ESTIMATING THE DISTRIBUTIONAL IMPACT OF TIME-OF-DAY PRICING OF ELECTRICITY

E. Philip HOWREY and Hal R. VARIAN

University of Michigan, Ann Arbor, MI 48109, USA

We consider the general problem of recovering estimates of welfare measures such as willingness to pay, price indices, etc. from demand data with particular emphasis on the problem of unobserved taste variation across households. We model taste differences across households in a variance components framework and derive an iterative estimator for systems of demand equations with variance components which is relatively easy to compute. Finally we present an example of the methods we propose which involves time of day pricing of electricity. We are able to calculate the fraction of the population which would prefer such pricing policies to flat rate pricing.

1. Introduction

In the standard approach to demand estimation the observed demand behavior is assumed to be generated by maximization of a representative consumer's utility function subject to a budget constraint. The assumption of a representative consumer is adopted not for its inherent realism but for its analytical convenience; and in many cases it seems to work well as a tentative hypothesis.

However, in some situations we may well desire a more general model that allows for differences in tastes across households. The most general alternative specification would be a model with all consumers having arbitrarily different utility functions. However, such a general model is usually impractical to specify and estimate.

A reasonable intermediate case is one where tastes are allowed to vary across the population according to some parametrically specified distribution. In this case we may well be able to estimate the parameters of the frequency distribution that reflects the variation in tastes across the population.

A circumstance where the estimation of the distribution of tastes is of particular interest is when we want to evaluate the distributional impact of some proposed policy change. In many situations we would like to know not only the *average* impact on consumer welfare of a policy change, but also the *distribution* of some welfare measure across the population.

King (1982) has described how one might go about estimating the variation in willingness to pay across the population when this variation depends on differences in *observed* demographic and economic variables. However, there

will typically be further differences in willingness to pay that are not directly attributable to observed demographic differences in households. We interpret these differences as differences in tastes, although other interpretations may be possible.

Burtless and Hausman (1978) have estimated a model incorporating variation in tastes in the context of labor supply. They specified that the frequency distribution of an income elasticity was truncated normal and estimated the parameters of this distribution by an iterative maximum likelihood technique. They did not explicitly calculate the welfare distribution implied by their estimated parameter distribution but were well aware that this would be possible. Below we show that much simpler estimation techniques can be used when the distribution of tastes can be assumed to be normal, rather than truncated normal, and we use the parameters derived by our estimation procedure to calculate the distribution of an appropriate measure of welfare.

The remainder of the paper proceeds as follows. First we examine the concept of willingness to pay and related welfare measures and show how these measures can be explicitly calculated as a function of the unknown parameters of the utility function in the one consumer case. We then postulate a model where the parameters of the utility function vary across the population according to some frequency distribution. We can then derive the implied demand equations and estimate the unknown parameters of the distribution of tastes using an error components model. The estimated parameters can then be used to calculate the distribution across the population of the willingness to pay for any particular policy change. Finally we illustrate these methods using some data involving time-of-day pricing of electricity.

2. The compensation function

What do we mean by the willingness to pay and related measures of welfare? In this section we attempt to give a meaningful empirical content to this concept. Further discussion can be found in King (1983) and Varian (1979), (1984).

We begin with the indirect utility function for some specific individual which we denote by $v(p, e)$. The indirect utility function measures the maximum utility the consumer can attain given prices p and expenditure e. Associated with this indirect utility function is its inverse, the expenditure function, denoted by $e(p, u)$. The expenditure function measures the minimum expenditure necessary to achieve a particular utility level u.

Suppose now that we are comparing two possible configurations of prices and expenditure which we denote by (p, e) and $(\bar{p}, \bar{e})$. We can ask how much money the consumer would need at prices p to be as well off as he would be in the situation described by $(\bar{p}, \bar{e})$. We denote this number by $\mu(p; \bar{p}, \bar{e})$. From the definition of the indirect utility function and the expenditure function, we

have

$$\mu(p; \bar{p}, \bar{e}) \equiv e(p, v(\bar{p}, \bar{e})).$$

Following Hurwicz and Uzawa (1971) we refer to the function $\mu(p; \bar{p}, \bar{e})$ as the 'income compensation function' or sometimes just as the 'compensation function'. King (1983) refers to the same concept as the 'equivalent income function'.

A reasonable measure of the willingness to pay to avoid a movement from the situation (p, e) to the situation $(\bar{p}, \bar{e})$ is given by

$$W = \mu(p; p, e) - \mu(p; \bar{p}, \bar{e}) = e - \mu(p; \bar{p}, \bar{e}).$$

By construction, a consumer who has income $e - W$ at prices p can reach the same level of utility as he could with income $\bar{e}$ facing prices $\bar{p}$. Hence this seems like a sensible way to measure the welfare impact of some policy change. Of course, W as we have defined it above, is simply the negative of Hicks' notion of the 'equivalent variation' – it is how much expenditure would have to change at prices p so as to make the welfare situation of the consumer at prices p *equivalent* to that obtained at $(\bar{p}, \bar{e})$.

The compensation function can also be used in ratio form to define various measures of the 'change in the cost of living'. Consider, for example, the expression

$$\pi(p, \bar{p}) = \mu(p; \bar{p}, \bar{e}) / \mu(\bar{p}; \bar{p}, \bar{e}) = \mu(p; \bar{p}, \bar{e}) \bar{e}.$$

The price index π measures how much expenditure one would need at prices p to be as well off as one would be at $(\bar{p}, \bar{e})$, relative to the actual expenditure at prices $\bar{p}$.

If π is greater than one it costs more at prices p to reach the utility level $v(\bar{p}, \bar{e})$ than it did at prices $\bar{p}$. In this sense the overall price level has increased in the movement from $\bar{p}$ to p.

If the underlying preferences happen to be homothetic, the indirect utility function can be chosen to be multiplicatively separable: $v(p, e) = \nu(p)e$. This in turn implies that the compensation function also has a multiplicatively separable form,

$$\mu(p; \bar{p}, \bar{e}) = \bar{e}\nu(\bar{p}) / \nu(p),$$

and that the price index has the form

$$\pi(p, \bar{p}) = \nu(\bar{p}) / \nu(p).$$

Thus in the homothetic case the price index is a function only and does not involve the level of expenditure.

It is worthwhile to note that $\mu(p; \bar{p}, \bar{e})$ behaves exactly like an expenditure function with respect to variations in p, holding $(\bar{p}, \bar{e})$ fixed. It also behaves like an indirect utility function with respect to $(\bar{p}, \bar{e})$ holding p fixed. This can easily be seen from the definition: for fixed p, $e(p, u)$ is an increasing function of u – if you want to get more utility at fixed prices you have to spend more money. Hence $e(p, v(\bar{p}, \bar{e}))$ is simply a monotonic transformation of the indirect utility function $v(\bar{p}, \bar{e})$ and is therefore itself an indirect utility function.

As an example of the above ideas, suppose that the indirect utility function is given by

$$v(p, e) = G(p) + e^{1-b}/(1-b),\qquad\qquad(1)$$

where $G(p)$ is some negative monotonic, quasi-convex function of prices. Such a utility function is of special interest because it generates demand functions which exhibit constant income elasticity. By Roy's law the demand for good j is given by

$$\ln x_j(p, e) = \ln\left(-\partial G(p)/\partial p_j\right) + b\ln e,$$

If $b = 1$ (the case of homothetic demand) then the indirect utility function in (1) takes the form

$$v(p, e) = G(p) + \ln e,$$

so that the demand functions have the form

$$\ln x_j(p, e) = \ln\left(-\partial G(p)/\partial p_j\right) + \ln e.$$

The expenditure function for an indirect utility function of form (1) can be found by solving for expenditure as a function of utility,

$$e(p, u) = \left[(1-b)(u - G(p))\right]^{1/(1-b)}.$$

Substituting $v(\bar{p}, \bar{e}) = u$, we have the income compensation function

$$\mu(p; \bar{p}, \bar{e}) = \left[(1-b)(G(\bar{p}) - G(p)) + \bar{e}^{1-b}\right]^{1/(1-b)}.$$

In the homothetic case, similar calculations show that

$$\mu(p; \bar{p}, \bar{e}) = \exp\left[G(\bar{p}) - G(p)\right]\bar{e}.$$

Thus the parametric specification and estimation of the demand functions and the indirect utility function is sufficient to identify and calculate the

compensation function $\mu(p; \bar{p}, \bar{e})$. For a specific example, which we will refer to later, consider the CES specification in which $v(p, e) = (\sum_j \alpha_j p_j^r)^{-1/r} e$. Then straightforward application of the above calculations shows that

$$\mu(p; \bar{p}, \bar{e}) = \bar{e}\left(\sum_j \alpha_j p_j^r\right)^{1/r} \bigg/ \left(\sum_j \alpha_j \bar{p}_j^r\right)^{1/r}.$$

3. Variations in tastes

We turn now to the specification of taste variation. Suppose that household i has an indirect utility function $v(p, e, \delta, \varepsilon)$, where δ is a vector of parameters specific to the household, and ε is a non-household specific error term. We suppose that δ is distributed across households according to the frequency function $h(\delta, \Delta)$, where Δ is a vector of unobserved parameters, and that ε has the usual properties of an error term.

The demand function for the good j by the household with characteristics δ and error term ε is given by

$$x_j = \frac{-\partial v(p, e, \delta, \varepsilon)/\partial p_j}{\partial v(p, e, \delta, \varepsilon)/\partial e}.$$

Given observations on (p, x, e) for a number of households, it will typically be possible to estimate the parameters in Δ and thereby construct an estimate of the variation in tastes across the population.

Suppose for example that we observe several choices made by household i over time and that the indirect utility function for household i takes the CES form described earlier,

$$v_i(p, e_i) = \left(\sum_j \alpha_{ij} p_j^r\right)^{-1/r} e_i,$$

where

$$\alpha_{ij} = \beta_j \exp(\varepsilon_{ijt} + \delta_{ij}).$$

In this case, the share equations for household i for good j at time t take the form

$$w_{ijt} = \left(\beta_j \exp(\delta_{ij} + \varepsilon_{ijt})\right) p_j^r / K,$$

where K is a constant depending on the parameters and prices.

The random variable δ_{ij} is specific to household i and remains fixed over time. The random variable ε_{ijt} is a disturbance term that varies over both

households and time. The variation of δ_{ij} over households is what we refer to as variation in tastes.

In order to estimate this model, it is convenient to normalize $\alpha_{ikt} = 1$, and consider the ratio of each share to the kth share. Taking logarithms we find

$$\log(w_{ijt}/w_{ikt}) = \log\beta_j + r\log(p_j/p_k) + \delta_{ij} + \varepsilon_{ijt}.$$

This gives us $k - 1$ equations that are linear in the relevant parameters with one cross-equation parameter restriction and an additive error components disturbance term.

4. Estimation of systems of equations with error components

We now consider how to estimate the parameters of a system of equations with error components using panel data. The general form that we consider derives from the expenditure share equations shown above; namely,

$$y_{ijt} = x_{ijt}\beta_j + u_{ijt},$$

where y_{ijt} is the observed value of the dependent variable in equation j at time t for household i, x_{ijt} is a vector of k_j explanatory variables, and β_j is a vector of k_j regression coefficients. (We have changed our notation a bit to conform with econometric practice.) The disturbance term u_{ijt} is assumed to be of the form

$$u_{ijt} = \delta_{ij} + \varepsilon_{ijt},$$

where δ_{ij} is that part of the disturbance term specific to equation j of household i.

The T time-series observations for equation j of household i can be written in matrix form as

$$y_{ij.} = x_{ij.}\beta_j + u_{ij.},$$

where $y_{ij.}$ is a column vector with elements $(y_{ijt}; t = 1, 2, \ldots, T)$, $x_{ij.}$ is a $T \times k_j$ matrix with x_{ijt} in row t, and $u_{ij.}$ is a column vector with elements $(u_{ijt}; t = 1, 2, \ldots, T)$. The vector $u_{ij.}$ can similarly be written as

$$u_{ij.} = \delta_{ij}e_T + \varepsilon_{ij.},$$

where e_T is a column vector of ones and $\varepsilon_{ij.}$ is a column vector with elements $(\varepsilon_{ijt}; t = 1, 2, \ldots, T)$. The mT observations for individual i can now be written

as

$$Y_i = Z_i \beta + U_i,$$

where

$$Y_i = \begin{bmatrix} y_{i1.} \\ y_{i2.} \\ \vdots \\ y_{im.} \end{bmatrix}, \qquad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}, \qquad U_i = \begin{bmatrix} u_{i1.} \\ u_{i2.} \\ \vdots \\ u_{im.} \end{bmatrix},$$

and

$$Z_i = \begin{bmatrix} x_{i1.} & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & x_{i2.} & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & x_{im.} \end{bmatrix}.$$

The disturbance vector can be expressed as

$$U_i = \begin{bmatrix} \delta_{i1} e_T \\ \delta_{i2} e_T \\ \vdots \\ \delta_{im} e_T \end{bmatrix} + \begin{bmatrix} \varepsilon_{i1.} \\ \varepsilon_{i2.} \\ \vdots \\ \varepsilon_{im.} \end{bmatrix} = \delta_i \otimes e_T + \varepsilon_i,$$

where $\delta_i \otimes e_T$ denotes the Kronecker product of the vectors δ_i and e_T.

Linear restrictions on the vector of regression coefficients,

$$R\beta = r,$$

are easily handled by transforming the variables. If R is a $g \times k$ matrix of rank g, it is possible to express g of the elements of β in terms of the remaining $k - g$ elements. Thus, by reordering the variables in x_{ijt} if necessary, we can partition β in such a way that

$$R_1 \beta^* + R_2 \beta^{**} = r,$$

were R_2 is a non-singular $g \times g$ matrix. Solving for β^{**} yields

$$\beta^{**} = R_2^{-1} r - R_2^{-1} R_1 \beta^*,$$

so that

$$\beta = \begin{bmatrix} I \\ -R_2^{-1} R_1 \end{bmatrix} \beta^* + \begin{bmatrix} 0 \\ R_2^{-1} r \end{bmatrix} = R^* \beta^* + r^*.$$

Substituting this expression into the regression model yields

$$Y_i^* = Z_i^*\beta^* + U_i,$$

where

$$Y_i^* = Y_i - Z_i r^* \quad \text{and} \quad Z_i^* = Z_i R^*.$$

Thus homogeneity and symmetry restrictions are easily handled for linear systems of demand equations.

The standard assumptions about the error components, namely,

$$E(\delta_{ij}) = 0,$$

$$E(\delta_{ij}\delta_{ik}) = \Delta_{jk},$$

$$E(\varepsilon_{ijt}) = 0,$$

$$E(\varepsilon_{ijt}\varepsilon_{iks}) = \Omega_{jk}, \qquad t = s,$$

$$= 0, \qquad t \neq s,$$

$$E(\delta_{ij}\varepsilon_{ikt}) = 0,$$

imply that

$$E(U_i) = 0,$$

and

$$E(U_i U_i') = \Psi = \Omega \otimes I_T + \Delta \otimes J_T,$$

where I_T is a $T \times T$ identity matrix and $J_T = e_T e_T'$ is a $T \times T$ matrix with each element equal to one. These assumptions allow for correlation among the individual specific effects δ_{ij} as well as contemporaneous correlation of the disturbances ε_{ijt} across equations. The assumption that the error components are normally distributed and independent across individuals completes the specification of the model.

It is shown in the appendix that maximum likelihood estimates of the parameters β (or β^*), $\Delta = (\Delta_{jk})$, and $\Omega = (\Omega_{jk})$ can be obtained by iterating the usual generalized least squares estimation procedure until convergence is achieved. In addition the structure of the covariance matrix Ψ can be used to simplify estimation of the elements in Δ and Ω. A relatively simple expression for Ψ^{-1} which is needed to calculate the generalized least squares estimates of β is also given in the appendix.

5. Empirical results

We now turn to an investigation of the demand for electricity by time-of-day. We consider a two-stage budgeting process of the form

$$q_i = q_i(p_1, p_2, p_3, e), \qquad e = e(\phi_1, \phi_2, y),$$

where

q_i = demand for electricity during period i ($i = 1, 2, 3$),
p_i = price of electricity during period i,
e = expenditure on electricity,
ϕ_1 = price index of electricity,
ϕ_2 = price index of all other goods, and
y = household income.

Our empirical results are based on data collected in 1976 by the Arizona Public Service Company. A random sample of 80 households in the Phoenix and Yuma service areas were assigned at random to the sixteen time-of-day rates shown in table 1. Electricity usage of these households was recorded for a six-month interval. The following results are based on the records of sixty of these households for the last five months of the experiment.[1]

The subutility function for electricity consumption on which our empirical results are based is of the CES form described earlier,

$$v(p, e) = \left(\sum_j \alpha_j p_j^r \right)^{-1/r} e.$$

This leads to the expenditure share equations

$$\log(w_{ijt}/w_{ikt}) = \log \beta_j + r \log(p_j/p_k) + \delta_{ij} + \varepsilon_{ijt},$$

for individual i at time t. Note that only two of the three share equations need to be estimated. The maximum likelihood estimates of the parameters are shown in table 2.

These parameter estimates exhibit two interesting features. First, the variation in the δ_{ij} across households is substantial. The estimated variance of δ_{i1}, for example, is 0.1450 which is nearly the same as the estimated variance of ε_{i1t}, which is 0.1414. Second, the estimated covariance between δ_{i1} and δ_{i2} is positive. Thus households that spend relatively less for electricity during one of these periods would also generally spend less during the other period.

---

[1]A detailed description of the experimental design is given in Hill et al. (1979). Incomplete data prevented the use of all eighty households in our analysis.

Table 1

Arizona TOD experimental rate schedules (¢/kWh).[a]

Rate group	Periods		
	Peak 2 p.m.–5 p.m.	Shoulder 9 a.m.–2 p.m. 5 p.m.–10 p.m.	Base 10 p.m.–9 a.m.
1	16	5	3
2	15	4	2
3	15	7	4
4	14	4	2
5	14	6	4
6	13	3	3
7	13	4	2
8	13	7	3
9	12	5	1
10	12	6	3
11	11	4	2
12	11	7	4
13	10	4	1
14	10	6	3
15	9	5	2
16	8	4	1

[a]*Source:* Taylor (1979).

Table 2

Maximum likelihood estimates of the expenditure share equations and error variances.

Coefficient	Estimate
$\log \beta_1$	−0.5551
$\log \beta_2$	0.4727
r	1.0335
Δ_{11}	0.1450
Δ_{22}	0.0912
Δ_{12}	0.0697
Ω_{11}	0.1414
Ω_{22}	0.1015
Ω_{12}	0.1047

6. Estimating the change in the price of electricity

The Arizona time-of-day pricing experiments are an interesting example for our purposes since they were specifically conducted in order to determine the feasibility of time-of-day pricing of electricity. The feasibility of time-of-day pricing depends, at least in part, on the willingness of households to accept those rates. Any initial resistance to time-of-day rates would presumably evaporate if households were to find that they were better off with TOD rates. Furthermore, the optimal design of time-of-day prices depends on consumers' utility functions for electricity consumption. Thus the estimation of the welfare impact of TOD pricing seems of considerable interest.

Hence we consider now how to measure the expenditure necessary at some time varying prices p to achieve the same level of subutility for electricity expenditure achieved at a flat rate schedule $\bar{p}$ and initial expenditure; that is, we wish to calculate $\mu(p; \bar{p}, \bar{e})$.

Recall the expression for the CES compensation function derived in section 2,

$$\mu(p; \bar{p}, \bar{e}) = \bar{e} \left(\sum_j \alpha_j p_j^r \right)^{1/r} \bigg/ \left(\sum_j \alpha_j \bar{p}_j^r \right)^{1/r}.$$

Using this expression, we can compute the fraction

$$\pi(p, \bar{p}) = \mu(p; \bar{p}, \bar{e}) / \mu(\bar{p}; \bar{p}, \bar{e}) = \mu(p; \bar{p}, \bar{e}) / \bar{e}$$

$$= \left(\sum_j \alpha_j p_j^r \right)^{1/r} \bigg/ \left(\sum_j \alpha_j \right)^{1/r} \bar{p},$$

which measures the relative change in the compensation function when moving from the flat rate $\bar{p}$ to time-of-day rates. It measures how much money one would need to have at the TOD rate schedule to have the same subutility one had at the flat rate schedule, expressed as a fraction of the expenditure at the flat rate schedule.

Since electricity consumption is only part of the entire consumption bundle, we cannot interpret $\pi(p, \bar{p})$ as a measure of the change in overall welfare. However, $\pi(p, \bar{p})$ is related to a particular price index for electricity consumption. We will briefly describe this interpretation below.

Recall the two-stage budgeting process mentioned earlier. The utility maximization problem involved can be written as

$$\max_{x, q} u(x, w(q)),$$

subject to

$$rx + pq = y,$$

where (r, x) are the vectors of prices and quantities of non-electricity consumption, and (p, q) are the analogous vectors for electricity consumption. The subutility function for electricity consumption, $w(q)$, is assumed to be homothetic. It follows that the compensation function will be of the form $\mu(\bar{p}; p, e) = ev(p)/v(\bar{p})$.

Using the compensation function as an indirect subutility function for electricity consumption, we can rewrite the consumer's maximization problem as

$$\max_{x, e} u(x, ev(p)/v(\bar{p})),$$

subject to

$$rx + e = y.$$

Letting $Q = v(p, \bar{p})\bar{e}$ be a 'quantity index' for electricity consumption, we can write this problem as

$$\max_{x, Q} u(x, Q),$$

subject to

$$rx + Qv(\bar{p})/v(p) = y.$$

Thus, $\pi(p, \bar{p}) = v(\bar{p})/v(p)$ serves as a price index for electricity consumption.

If π is greater than 1, then the price of electricity consumption has risen in the move from flat to TOD rates, and if it is less than 1, the price of electricity has fallen. Note that this price index can be given a welfare interpretation: if $\pi(p, \bar{p})$ is greater than 1, and no other prices change, the consumer is definitely worse off at time-of-day prices than at flat rate prices.

The above discussion is true for an arbitrary homothetic subutility function. For the CES case used in our empirical study, $\pi(p, \bar{p})$ is given by the explicit formula derived earlier,

$$\pi(p, \bar{p}) = [\alpha_1 p_1^r + \alpha_2 p_2^r + p_3^r]^{1/r} / [\alpha_1 + \alpha_2 + 1]^{1/r} \bar{p}.$$

The value of $\pi(p, \bar{p})$ is shown in table 3 for each of the sixteen rate schedules of table 1 and for four different flat rates: $\bar{p} = 4, 6, 8, 10$. The values in this table are computed for the average, or 'representative', household in the sample.

Prior to the experiment, households faced a declining block rate schedule and paid an average of approximately 4¢ per kWh. It is clear from the entries in table 3 that from the point of view of households only one of the TOD rate schedules is superior to the 4¢ flat rate. Discounts of nearly 50% are required to make households indifferent between the TOD rates and the 4¢ flat rate. As the flat rate increases, several of the TOD schedules become quite attractive, as we would expect. All of the TOD schedules are superior to a 10¢ flat rate.

The last column in table 3 gives the flat rate that is equivalent to the corresponding TOD schedule in the sense that the same expenditure on electricity gives the same level of subutility with the TOD rates and the equivalent flat rate. For example, the equivalent flat rate for rate schedule 1 is 6.40¢ per kWh. Flat rates below 6.40¢/kWh are preferable to TOD schedule 1, whereas for flat rates above 6.40¢/kWh, the TOD schedule is preferable.

The above figures are presented for the representative household. Since our econometric results indicated significant dispersion of tastes across households we also examine the variation in willingness to pay for TOD rates across households. The variation in tastes induces a variation in the price index $\pi(p, \bar{p})$, which can be easily calculated numerically.

Table 3

Estimate of the change in the cost of electricity.

Rate schedule	Flat rates				*PEQ*
	4	6	8	10	
1	1.6007	1.0672	0.8004	0.6403	6.40
2	1.3526	0.9018	0.6763	0.5411	5.41
3	1.8823	1.2549	0.9412	0.7529	7.53
4	1.3059	0.8706	0.6530	0.5224	5.22
5	1.7103	1.1402	0.8552	0.6841	6.84
6	1.2114	0.8076	0.6057	0.4846	4.85
7	1.2593	0.8395	0.6297	0.5037	5.04
8	1.7132	1.1421	0.8566	0.6853	6.85
9	1.2630	0.8420	0.6315	0.5052	5.05
10	1.5409	1.0272	0.7704	0.6163	6.16
11	1.1662	0.7775	0.5831	0.4660	4.66
12	1.6981	1.1321	0.8490	0.6792	6.79
13	1.0439	0.6960	0.5220	0.4176	4.18
14	1.4487	0.9658	0.7243	0.5795	5.79
15	1.1996	0.7997	0.5998	0.4798	4.80
16	0.9511	0.6341	0.4756	0.3805	3.80

Table 4

Probability of benefit from a switch from flat to TOD rates.

Rate schedule	Flat rates			
	4	6	8	10
1	0.000	0.198	0.996	1.000
2	0.003	0.850	1.000	1.000
3	0.000	0.003	0.834	1.000
4	0.005	0.938	1.000	1.000
5	0.000	0.013	0.995	1.000
6	0.006	0.994	1.000	1.000
7	0.006	0.975	1.000	1.000
8	0.000	0.020	0.992	1.000
9	0.011	0.970	1.000	1.000
10	0.000	0.316	1.000	1.000
11	0.024	0.998	1.000	1.000
12	0.000	0.006	0.999	1.000
13	0.301	1.000	1.000	1.000
14	0.000	0.725	1.000	1.000
15	0.005	0.999	1.000	1.000
16	0.713	1.000	1.000	1.000

Since δ_{ij} is a normal random variable, $\pi(p, \bar{p})$ will involve sums and ratios of lognormal variables, and its distribution function will not have a simple closed form expression. One useful way to characterize the variability of $\pi(p, \bar{p})$ is to calculate the probability that $\pi(p, \bar{p}) \leq 1$. This is the probability that a household selected at random would benefit from a switch from flat to TOD rates. The results of this calculation are shown in table 4.

For rate schedule 1, for example, virtually none of the households would benefit from a switch to TOD rates from a 4¢ flat rate. If the flat rate were 6¢/kWh, 19.8% of the households would benefit from a switch to the TOD schedule. The corresponding entry in table 3 indicates that the representative household would made worse off from this change. Thus while the representative household would be made worse off with the TOD rates, nearly 20% of the households would be made better off.

7. Summary

We have shown how one can estimate the distribution of willingness to pay across the population using panel data. In our example, there seems to be a significant dispersion of willingness to pay for time-of-day pricing of electric-

ity. Such dispersion can be taken into account in examining the welfare implications of policy choices by using the methods we have described above.

Appendix

In this appendix we sketch the derivation of maximum likelihood estimates of the parameters of a set of regression equations with additive error components. This provides a generalization of the single-equation results given by Graybill (1961) for the model

$$y_{it} = \mu + \delta_i + \varepsilon_{it}.$$

The recent work by Avery (1977) and Baltagi (1980), building on earlier results obtained by Wallace and Hussain (1969), Amemiya (1971), Nerlove (1971), and Maddala (1971), deals with systems of equations of the form

$$y_{it} = x_{it}\beta + \delta_i + \eta_t + \varepsilon_{it}.$$

In the model considered here, the η_t term is missing. The essential feature of this simpler model is that for a given value of β, maximum likelihood estimates of the variances of the error components can be calculated recursively. The operational result is that maximum likelihood estimates can be obtained by iterating the usual generalized least squares estimation procedure with analysis of variance estimates of the covariance matrices of the error components.

We use the notation of the text. Omitting the inessential constant term and multiplying by two, the (modified) log-likelihood for $(Y_1, Y_2, \ldots, Y_n)$ is

$$\mathscr{L} = -n \log |\Psi| - \sum_{i=1}^{n} (Y_i - Z_i\beta)' \Psi^{-1}(Y_i - Z_i\beta).$$

Setting the derivative of $\mathscr{L}$ with respect to β equal to zero, we find the likelihood equation for β,

$$\beta = \left(\sum_{i=1}^{n} Z_i' \Psi^{-1} Z_i \right)^{-1} \left(\sum_{i=1}^{n} Z_i' \Psi^{-1} Y_i \right).$$

Thus $\hat{\beta}$ is simply the generalized least squares estimate of β based on the maximum likelihood estimate $\hat{\Psi}$ of Ψ.

In order to obtain the likelihood equations for Δ and Ω, we write the likelihood function as

$$\mathscr{L} = -n \log |\Psi| - \sum_{i=1}^{n} U_i' \Psi^{-1} U_i,$$

where $U_i = Y_i - Z_i \beta$ is a function of β but does not depend on Δ or Ω. The following results will be used to rewrite the likelihood function in a more convenient form.

Proposition 1.   The determinant and the inverse of $\Psi = \Omega \otimes I_T + \Delta \otimes J_T$ are given by

$$|\Psi| = |\Omega|^{T-1}|\Omega + T\Delta|,$$

and

$$\Psi^{-1} = \Omega^{-1} \otimes I - (\Omega + T\Delta)^{-1}\Delta\Omega^{-1} \otimes J.$$

Proof.   Available from the authors by request.   ■

Using these results we can rewrite $\mathscr{L}$ as

$$\mathscr{L} = -n(T-1)\log|\Omega| - n\log|\Omega + T\Delta| - \sum_{i=1}^{n} \hat{\varepsilon}_i'(\Omega^{-1} \otimes I)\hat{\varepsilon}_i$$

$$-T\sum_{i=1}^{n} \hat{\delta}_i'(\Omega + T\Delta)^{-1}\hat{\delta}_i.$$

where

$$\hat{\delta}_{ij} = T^{-1}\sum_{t=1}^{T} u_{ijt} \quad \text{and} \quad \hat{\varepsilon}_{ijt} = u_{ijt} - \hat{\delta}_{ij}.$$

Thus the likelihood equation for Δ is

$$\Delta = n^{-1}\sum_{i=1}^{n} \hat{\delta}_i\hat{\delta}_i' - T^{-1}\Omega.$$

When Δ is determined according to this equation, the likelihood function becomes

$$\mathscr{L} = k - n(T-1)\log|\Omega| - \sum_{i=1}^{n} \hat{\varepsilon}_i'(\Omega^{-1} \otimes I)\hat{\varepsilon}_i,$$

where k is a function of $\hat{\delta}_i$ and hence β but not a function of Ω. A direct calculation reveals that

$$\sum_{i=1}^{n} \hat{\varepsilon}_i'(\Omega^{-1} \otimes I)\hat{\varepsilon}_i = \text{tr}(\Omega^{-1}E),$$

where the typical element of E is

$$E_{jk} = \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{\varepsilon}_{ijt}\hat{\varepsilon}_{ikt}.$$

Hence the likelihood equation for Ω is

$$\Omega = n^{-1}(T-1)^{-1}E.$$

Collecting these results, the maximum likelihood estimates must satisfy three sets of equations,

$$\beta = \left(\sum_{i=1}^{n} Z_i'\Psi^{-1}Z_i \right)^{-1} \left(\sum_{i=1}^{n} Z_i'\Psi^{-1}Y_i \right),$$

$$\Delta = n^{-1} \sum_{i=1}^{n} \hat{\delta}_i\hat{\delta}_i' - T^{-1}\Omega,$$

$$\Omega = n^{-1}(T-1)E.$$

The usual generalized least squares estimator provides the first step of an iterative process which, if it converges, will produce maximum likelihood estimates of the parameters. In particular, let $\hat{\beta}(1)$ be the least squares estimator of β. Define $\hat{U}_i(1)$, $\hat{\delta}_i(1)$ and $\hat{\varepsilon}_i(1)$ corresponding to this value of β. Initial estimates of $\hat{\Omega}(1)$ and $\hat{\Delta}(1)$ can now be obtained from the likelihood equations. This yields an initial estimate $\hat{\Psi}(1)$ of Ψ which can be used to obtain $\hat{\beta}(2)$, the generalized least squares estimate of β. This estimate can then be used to define $\hat{U}_i(2)$, $\hat{\delta}_i(2)$, and $\hat{\varepsilon}_i(2)$, and from these $\hat{\Omega}(2)$, $\hat{\Delta}(2)$, and $\hat{\Psi}(2)$ can be obtained. Continuing in this way until convergence is achieved we obtain the maximum likelihood estimates.

References

Amemiya, T., 1971, The estimation of the variances in a variance components model, International Economic Review 12, 1–13.

Anderson, T.W., 1958, An introduction to multivariate statistical analysis (Wiley, New York).

Avery, R.B., 1977, Error components and seemingly unrelated regressions, Econometrica 45, 199–209.

Baltagi, B.H., 1980, On seemingly unrelated regressions with error components, Econometrica 48, 1547–1551.

Burtless, G. and J. Hausman, 1978, The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment, Journal of Political Economy 86, 1103–1130.

Graybill, F.A., 1961, An introduction to linear statistical models, Vol. I (McGraw-Hill, New York).

Hill, D.H., R.M. Groves, E.P. Howrey, A.C. Kline, D.F. Kohler, J.S. Lepkowski and M.A. Smith, 1979, Evaluation of the Federal Energy Administration load management and rate design demonstration projects (Electric Power Research Institute, Palo Alto, CA).

King, M., 1983, Welfare analysis of tax reforms using household data, Journal of Public Economics 21, 183–214.

Hurwicz, L. and H. Uzawa, 1971, On the integrability of demand functions, in: Chipman et al., Preferences, utility, and demand (Harcourt Brace Jovanovich, New York).

Maddala, G.S., 1971, The use of variance components in pooling cross section and time series data, Econometrica 39, 341–358.

Nerlove, M., 1971, A note on error components models, Econometrica 39, 383–396.

Taylor, L.D., 1979, On modeling the residential demand for electricity by time of day, Journal of Econometrics 9, 97–115.

Theil, H., 1971, Principles of econometrics (Wiley, New York).

Varian, H., 1979, Notes on cost–benefit analysis, Working paper (University of Michigan, Ann Arbor, MI).

Varian, H., 1984, Microeconomic analysis, 2nd ed. (W.W. Norton and Co., New York).

Wallace, T.D. and A. Hussain, 1969, The use of error components models in combining cross section with time series data, Econometrica 37, 55–72.