in the identification of those customers for whom ΔW exceeds the additional costs associated with implementing the TOU rate structure. This desired set of customers may be identified by size, by appliance ownership, or by some other meaning-

13. See Acton and Mitchell (1980, 7) for more discussion.

ful stratifying variable. In their application to the LADWP service area, Acton and Mitchell found that under an illustrative mandatory program, the change in welfare was sufficiently large to offset the traditional metering costs only for customers whose average monthly consumption exceeds 1100 kWh. Although this is a small fraction of the population of LADWP residential customers (4.3 percent), they account for approximately 17 percent of residential consumption.

Offering TOU pricing on a voluntary basis creates a further difficulty in that the welfare-efficient set of customers may not be the ones attracted to the program. For example, in the Acton and Mitchell illustration, if customers were to be charged the appropriate monthly rate for TOU metering, the monthly bills of all but the very largest members of the $\geqslant$ 1100 kWh/mo. group would experience bill increases at initial consumption levels. Unless customers could be persuaded that their ultimate shifting response will result in sufficiently more "welfare" to make participation worthwhile—no doubt a hard concept to sell—the voluntary program will not work.

However, there may be other population subgroups that are almost as attractive in the welfare-efficiency sense whose bills at initial consumption levels show decreases large enough to offset metering charges, thereby making participation in a voluntary program attractive. In the LADWP case, Acton and Mitchell point to those households with swimming pools as likely candidates, but in this case also, bill changes at initial consumption levels do not make a voluntary program attractive, even though bill decreases would be experienced by most of this group once their peak and off-peak period consumption had adjusted to the TOU rates. Many households with electric space heating, on the other hand, would enjoy initial bill decreases in excess of additional metering costs and would, therefore, find a voluntary program attractive. However, only for the largest of these users are the welfare gains sufficient to offset metering costs. Therefore, the utility faces the problem of attracting a potentially large number of customers into a voluntary program for whom the ultimate composite welfare benefits are such that they should not have been included.

1.4 The Risks of Implementation

The previous sections have laid a foundation for possible implementation of TOU rates. But there are various uncertainties to cope with in assessing the risks associated with that course of action, given that the elasticity estimates upon which it depends are subject to sampling error at the very least (assuming there was a relevant experiment to provide elasticity estimates for the target population) and, more generally, subject to other risks, under the assumption that transference of results outside the target population and/or service territory is involved.

An important aspect of the risks involved in an implementation decision is that even if there is a relevant experiment available that can provide elasticity estimates for the target population of interest, the sampling errors associated with those estimates imply uncertainties for the revenue and welfare calculations. These can and should be traced through, by deriving the statistical confidence intervals on, say, revenue from the confidence intervals available for the estimated elasticities.

Finally, the more widespread problem of transferring results from an experiment of relevance to a different target population and, probably, a different service territory involves yet addititional uncertainty, still of the statistical variety. The methodology for handling the transference problem is available but has not yet been adapted to the issue at hand.

So, while there can be general discussion on all these points, the particular use of available techniques for applying them in specific cases remains one of custom fit.

1.4.1 Uncertainty in Using Experimental Results

In those few service areas where a reasonably good TOU experiment has been run, there still are some sticky issues to contend with in using the demand elasticities or, more generally, the estimated set of demand equations in evaluating the welfare and revenue impacts of adopting a TOU pricing structure. These issues are developed and illustrated herein. In those service areas where an experiment without induced price variation was completed, a similar exercise could be accomplished, but only if the contemplated rate structure were exactly the same as the experimental rate structure. Otherwise the situation is such that an evaluation would require the transference of results from another service area, a topic we address in the second part of this section.

A crucial feature of any of the experiments is the effective population to which they apply. In their initial evaluation of the available TOU-pricing projects, Research Triangle Institute (RTI) does a good job of addressing this point (U.S. Department of Energy 1978), and they reemphasize it in a recent report summarizing their findings on estimated price effects for several of the projects.

> The sample was drawn from a population consisting of all 1976 WPSC non-farm non-seasonal residential customers who had a 12-month billing history and whose average monthly usage exceeded 100 kWh in 1975. This sampled population (approximately 217,000) was stratified into nine groups according to their 1975 annual consumption. Simple random sampling was used within each stratum to produce the desired number of participants. Approximatley 43 percent of all WPSC residential accounts were represented by the sample.

> Participation in the study was mandatory. Customers were billed under revenue-neutralized rates. The rates ensured that the average household's bill would be nearly identical to its bill under the existing rates, if its consumption level and pattern were unchanged during the experimental period. (Miedema and White, 1980, 66–67)

According to this quotation, 43 percent of Wisconsin Public Corporation accounts are represented by the sample; that is, this 43 percent is the effective population for which the sample results can be generalized. Since participation was mandatory, there is no apparent problem of sample self-selection to contend with.

In Ohio, the experiment was voluntary, and, according to Miedema and White,

> Selection of sample customers excluded those without a 12-month billing history of April 1974 to April 1975, bulk-metered residences, company employees, residents on "frozen" rates, and customers with less than 6,000 kWh annual usage. The group that was sampled contained approximately 196,000 of the 337,000 residential customers living in the DPL service area.
>
> About 80 percent of the customers selected to participate in the study were not included for such reasons as meter installation problems, refusal to participate, and moving plans. Based on the combined number of rejections for the experimental and control groups, customers in the final sample represent about 39,000 residential customers (the effective population). This group constitutes approximately 20 percent of the sample population and 12 percent of all residential customers. (1980, 55–56)

In this instance, many restrictions were imposed on the population ultimately sampled; so many, in fact, that the results (which are only comparative—one TOU rate was used) apply at best to 12 percent of all residential customers.

The point is simply that statistical inference involves generalizing sample results in the population from which the sample was selected. If that population is different from the target population, then the inference will be biased. In some situations the problem can be handled analytically,[14] but otherwise the best that can be hoped for is knowledge of the direction of the bias.

For the effective population, it is relatively simple to trace the effects of imprecision in the estimated demand elasticities on the revenue or welfare calculations discussed previously. For example, the revenue change derived from shifts in off-peak consumption in moving from price $\bar{p}$ to TOU prices p_1 and p_2 is a function of the quantity change $\Delta x_1 = \bar{x}_1 - x_1$.

14. See, for example, one of the attempts to correct the Arizona experimental results for bias introduced by the incentive scheme used (Aigner and Hausman 1980). While these authors do not specifically address the target-population–effective-population issue, the methodology developed is pertinent.

And that quantity change, as a function of the price changes p_1 and p_2, can be written

$$\Delta x_1 = \frac{\partial x_1}{\partial p_1}\Delta p_1 + \frac{\partial x_1}{\partial p_2}\Delta p_2 . \tag{7}$$

The required partial derivatives, if not directly estimated from a set of fitted demand functions, are usually simple functions of other estimated model parameters, and a measure of the precision of their estimation can be readily obtained. For example, given direct estimates of $\partial x_1 / \partial p_1$ and $\partial x_1 / \partial p_2$ with their estimated variances and covariance, the variance of Δx_1 can easily be derived and used to put a confidence band around the predicted revenue or welfare change.[15] This would seem to be a very useful way to characterize the implications of uncertainty (imprecision) in estimated model parameters (slopes, elasticities) on the ultimate criterion for judging whether a particular TOU rate program should be adopted or not.

Unfortunately, there is not just one way to estimate the important parameters. In a previous section we saw how model specification had a pronounced influence on the estimated elasticities in the Arizona experiment. Each set of parameter estimates rests on different statistical assumptions, and therefore so do the inferences emanating from them. And, for the most part, these underlying assumptions cannot be tested. What is hoped for is robustness in the estimated parameters over models, but this almost never seems to happen. The range of parameter estimates over models, which implies a range of revenue and/or welfare changes, poses yet another source of uncertainty to cope with in evaluating a proposed TOU rate. In all such situations the most reasonable thing to do is to trace out the range of implications and assess their consequences in order to fully understand the risks involved, if not to completely resolve the matter.

1.4.2 The Transferability Problem[16]

When sampling is used to answer a question—test a hypothesis or estimate a particular population parameter—there is always the possibil-

15. If $V(b_1)$ and $V(b_2)$ are the variances of the estimates for $\partial x_1/\partial p_1$ and $\partial x_1/\partial p_2$, respectively, and $C(b_1,b_2)$ is their covariance, then

$$V(\Delta x_1) = \Delta p_1^2 V(b_1) + \Delta p_2^2 V(b_2) + 2\Delta p_1 \Delta p_2\, C(b_1,b_2) .$$

Extending to get a measure of precision for the revenue change from alterations in off-peak consumption, we have that

$$\Delta R_1 = p_1 x_1 - \bar{p}\bar{x}_1 = p_1\, \Delta x_1 + \Delta p_1 x_1 ,$$

and that

$$V(\Delta R_1) = p_1^2 V(\Delta x_1) .$$

16. In preparing this subsection I have benefited greatly from discussions with E. Leamer.

ity of using previous studies or experience on the same topic in order to "guide" the sample design, to supplement it, or to replace it completely by prior information. If it is assumed that in any such application there is a primary population model that describes the relationships between dependent and independent variables and the statistical properties of error terms, and that this population relationship applies, with differences in parameters only, to all subpopulations, then all available information can be represented in terms of sets of estimated parameters of a "grand" model. The analytical question to be addressed is how the various sets of estimates are to be weighted in importance.

To be specific, consider the situation whereby a utility is considering increasing the sample size of its residential load research sample. The choices are limited to the addition of some calculable number of magnetic tape-recording meters and/or dependence on the load research data of a neighboring utility or group of utilities to use in lieu of an addition to its own sample. Assuming there is a statistical model that relates the load in any predescribed small interval (the meter resolution capability) to weather and various other exogenous factors (physical characteristics of homes, socioeconomic characteristics of families, appliance ownership patterns) for both utilities, it can be shown that there is an optimal estimator for the load at any moment which is a linear combination of the load data from both utilities.[17] This is not to say that additional sampling is not necessary; indeed, it may be. But at least a methodology exists that makes it possible to combine information in an appropriate fashion. Likewise, in any decision problem prior information can be considered in combination with sample data to decide whether further sampling is required, and, if so, how many additional additional observations should be taken.[18]

The transferability problem with respect to an analysis of the implications of a TOU rate can be viewed in exactly the same way. In general, the problem should be approached by allowing for the possibility that a fresh sample will be taken, that is, that the subject utility will design and operate its own pricing experiment. Then, given prior information based on the results of other experiments, the question is whether a new sample should be taken, and if so, what its size should be. Otherwise the decision problem is based solely on the available prior information. In a similar fashion, the combination of existing elasticity estimates for utilities that have run experiments with information from other utilities can be achieved.

A major problem in combining data sets or estimates concerns data

17. Lindley and Smith (1972). A very recent application of this methodology to a "transferability" problem is contained in the paper by DeMouchel and Harris (1981), which I became aware of at the conference.

18. Aigner (1979); a general reference is Raiffa and Schlaifer (1961).

quality and sample design. While it is possible to combine data from different utilities according to the Lindley and Smith (1972) methodology, a presumption must be made about their quality. Moreover, while the data themselves may be comparable, no doubt there will be differences in design—as summarized by which stratifying variables have been used, on which other explanatory variables data have been collected, etc. Thus in the combining process, recognition must be made of the fact that various specification errors will exist in fitting the so-called grand model. This problem has not been addressed so far in the literature that has grown out of the Lindley and Smith work. However, it must be confronted in order to properly characterize the data that are to be combined and to thereby determine whatever constraints may exist on the combining process from data that lack certain essential characteristics or "quality."

A general proposition is that "similar" utilities can properly make use of each other's load data or experimental results. The statistical modeling problem is how best to define "similar" and how to "partially transfer" information among dissimilar utilities. As an example, initially one might suppose that data could be transferred among warm-climate utilities and among cold-climate utilities but not between utilities drawn from different climates. However, climate might affect the average load in a given time period while leaving unchanged the responsiveness of load to increases in price. In that case information about the responsiveness of loads to price, possibly for predicting load increases, could be transferred between warm- and cold-climate utilities, but information about the average loads could not. Next suppose that after a study of several utilities it was found that the average load increased by 1 percent every time the average daily temperature departed from 68° by 2°. Then even the information about average load could be transferred between climates. This means that the transferability problem has to be addressed within the context of an econometric model of load demand.

As an example data-pooling problem, consider the one studied by Efron and Morris (1975), from which table 1.4 is taken. The batting average of eighteen major-league players after 45 turns at bat during 1970 are reported in column 1 of table 1.4. Anyone familiar with baseball statistics will recognize immediately that these averages are too dispersed to be end-of-season averages. In fact no one has batted over .400 for the season since Ted Williams did in 1941. A sensible prediction of Clemente's rest-of-season average would surely be less than his average of .400 after only 45 at bats. What Efron and Morris do is to pool Clemente's average with the averages of the other seventeen batters. An extreme possibility would be to assume that all players have the same batting ability and to predict the average for the remainder of the season to be the same for each, thus equal to the overall average of the 45×18 at

Table 1.4 **1970 Batting Averages and Predictions for Eighteen Major League Players**

Player	Average after First Forty-five At Bats	Average for Remainder of Season	Pooled Prediction Based on Stein's Estimator	Limited Translation Pooled Prediction
Clemente (Pitts, NL)	.400	.346	.290	.351
F. Robinson (Balt, AL)	.378	.298	.286	.329
F. Howard (Wash, AL)	.356	.276	.281	.308
Johnstone, (Cal, AL)	.333	.222	.277	.287
Berry, (Chi, AL)	.311	.273	.273	.273
Spencer (Cal, AL)	.311	.273	.273	.273
Kessinger (Chi, NL)	.289	.263	.268	.268
L. Alvarado (Bos, AL)	.267	.210	.264	.264
Santo (Chi, NL)	.244	.269	.259	.259
Swoboda (NY, NL)	.244	.230	.259	.259
Unser (Wash, AL)	.222	.264	.254	.254
Williams (Chi, AL)	.222	.256	.254	.254
Scott (Bos, AL)	.222	.303	.254	.254
Petrocelli (Bos, AL)	.222	.264	.254	.254
E. Rodriguez (KC, AL)	.222	.226	.254	.254
Campaneris (Oak, AL)	.200	.285	.249	.242
Munson, (NY, AL)	.178	.316	.244	.218
Alvis (Mil, NL)	.156	.200	.239	.194

bats, namely .265. But the dispersion of the 18 averages after 45 at bats is much greater than would be predicted if all players had identical abilities. A better assumption would be that the players' abilities come from a distribution with unknown mean and unknown variance. This assumption leads to the partial pooling of the 18 averages. The prediction of the batting average for the remainder of the season is then a weighted average of own average and overall average (.265). Clement's estimated average is thereby reduced from .400 to .290. The last two columns of table 1.4 contain different pooled predictions, the latter one being a "limited translation estimator" that was designed to limit the pooling effect on extreme averages, Clemente's for example. By a measure of overall error, the pooled estimators perform better than the players' own scores after 45 at bats as predictions of their future performance.

The statistical problem of predicting the rest-of-season averages of these eighteen ball players is essentially the same as the problem of data transferability for utilities. If no data are transferred from the other seventeen players to Clemente, a prediction based on his average alone is likely to be too extreme. If the other data are fully transferred with no allowances made for individual differences, the resultant prediction of

.265 is likely to be too conservative. The statistical technique described by Efron and Morris allows the data to select the degree of transferability. If the 18 averages are sufficiently similar to support the hypothesis that all batters have the same ability, then the data are fully transferred in the sense that Alvis's .156 average has the same weight for predicting Clemente's rest-of-season performance as Clemente's own average. But if the 18 averages after 45 at bats are very dispersed, the degree of transferability can drop to zero, with Clemente's prediction being equal to his current average of .400.

This simple problem of pooling batting averages can be generalized in several ways that bring it closer to the complexities that must be handled in the problem of data transferability. The batting averages used for prediction purposes are all based on 45 at bats. In the case of the TOU experiments, sample sizes vary widely. And for most utilities that desire to evaluate a TOU pricing structure, sample size is zero. This is akin to a new player entering the league.

Of course, in the case of utility companies there will be many variables which could suggest reasons why a utility without data should not be treated as an average utility. Continuing the baseball example, we might expect to observe two kinds of batters—outfielders and infielders, with the latter selected with greater concern for their defensive abilities. A player with an initial average of .400 may with justification be considered an outfielder, and therefore his average should be shrunk not toward the overall mean of .265 but rather to the mean of the outfielders' averages. If it is not known which batters are offensive specialists and which defensive specialists, the data may nonetheless suggest the dichotomy.

Theoretically, the limitations on data transference will be defined by the weight placed on data from other utilities in the pooling process just described. In a sense, if all relevant variables are observed both in the parent utility and in the collection of utilities at large, there are no limitations to transferability. But because the data on concomitant variables may not be comparable, certain difficulties arise in the pooling process. As mentioned previously, the nature and extent of these difficulties remain to be worked out. From this research will come guidance as to what the limitations to transferability are. They will be defined in terms of geographical location, specific weather patterns, customer demographics, etc.

To conclude, there are no *conceptual* limitations on data transference but there are strong assumptions which may not be very attractive that are required to accomplish it through the Lindley and Smith (1972) framework. The existing methodology has not yet been applied to the case of transferring experimental results between utilities or, for that matter, to load research data in general. This major research focus remains to be undertaken. Finally, within the context of a general framework for transferability, even the information from experiments without price variation may be important and usable.

1.5 Conclusions

What are the primary shortcomings of the available estimates of the relevant short-run elasticities? First, there is the matter of discrepancies among the estimates within experiments. From a qualitative viewpoint, we can conclude that TOU- pricing "worked." Quantitatively, only the Arizona and Wisconsin tests provide enough integrity of design and at least a semblance of common findings on which to base a decision about implementation. Soon, but not just yet, detailed results will be available from Oklahoma, North Carolina (CP&L), LADWP, and Southern California Edison Co. These results may provide enough evidence for implementation, but we should also like to know if a better strategy than overall implementation might be phased or selective implementation, concentrating on large customers first. Presumably every experiment's results could be organized and reported according to size of customer (at least in some broad intervals) as in the LADWP case. Thus proper information to support selective implementation could be made available.

What then of the fact that the available reliable findings are confined to such a limited area (Arizona and Wisconsin—or Arizona, Wisconsin, Oklahoma, Southern California [LADWP and SCE], and North Carolina, when these latter results are available)? Ideally, one would hope that once all conditioning variables have been accounted for, there will be a unanimous verdict, even down to the actual numbers themselves. Yet there seems to be enough room for a lack of unanimity that, due to the basic latitude for mismodeling, present but undiscovered faults in the experiments, etc., one shouldn't be too surprised if in fact there is a hung jury on the specific magnitude of TOU effects across service territories.

Research on the transferability problem, made even more important by practical considerations imposed by the National Energy Act, may produce the comprehensive framework within which all such issues can be resolved. But the electric utilities will have to act (or at least defend their nonaction) on TOU- pricing very soon, probably before the transferability problem is resolved. In the interim, given a clear verdict on the qualitative results of the experiments, implementation may take place, for the rate setting and regulatory process itself is an experiment with inherent and accepted risks associated with errors of calculation and judgment. A decision made on TOU- pricing is no more or less irreversible than any other.[19]

19. Following this line of reasoning, presumably there are now quite a few examples of implementation, primarily of a voluntary nature, including Arizona Public Service Co., LADWP, and Wisconsin Power & Light.

Finally, what of the design issues raised at the outset? How have they contributed to or limited our knowledge of TOU- pricing effects from the battery of available experiments? Clearly the lack of significant price variation and the fact that in most experiments with price variation, peak, off-peak, and midpeak prices are not varied independently substantially limits the generalizability of the results even on a local level. Sample-size considerations are at issue here too, since more "treatments" (prices, definitions of pricing periods, etc.) require more observations in order to achieve a given level of precision for parameter estimates.

In these several respects, one of the experiments stands clearly above the rest: the Rand-LADWP experiment. It has a large sample size with wide (and, to some extent, independent) variation in prices. Definitions of pricing periods (peaks, off-peak, midpeak) were not rigid. Moreover, optimal design methods were used to further squeeze information out of the sample, but not to such an extent as to limit the variety of analysis models that could be applied once the data were in hand. Probably the only clear limitation on the LADWP experiment was its use of volunteers and an incentive scheme for participants.[20]

All the other experiments represent some compromise on the best features of the LADWP test, but some also improve on the LADWP experiment's major weaknesses. For example, both the Wisconsin and North Carolina (CP&L) projects have mandatory participation and no compensation payments.

In all cases, however, there are other, perhaps more severe limitations that customers face, imposed by the experimental environment and the fact that the experiments are short-run by nature. Many people contend that with a full-blown commitment to TOU- pricing, the responsiveness of customers will be different than that estimated by the experiments. This difference will occur because appliance choices will be made with an eye to TOU response; new appliances will become widely available, and there will be no choice but for customers to respond (assuming mandatory implementation). But it seems impossible to estimate these effects. The prevailing attitude is that if TOU- pricing can be shown to be cost-effective based on the available experimental evidence, the pricing strategy surely must be even more desirable in the long- run.

While both of these views are reasonable, it is important to note that none of the experiments allows us to estimate the effects (apart from TOU response) of changing the average price of electricity on the consumption of other goods. These expenditure elasticities, which convert partial price elasticities to full price elasticities, can mitigate substantially

20. It should be noted that the issues of independent price variation, voluntary participation, and incentive payments are not separate. Rates designed with revenue neutrality in mind (which therefore eliminates independent variation in peak and off-peak prices), of course eliminates the need for compensation payments and may be an important ingredient in achieving mandatory participation.

the anticipated long-run effects from mandatory implementation of TOU rates. It may well be that the apparent overall conservation effect of TOU- pricing observed in many of the experiments is merely an artifact of their short-run nature, wherein response was constrained by the present set of household appliances and closely reflects the potential response to higher electricity prices without regard to TOU influences.

Appendix: Project Evaluations

In this appendix we discuss the fifteen demonstration projects in more detail and explain our evaluation of each project. The projects may be judged on many facets of their experimental design and, as mentioned in the text, the Research Triangle Institute and Michigan Survey Research Center (SRC) studies do just that. The SRC paper, for example, evaluates projects on the basis of appropriateness of experimental conditions, noncoverage error, sampling error, treatment error, nonresponse error, and measurement error. The judgments we make in table 1.1 of the text are based on the descriptions and evaluations given in the SRC study, with the design of the experiment's price treatments being of primary importance in making our decisions.

In classifying projects on the basis of their ability to provide quality data for estimates of TOU price effects, we use the following procedure. First we divide the projects into two groups, depending on whether they employ a single price treatment or a wide variation in treatments. We then examine the nonprice aspects of each project's experimental design. Those projects with a single price treatment which also have other design problems are classified as being "of no interest." Those with a single price treatment which are otherwise generally well designed are termed "of limited interest." Any projects that have multiple price treatments but other design drawbacks are also considered to be "of limited interest." Finally, projects with multiple price treatments that are well designed overall are viewed as being "of interest."

Projects with a Single Price Treatment and Other Design Flaws ("Of No Interest")

Projects in this category are those in New York, Ohio, Rhode Island, and Vermont. All employ single price treatments. The Vermont experiment is not even evaluated by the SRC study because of its unscientific design. The New York project has a severe problem with a biased sample, as the sample is made up in large part by nonrandomly selected volunteers who responded to a bill insert. The Ohio project is somewhat better in this regard, since the sampling was conducted randomly (after some customers were excluded for various reasons). However, because participation was voluntary, about 70 to 80 percent of the selected customers decided not to take part in the experiment. Other problems include possible customer confusion about the price they were being charged during a labor strike and insufficient documentation of many of the details of the experiment. The Rhode Island project appears not to have been designed to examine demand responses to TOU rates, as it employs a complex rate structure that is never varied. At the time of the SRC study, documentation was insufficient to judge the project on many other design characteristics.

Projects with a Single Price Treatment Which Are Otherwise Well Designed ("Of Limited Interest")

The Arkansas, Connecticut, and North Carolina (BREMC) projects come under this heading, as all are generally well designed except that they have only a single price treatment. The Arkansas project's favorable design characteristics include mandatory participation and no participation incentives. The Connecticut experiment employs a sampling procedure that results in high-quality consumption and survey data.

Projects with Multiple Price Treatments but Other Design Flaws ("Of No Interest")

The Puerto Rico project falls in this category. While it features varying time-of-day rates that allow for elasticity estimation, it has a number of unappealing aspects. There are many unnecessary exclusions from the sample, it is unclear whether allocation to treatment groups is random, and response is voluntary and quite low.

Well-Designed Projects with Multiple Price Treatments ("Of Interest")

The projects in this group are those in Arizona, California (LADWP and SCE), North Carolina (CP&L), Oklahoma, and Wisconsin. All employ multiple price treatments and generally have other favorable design characteristics. The North Carolina and Wisconsin projects are the best designed of these featuring widely varying rates since they have mandatory participation. These two experiments also use rate structures that contain demand charges, enabling researchers to estimate price effects on peak kW demand. The other experiments offer slightly lower quality data, and researchers have to contend with problems of volunteer bias and the effects of incentive payments on consumption.

Other Projects

TOU pricing experiments were also scheduled to take place in New Jersey and San Diego, California. Both were scrapped because of equipment problems.

Comment Paul L. Joskow

Dennis Aigner has provided us with a useful paper that reviews the various residential time-of-use (TOU) pricing experiments and discusses a variety of important implementation issues. After nearly a decade of debate before state and federal regulatory agencies about the benefits and costs of TOU pricing and the expenditure of tens of millions of dollars on experiments, it is certainly an appropriate time to take stock of what we have learned.

The Aigner paper leads me to several general conclusions about the TOU experiments and their usefulness for public policy and electric-utility decision making. First, it is fairly clear that many of the experiments have serious flaws that limit their usefulness for estimating the price elasticities of interest. While there are numerous sampling, experimental design, and duration problems, the most important general problem is that many of the experiments failed to provide a sufficient number of different TOU price treatments to make econometric estimation of price elasticities possible. Of the fifteen experiments discussed, only six appear to be particularly useful for the estimation of TOU own-price and cross-price elasticities, and data from only two of these (Arizona and Wisconsin) have been made widely available for analysis to date.

Second, most studies that have analyzed the data from experiments which were structured so that price elasticities could be estimated have found that the own-price elasticities at various times (peak, shoulder, off-peak) are negative. This result comes as no great shock to most economists. The point estimates for the own-price elasticities vary widely, however. Peak-period elasticity estimates have a range of something like -0.2 to -0.8, and off-peak elasticity estimates range from about -0.1 to -0.8. Differences among the experiments in the definition of pricing periods makes useful comparisons of peak versus off-peak elasticities very difficult. Cross-price elasticity estimates also vary widely across studies, both in absolute value and sign. Analysts working with similar data come up with very different elasticity estimates because the analytic approaches differ with regard to the specification of the demand system to be estimated, the experimental time period for which data are utilized, and the extent to which they account for the compensation scheme used to "protect" participants from increases in their electricity bills.

Finally, at least in principle, the results of these experiments can be used by public utilities and their regulatory commissions to help decide on whether to institute voluntary or mandatory residential TOU pricing.

Paul L. Joskow is professor of economics, Massachusetts Institute of Technology.

There are three major issues that must be tackled here. First, since the implementation of residential TOU pricing requires relatively large expenditures on metering equipment, we want to determine the characteristics of those residential customers whose responses to TOU pricing will yield welfare gains that are greater than the additional transaction costs associated with TOU pricing. Aigner's discussion of the calculation of the welfare gains and losses from TOU pricing, that draws on the familiar Steiner/Boiteux peak-load pricing model and the application of this model to preliminary data from the Los Angeles experimental data by Acton and Mitchell, provides a simple methodology for making such calculations. Second, even when good experimental data are available to a regulatory commission, the elasticity estimates are uncertain, and statistical confidence intervals must be correctly calculated and applied in a meaningful way in any welfare analysis. Third, techniques must be developed to allow us to transfer what we have learned from the very small number of good experiments to other areas of the country with diverse economic and demographic characteristics. While Aigner sketches out possible solutions to these last two issues, it seems to me that a lot of work remains to be done before they can be adequately resolved.

The residential TOU experiments were conducted primarily at the behest of state and federal regulatory authorities and electric utilities interested in obtaining information to help resolve public policy debates over TOU pricing and to estimate the effects of TOU pricing on load patterns and revenues. These are the ultimate customers for the results of these experiments. I believe it is useful to evaluate them from the perspective of the regulators and the firms in the context of the problems they have been trying to grapple with at least for the past ten years.

The first issue that I want to discuss is why so many of the experiments were so poorly designed. I believe the answer lies in the true political and economic origins of these experiments. When the earliest experiments were structured, those involved had simply not thought very deeply about what the data generated might be used for. The earliest experiments were motivated more by narrow adversarial and litigation concerns than by a serious interest in sound economic analysis. The earliest efforts to estimate the effects of TOU pricing on residential electricity consumption patterns arose in the early 1970s in the context of enormous regulatory controversy about the desirability of marginal-cost pricing in general and TOU pricing based on marginalist principles in particular.

On the one hand, several environmental groups and a number of economists appearing before state regulatory commissions were pressing for the use of marginal-cost pricing principles in place of conventional average-cost pricing principles and argued that mandatory TOU rates for industrial customers and mandatory or voluntary TOU rates for residential consumers should be implemented. Proponents of TOU pricing pointed to the experiences in France, England, and Germany as examples of situations in which TOU pricing had been used successfully at both the industrial and residential levels. The efficiency rationale for marginal-cost pricing and the extension of marginalist principles to the development of peak-load pricing schemes for electricity service had long been a part of the economic literature. The major interest of environmental groups in TOU pricing was to give incentives to reduce the rate of growth in peak demand so as to reduce the need for additional power plants. These arguments were often congenial to state regulatory commissions because they too were looking for ways to reduce the need for new generating capacity, with costs two to three times average historical costs, as a way to moderate the need for rate increases. Most electric utilities opposed TOU pricing initially. Among other things, they argued that consumers would not respond to higher prices by reducing consumption on peak, that TOU metering was too costly and impractical, that it would increase uncertainty about revenues and profits, that it would only lead to shifting peaks, etc. Large industrial customers opposed marginal-cost pricing and, initially, TOU pricing primarily because they were concerned that major changes in rate structures would be used by state regulatory commissions to redistribute the relative contributions to total utility revenue requirements so that the industrial classes would pay more and the residential class pay less. The early TOU experiments were really fire-fighting exercises aimed at developing some crude U.S. evidence that the elasticity of demand for electricity was negative (yes indeed, back in the early 1970s some regulatory commissions and utilities refused to believe this). Furthermore, the experiments were motivated by a misperception that the residential class represented the greatest target of opportunity for TOU rates, despite the fact that countries like France had devoted most of their efforts to the industrial class. The early TOU experiments were also viewed by some as a convenient way of delaying regulatory decisions on TOU pricing.

These experiments had their origin in a heated regulatory controversy, and little thought was given to the kinds of cost-benefit analyses that Aigner spends a good deal of time discussing. There was no *inherent* reason for these early experiments to have been so poorly designed. There already existed a reasonably good model to build on in the English Domestic Tariff Experiment conducted between 1967 and 1972, and the analysts working with these early experimental residential data used precisely the same welfare model that Aigner presents in his paper. In short, the early experiments were poorly designed because they were poorly motivated and had very narrow objectives.

Some of the more recent experiments have been better structured and have yielded more useful information, both because the Department of Energy established some minimal guidelines and because some utilities have come to realize that it is in their interests to better understand how residential consumers respond to TOU rates, how such rates affect short-run revenues, and how such rates can be used strategically for system planning and load forecasting. Although there remains some utility hostility to TOU rates and skepticism about the value of the information generated by experiments like this, many utilities have come to realize that it is in their interest to squeeze as much information as they can out of such experiments since this information is potentially useful to them for planning and regulatory purposes. The best experiments have been done in situations where the ultimate consumers of the information (utilities and regulators) really cared about using it effectively and where the experiments were part of a broader-load research-and-system-planning effort. Furthermore, a major contribution of these experiments was to get utilities and regulators to begin to think seriously about rate-structure reforms and consumer responses to changing rate structures.

A second issue involves the neoclassical welfare analysis discussed at length in Aigner's paper. Are regulatory commissions and utilities likely to be guided by these kinds of calculations? I believe the answer is almost certainly no. Regulatory commissions appear to be guided in their policy decisions by three considerations: Can TOU pricing reduce the need for additional generating capacity and help to moderate requests for rate increases? Can the results of TOU pricing experiments be used to determine who gains and who loses as a result of changes in rate structure? Will TOU pricing reduce customer bills in the short run?

There is general acceptance among regulatory commissions of the notion that on the margin, all consumers should face prices that reflect the replacement costs of electricity. However, this intuitive understanding of the role of prices in consumer decisions and the relationship between consumption decisions and electricity supply has not been translated into broad acceptance of marginal-cost pricing principles by regulatory commissions. Whatever the academic interest in more refined welfare calculations, regulatory commissions do not understand what deadweight losses are, would not care much about them if they did, and, as a result, more refined calculations are unlikely to have any policy effects. I should note here that the decision to offer general TOU rates in England (the White Meter Tariff) was made before the Domestic Tariff Experiment was completed. Furthermore, the welfare calculations performed after the experiment was completed indicated that the welfare gains from TOU pricing were insufficient to cover the additional metering costs. The new rates went into effect before the experimental information was in and stayed in effect despite the negative welfare calculations. Regulatory commissions in the United States are primarily concerned with the average level of the electricity prices and issues associated with the distribution of the revenue burden among different types of customers. Their decisions for or against marginal-cost pricing and TOU pricing reflect these concerns. Any results based on conventional welfare calculations are only likely to have an impact if they can be placed in this context.

A third set of issues concerns the general usefulness of the TOU pricing experiments for long-range planning and load forecasting by utilities and their regulators. Most utility planners who understand the economic rationale for TOU pricing and are even sympathetic to the underlying objectives of efficiency based TOU pricing, do not find the experimental results to be particularly useful. A major reason for this view is the correct perception that by their very nature these experiments only allow us to estimate short-run elasticities of demand, given existing appliance stocks. Many utility planners envision the potential for substantial changes in the composition of appliance stocks as consumers repond both to generally higher energy prices and to TOU electricity prices. The changing appliance stock will draw on both existing appliance technology and appliance innovations. Of special interest are storage heating and cooling systems, storage hot-water heaters, dual fuel heat pumps, as well as changes in conventional appliances that will allow for better exploitation of TOU rates. System planners are most interested in examining the impact of TOU rates in the context of disaggregated appliance-specific load-forecasting models which can be conveniently coordinated with system dispatch and probabalistic planning models. As utility planning and load forecasting has become more sophisticated, the most progressive utilities have come to follow the European example of trying to coordinate pricing policy with appliance research and the provision of appliance information to consumers. Overall, utility planners have not found the aggregate econometric demand work that has been forthcoming from the experiments particularly useful because it reflects only short-run responses and has been conducted at too high a level of aggregation.

The lack of interest in the experimental results also reflects a perception that electricity pricing systems based primarily on time of use do not really represent the most effective way to give consumers signals reflecting the true marginal cost of production. In a number of cases utilities have gone well beyond conventional TOU pricing and are developing more sophisticated and efficient pricing systems. For example, for those systems in which peak demands are very sensitive to variations in temperature and humidity, there is a desire to tie the price signals more directly to weather-related periods of coincident peak demand than can be done with predetermined prices established for broad time periods of the day during several months of the year. A variety of radio, cable, and transmission-line communication devices are being developed. These will allow for more flexible interactive pricing systems, including interruptible tariffs that provide reduced rates but allow the utility to directly control appliances during peak periods. Coincident peak-sensitive pricing schemes are being offered to some industrial customers and experiments are under way with controlled storage heating, cycling of air conditioners, dual fuel heat pumps and other appliances. The residential TOU experiments simply do not provide the information on long-run consumer behavior that is of most importance in this effort.

TOU pricing experiments have had and will have little if any positive impact on regulatory decisions to implement TOU rates and have provided information on consumer behavior that is of only limited value in planning and load forecasting. The experiments have shown that consumers respond to higher prices by reducing consumption in the short run. For those who really doubted that such responses would occur, perhaps the results will finally convince them. Those studies that have found that peak and off-peak consumption were complements are of some potential interest, but I suspect that this result is an artifact of the focus on short-run responses given appliance stocks and are in any event presented at too high a level of aggregation to be of general use. Even where short-run elasticity estimates might be of value, the uncertainty ranges associated with the estimates obtained from the same data sets are so large that they are of limited use for pinning down anything of interest with great precision. To some extent the residential TOU experiments may have been counterproductive. They have led to too much of a focus on residential customers, where metering and other transactions costs are relatively high, and have diverted attention from industrial customers where TOU rates are likely to be more productive and where we have the most evidence based on foreign experience. Furthermore, the existence of ongoing TOU rate experiments may have served as a convenient excuse to avoid making decisions to reform electric utility rates so that they better promote economic efficiency.

This is not to say that there has been no progress on the rate-reform front. TOU rates are now available on a mandatory or voluntary basis in many states. Regulatory commissions and utilities have come to take rate-structure reforms, including TOU rates, much more seriously than they did ten years ago. The Public Utility Regulatory Policy Act requires states to consider the cost basis on which electricity rates are based and to determine whether TOU rates should be instituted. Recent Federal Energy Regulatory Commission rules for determing the rates that utilities must pay for power generated by cogeneration and other small power production facilities are based on the kinds of marginal-cost principles that have motivated economists to advocate the general application of marginal-cost pricing to electric-power rate making. But these reforms have proceeded largely independently of the residential TOU experiments. Commissions and utilities that have gone forward have taken the bull by the horns; they have recognized that consumption behavior is sensitive to prices and that rates based on marginal cost give better signals to consumers than rates based on average historical costs. They have recognized that TOU rates are almost certainly justified for large customers, given conservative assumptions about the relevant elasticities, that it takes a long time to economically re-meter the system anyway, and, therefore, that the most sensible thing is to gradually introduce permanent mandatory TOU rates starting with the largest consumers and to follow the behavior of these consumers with a carefully structured load research program. Mandatory programs have often been supplemented by voluntary programs in which customers wanting TOU rates must pay for their own meters and perform their own cost-benefit analyses. Some care in structuring the voluntary programs and gradual adjustments in the basic residential rates can help to avoid adverse selection problems that might otherwise develop.

Perhaps the most important change that has occurred during the past decade is that proposals for fundamental changes in all electric power rates are now taken seriously by regulators and utilities. TOU rates have come to be seen as one of several potential pricing and contracting innovations that must be viewed in the broader context of the development of a better understanding of appliance utilization, appliance choice, the development of new appliances that can use energy more economically, better sampling and load research work, and the development of more sophisticated load-forecasting and system-planning models. The TOU experiments and the econometric estimates of demand elasticities based on these estimates have provided some useful impetus to these developments, but I think it has been and will continue to be a small impetus. The most important effects have been indirect. The experiments have fostered more thoughtful discussion of rate-making alternatives that include TOU pricing, but have gone beyond pure time-related rates. The economic and statistical techniques that have been developed to analyze the data generated by these experiments will also certainly prove to be useful in the evolution of our understanding of consumer behavior as these techniques are applied to more extensive load-research information and new developments in pricing, metering, control, and appliance technology.

In short, the experiments have helped to focus the discussion and analysis of theoretical and empirical issues that arise when we consider broad issues of efficient pricing, consumer behavior, and load forecasting in the electric utility industry. The estimated elasticities themselves are of limited theoretical interest or public policy significance.

Comment Lester D. Taylor

In April 1975, the Federal Energy Administration funded six demonstration electricity pricing experiments for the purpose of generating data that could be used in assessing the costs and benefits of pricing electricity according to the time of day it is used. TOU pricing was viewed by policy makers as a possibly important conservation-inducing response to the energy crisis that was triggered by OPEC and the Arab embargo. Economists, as is well- known, are drawn to TOU pricing because of its firm basis in theoretical welfare economics. Indeed, scratch an economist, and if he doesn't say "supply and demand," he will probably say "peak-load pricing," for if done properly TOU pricing is economically efficient and leads to a maximum social welfare.[1]

However, in any practical situation, this is a counsel of perfection, because whether or not an existing nonoptimal pricing system should be scrapped depends upon the benefits to be gained in relation to the costs. In other words, the implementation of TOU pricing can be justified on social-welfare grounds only if it can be demonstrated that the change in social welfare is positive. However, calculation of the benefits and costs that would be involved requires a great deal of very detailed information on the structure of demand and costs. Information on costs is in principle available in the utilities, but, unlike utilities in Western Europe, U.S. utilities in 1973 had never engaged in TOU pricing, so there were no historical data from which estimates of the demand for electricity by time of day could be obtained. This was the information that was to be forthcoming from the FEA time-of-use pricing experiments.

In his paper, Dennis Aigner has attempted to assess the knowledge that has been obtained in the experiments. Altogether fifteen experiments figure in Aigner's assessment—the six original demonstration projects, plus nine subsequent experiments. Aigner's was a difficult one, and he has done a very good job. Besides providing an excellent overview of results, he has produced a coherent description of a framework for calculating the benefits and costs of TOU pricing plus a statement of the problems that will be faced in transferring the results beyond the sampled population.

Drawing on earlier work at the University of Michigan and the Research Triangle Institute, Aigner gives a useful summary assessement of the strengths and weaknesses in the designs of the various experiments. By now, there is almost general agreement as to which are the well-designed experiments and which are the ones to avoid, at least in terms of

Lester D. Taylor is professor of economics, University of Arizona.

1. Social welfare in this context is defined as the sum of consumers' and producers' surplus.

yielding transferrable information about the structure of demand. As he notes, the Arizona, California (both LADWP and SCE), Oklahoma, North Carolina, and Wisconsin experiments have the most potential in this regard. Of these experiments, data from the Arizona experiment became available the earliest, and, together with the Wisconsin experiment, has been the most worked over to this point.

Having been a party to the Arizona experiment (Jack Wenders and I designed the tariffs and the incentive scheme), I agree with Aigner that the Los Angeles and Wisconsin experiments are more useful than the Arizona experiment, although the Arizona experiment does contain some prime information, especially concerning the effects of incentive payments. As Aigner notes, the first year of the Arizona experiment contained a complicated implicit incentive payment that was designed to ensure voluntary participation in the face of TOU rates that reached as high as 16¢/kWh during the peak period. Until the paper of Hill et al. 1980, the implicit incentive payment had not been properly modeled (including my own Aspen effort). That this is important to the price elasticities is evident by comparing the results in my Aspen paper, which Aigner kindly describes as "anomalous," with the results in Hill et al. 1980. However, to return to the comparison of the Arizona with the Los Angeles and Wisconsin experiments, the Los Angeles and Wisconsin experiments are much larger than the Arizona experiment and are better designed in their nonprice aspects.

The really important question in all of this, however, is the one raised in the title of Aigner's paper: What have we learned from the experiments? I had hoped that Aigner would conclude that enough information now exists to provide at least a preliminary assessment of the costs and benefits of implementing TOU pricing, but those hopes are clearly dashed. Although he doesn't say it in so many words, my conclusion from reading Aigner's paper is that any scientific assessment of TOU pricing is still a long way off. The experiments are in reasonable agreement on own-price TOU elasticities of demand, but there is virtually no agreement concerning cross elasticities. Until these are defined with reasonable precision, a scientific assessment of TOU pricing is not possible. However, my gut feeling is that events will not wait for this to occur.

References

Acton, Jan Paul, and Bridger M. Mitchell. 1980. Evaluating time-of-day electricity rates for residential customers. In *Regulated industries and public enterprise: European and United States perspectives*, ed. B. M. Mitchell and P. R. Kleindorfer. Lexington, Mass.: Lexington Books.

Aigner, Dennis J. 1979. Bayesian analysis of optimal sample size and a best decision rule for experiments in direct load control. *Journal of Econometrics*, 9: 209–21.

Aigner, Dennis J., and Jerry A. Hausman. 1980. Correcting for truncation bias in the analysis of experiments in time-of-day pricing of electricity. *Bell Journal of Economics* 11 Spring: 131–42.

Aigner, Dennis J., and Lee A. Lillard. 1980. Initial analysis of Southern California Edison's domestic time-of-use experiment. Department of Economics, University of Southern California.

Aigner, Dennis J., and Dale J. Poirier. 1979. Electricity demand and consumption by time-of-use. Report EA–1294, Electric Power Research Institute, Palo Alto, Calif.

Atkinson, Scott E. December 1979a. A comparative analysis of consumer response to time-of-use electricity pricing: Arizona and Wisconsin. In *Modeling and analysis of electricity demand by time-of-day*, ed. D. Aigner. Report EA–1304, Electric Power Research Institute, Palo Alto, Calif.

———. 1979b. Responsiveness to time-of-day electricity pricing: First empirical results. In *Forecasting and modelling time-of-day and seasonal electricity demands*, ed. Anthony Lawrence. Report EA–578–SR, Electric Power Research Institute, Palo Alto, Calif. (Also published in *Journal of Econometrics* 9 [January 1979]: 70–96.)

Brown, Michael, N. J. Dikeman, Jr., Upton Henderson, David Huettner, Charles Massey, Donald Murry, Suhas Patwardhan, and Frank Wert. 1980. A summary of the Edmond Electric Utility Demonstration Project and its findings. Paper presented at the April Electric Rate Demonstration Conference, Denver, Colo.

Caves, Douglas W., and Lauritis R. Christensen. 1980a. Econometric analysis of residential time-of-use electricity pricing experiments. *Journal of Econometrics* 14: 287–306.

———. 1980b. Residential substitution of off-peak for peak electricity. *Energy Journal* 2: 85–142.

De Mouchel, W. H., and J. E. Harris. 1981. Bayes and empirical Bayes methods for combining cancer experiments in man and other species. Technical report No. 24, Department of Mathematics, Massachusetts Institute of Technology.

Efron, Bradley, and Carl Morris. 1975. Data analysis using Stein's estimator and its generalization. *J. Amer. Stat. Assn.* 70: 311–19.

Electric Utility Rate Design Study. 1977a. Considerations of the price elasticity of demand for electricity. National Economic Research Associates.

———. 1977b. Elasticity of demand. J. W. Wilson & Associates.

Engle, Robert, Clive Granger, Allen Mitchem, and Ramu Ramanthan. 1979. Some problems in the estimation of daily load shapes and peaks.

In *Modeling and analysis of electricity demand by time-of-day. See* Atkinson 1979a.

Granger, Clive W. J., Robert Engle, Ramu Ramanathan, and Alan Andersen. 1979. Residential load curves and time-of-day pricing: An econometric analysis. *Journal of Econometrics*, 9 (January): 13–32.

Hausman, J. A., M. Kinnucan, and D. McFadden. 1979. A two level electricity demand model: Evaluation of the Connecticut time-of-day pricing test. In *Modeling and analysis of electricity demand by time-of-day. See* Atkinson 1979a.

Hendricks, Wallace, and Roger Koenker, 1979. Demand for electricity by time-of-day: An evaluation of experimental results. Paper presented at the Rutgers Conference.

Hendricks, Wallace, Roger Koenker, and Dale J. Poirier, 1978. Residential demand for electricity by time-of-day: An econometric approach. Report EA–704, Electric Power Research Institute, Palo Alto, Calif.

Hill, Daniel H., Robert M. Groves, E. Philip Howrey, A. Christopher Kline, Daniel F. Kohler, James M. Lepkowski, and Marc A. Smith. 1979. Evaluation of the FEA's load management and rate design demonstration projects. EA–1152, Electric Power Research Institute, Palo Alto, Calif.

Hill, Daniel H., Deborah A. Ott, Lester D. Taylor, and James M. Walker. 1980. The impact of incentive payments in time-of-day electricity pricing experiments: The Arizona experiment. Mimeo.

Lau, Lawrence J., and Lee A. Lillard. 1979. A random response model of the demand for electricity by time-of-day. In *Modeling and analysis of electricity demand by time-of-day. See* Atkinson 1979a.

Lawrence, Anthony, and Steven Braithwait. 1977. The residential demand for electricity by time-of-day: An econometric analysis. In *Forecasting and modelling time-of-day and seasonal electricity demands*, ed. A. Lawrence. EA–578–SR, Electric Power Research Institute, Palo Alto, Calif. (Also published in *Journal of Econometrics* 9 [January 1979]: 59–78.)

Lifson, Dale P. 1980. Comparison of time-of-use electricity demand models. Research Triangle Institute, N.C.

Lillard, Lee A., and Jan Paul Acton. 1980. Season electricity demand and pricing analysis with a variable response model. Rand Corporation working draft #WD–670–DWP.

Lindley, Dennis, and F. M. Smith. 1972. Bayes estimates for the linear model. *J. Royal Stat. Soc.* 34: 1–18.

Miedema, Allen K. 1980. North Carolina rate demonstration project: Preliminary results. Paper presented at the Electric Rate Demonstration Conference, Denver, Colo.

Miedema, Allen K., Jerome A. Olson, Dale P. Lifson, and Bryan Krakauer. 1978. Time of use electricity prices: Arizona. Report, Research Triangle Institute, N. C.

Miedema, Allen K. 1980. North Carolina rate demonstration project: Preliminary results. Paper presented at the Electric Rate Demonstration Conference, Denver, Colo.

Raiffa, H., and R. Schlaifer. 1961. *Applied statistical decision theory*. Cambridge: MIT Press.

Taylor, Lester. 1977. On modeling the residential demand for electricity by time of day. In *Forecasting and modelling time-of-day and seasonal electricity demands*, ed. A. Lawrence. EA–578–SE, Electric Power Research Institute. (Also published in *Journal of Econometrics* 9: 97–118.)

U.S. Department of Energy. 1978. Analytical master plan for the analysis of data from the Electric Utility Rate Demonstration projects. Research Triangle Institute, N. C.

White, S. B., C. A. Clayton, B. V. Alexander, and D. P. Duncan. 1978. Analysis of the effects of the time of use electricity rate in the Connecticut electricity utility demonstration project. Draft final report, Research Triangle Institute, N. C.

This Page Intentionally Left Blank