MODELLING RESEARCH GROUP

DEPARTMENT OF ECONOMICS
UNIVERSITY OF SOUTHERN CALIFORNIA
UNIVERSITY PARK
LOS ANGELES, CALIFORNIA
90089-0152

USC M8607

MODELLING RESEARCH GROUP

A COMPARISON OF THREE FLEXIBLE COST FUNCTIONS USING ESTABLISHMENT LEVEL ELECTRICITY USE DATA

JOSEPH G. HIRSCHBERG

MRG WORKING PAPER #M8607

DEPARTMENT OF ECONOMICS
UNIVERSITY OF SOUTHERN CALIFORNIA
UNIVERSITY PARK
LOS ANGELES, CALIFORNIA
90089-0152

A COMPARISON OF THREE FLEXIBLE COST FUNCTIONS USING ESTABLISHMENT LEVEL ELECTRICITY USE DATA

JOSEPH G. HIRSCHBERG

MRG WORKING PAPER #M8607

Abstract

This paper compares the performance of three related systems of input demand equations in their ability to model time-of-day electricity demand. They are derived from the Translog, Generalized Leontief, and the Quadratic-Squareroot flexible cost function specifications. The data used in the estimation are firm-level monthly observations of seven time differentiated electricity inputs. Linear demand equations are specified in order to facilitate the firm-level estimation. The assumption of symmetric separability so often employed in demand studies of this nature has been relaxed in favor of non-symmetric separability which allows the substitution of non-electric inputs for electricity inputs. Comparisons of these specifications are made on the basis of the log-likelihood, in-sample prediction, and the precision of the price elasticity estimates. The results are reported for all firms combined and in terms of individual firm model comparisons.

Revised June 1986

I wish to thank Dennis J. Aigner and Denzil G. Fiebig for their comments and suggestions. Prepared under the auspices of NSF Grant #SES-8319129.

1. Introduction

The recent availability of various types of detailed panel data has lead to the question of specification choice for the modelling of individual agent behavior. Although microeconomic theory is primarily concerned with individual agent behavior most applications of theory-based demand models have used aggregate data. The computer induced revolution in the structure of agent transaction data now yields the detailed information that makes the specification and estimation of individual agent models possible.

This paper considers the choice of a cost function specification for individual firm demand systems. In particular, we compare models used to estimate firms' responses to time-of-day electricity pricing. We limit our consideration to three flexible cost functions that are related by their similar Box-Cox transformation form.

The data used in this analysis consists of monthly measurements on seven inputs that define characteristics of hourly electricity usage. The data are available for 60 months for the period 1976 to 1980 which included more than a year and a half before TOU rates were put into effect. These data are for the largest (in electricity usage) 103 firms in the Southern California Edison service area. First four inputs (A, B, C, and D) are defined respectively as the energy purchased in the periods: 8 a.m.-12 p.m. weekdays, 12 p.m.-5 p.m. weekdays, 5 p.m.-10 p.m. weekdays and all other weekday and weekend hours. The last three inputs (E, F, and G) are the levels of potential energy consumed* in the same periods as the inputs A,B, and C.

The paper will proceed as follows: Section two provides the general form of the class of flexible functional forms considered. Also given is the

---

*Potential energy consumption is defined as the deference between the maximum kW and the average kWh. See Hirschberg and Aigner (1983) for a detailed discussion of this concept.

rationale for loosening the separability assumptions usually employed when considering a subset of the entire set of goods or inputs demanded. Section three describes the estimation procedure and provides the error structure assumptions. In the fourth section we define criteria for comparing the estimates based on measures of fit, as well as the log-likelihood. An alternative to Theil's entropy based measure is proposed. These comparisons are made across establishments and specifications by firm type. Section five defines and reports the values of two comparison measures based on the precision of the price elasticity estimation. The final section summarizes the various comparisons and relates these findings to two earlier studies of aggregate data.

2. The Cost Function Specification

We use a cost function specification because the group of firms being modelled is so diverse. Military bases and manufacturing plants span the spectrum of the firms. (The distribution of firms by SIC is available in table 2.) In addition, the use of a profit function would also require the knowledge of output prices that are not available.

A general form of the Box-Cox type flexible functional form for cost-functions has been proposed by Berndt and Khaled (1979) and also by Applebaum (1979). In both cases a non-linear model was estimated using aggregate macro data involving only three or four inputs. The model we use is different from either of these in some specification details, but the essential form is followed. Due to the large number of firms (and inputs) in our sample we employ linear input demand equations based on three common forms of the general Box-Cox flexible function.

The general form of the Taylor series family of flexible functional forms can be given as:

$$s(2r) = a_o + A*Q(r) + (1/2)*(Q(r)'*B*Q(r)) \qquad (1)$$

where s is the resulting scalar, cost for cost functions and output for production functions, A is a 1×m vector of constant parameters, B is an m×m matrix of constant parameters, Q is the appropriate vector of arguments, prices for cost functions and inputs for production functions, r is the Box-Cox parameter used in the determination of the variables in the model, and a_o is a scaling constant, and the Box-Cox transformation is of the form:

$$Q_i(r) = (Q_i^{r/2} - 1)/(r/2) \qquad (2)$$

By setting the value of the Box-Cox parameter various functional forms are obtained. If $r \to 0$ then we get Christensen, Jorgensen and Lau's (1971) Transcendental Logarithmic (TL) production and cost functions. With $r = 1$ we obtain Diewert's (1971) Generalized Leontief (GL) functional form and $r = 2$ provides us with another of Diewert's functions — the Quadratic Square Root function (SQ).

Our cost function specifications have the additional properties of a) linear homogeneity in the prices, b) positive valued first ordered derivatives, c) concave in the prices (which implies a negative semi-definite Hessian). Linear homogeneity in the prices can be imposed via the following restrictions on the values of the parameters: $\Sigma_{i=1}^{N} b_{ij} = 0$, $\Sigma_{i=1}^{N} a_i = 1 + ra_o$. The symmetry of the price effects and the form of the approximation of the second order Taylor series dictates that B is symmetric which implies the additional restriction: $b_{ij} = b_{ji}$.

One departure in this study from some previous studies of electricity demand and other studies of a subset of all inputs is in the nature of the separability assumptions needed when estimating a set of linear demand equations (see Chung and Aigner (1981), Hirschberg and Aigner (1983)). Many studies employing individual agent data lack quantity observations for all inputs. This is hardly surprising since economic theory proposes markets with single prices and an infinite number of transactions. For this reason the estimation of a complete system of input equations, is only attempted with aggregate data. The assumption of separability between bundles of inputs is assumed to allow the estimation of a less than complete system of input equations. Because the information concerning electricity consumption and the prices at which it is consumed is much more detailed than the data for demand of other inputs, it is commonly assumed that the demand for the other inputs is not influenced by the relative price differences of the electricity inputs and — the symmetric effect — that relative price changes of other inputs do not effect the mix of electricity related input consumption.

In the present analysis we also require some type of separability assumption but we use an asymmetric assumption that can best be described as a long-run/short-run type of distinction (Blackorby, Primont and Russell (1978)). We assume that relative prices of various electricity inputs do not influence the mix of non-electricity inputs (same as the normal separability assumption), however we do not assume the converse. We allow the prices of non-electricity factors to influence the mix of TOU electricity inputs.

The most important factor in this regard is the wage rate. The efficiency of a shift in work schedules, which may be a reaction to changes in TOU prices for electricity, are mitigated by the resulting change in the wage bill. Thus, although we assume dependence of the mix of TOU electricity inputs on the wage rate, we also assume that TOU price differences for electricity do not have any influence on labor demand. Using this logic we include prices of such inputs as labor and other energy inputs in our cost function for TOU electricity inputs as factors which are exogenous to the total cost of electricity.

To incorporate the assumption of non-symmetric separability we add the non-electric prices or exogenous input prices, exogenous because these inputs are outside our cost specification. The cost function is then written as:

$$C(2r) = a_o + A'P(r) + \frac{1}{2}P(r)'BP(r) + P(2r)'GP_o(r) + P(2r)'WY(r) \tag{3}$$

where C is total cost, P is the vector of electricity input prices, the elements of P_o are the prices of the separable inputs: labor, gas, and petroleum products (transformed to constant dollars using the manufacturing price index), and Y is the proxy for output.

In addition to the restrictions on B and A linear homogeneity of the cost function also results in restrictions on G and W of the form: $\Sigma_{i=1}g_{ik} = 0$ (where k is the exogenous price subscript) and $\Sigma_{i=1}w_i = 0$.

We can now apply Shephard's Lemma (Shephard (1953)) to derive the appropriate input demand equations from our cost function.

We define the input equation specification as:

$$X_{it} = C^{1-r}H_{it}\theta_i , \qquad i = a, b, \ldots, g \qquad t = 1, 2, \ldots, T \ (T = 48 \text{ to } 60), \tag{4}$$

where the X_{it}'s are the dependent variables and the H_{it}'s are functions of the input prices, and θ_i is the parameter set for input i. The vector H_{it} is a function of prices defined when $r \neq 0$ as

$$H_{it} = \left(\frac{1}{P_i}\right)\left(\frac{2P_i^{r/2}}{r}\right)\left[(P_a^{r/2}, \ldots, P_g^{r/2}) \vdots P_i(P_{o1}(r), \ldots, P_{o3}(r), Y(r), 1)\right] \tag{5}$$

where P_i is price of electricity input i, the $P_{oj}(r)$ are the Box-Cox transformed exogenous prices, Y(r) is the Box-Cox transformed output proxy,* and r is the Box-Cox parameter. When r = 0;

$$H_i = (1/P_i)[\ell n(P_a), \ldots, \ell n(P_g) : \ell n(P_{o1}), \ldots, \ell n(P_{o3}), \ell n(Y), 1]. \tag{6}$$

---

*The output proxy was the predicted total electricity usage from a regression of a cubic in time.

In all cases:

$$\theta_i = (b_{i1},\ldots,\ b_{i7}:g_{i1},\ldots,\ g_{i3}:w_i:a_i). \qquad (7)$$

However due to the homogeneity restrictions $a_i = 0$ for all cases except the Translog (when $r = 0$).

The primary reason for assuming separability now becomes apparent in the form of equation (4). Because we need a value for total cost (C) we can solve this equation in two ways: First, one could employ equation (6) for the value of C and substitute it into equation 4.* Alternatively we could substitute the actual value of C directly into equation 4. The first method requires the consideration of a nonlinear estimation problem which could be quite large depending on the number of inputs, but the second raises the spector of inconsistent estimates. The next section will describe the use of the second method in detail.

3. Estimation of the Demand Equations

This section describes the error specification and the estimation procedure.

The general form of the equations to be estimated are:

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{bmatrix} = \begin{bmatrix} C(X)_1^{(1-r)} & & & 0 \\ & C(X)_2^{(1-r)} & & \\ & & \ddots & \\ 0 & & & C(X)_m^{(1-r)} \end{bmatrix} \begin{bmatrix} H_1 & & & 0 \\ & H_2 & & \\ & & \ddots & \\ 0 & & & H_m \end{bmatrix}^{\theta} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ em \end{bmatrix}, \qquad (8)$$

*When $r \to 0$ we would be multiplying all parameters in the cost function by those in the input equation, and when $r = 2$ we would divide the input equation parameters by all the cost function parameters.)

where $C(X)_i^{(1-r)}$ is the T×T diagonal matrix of the costs taken to the (1-r) power (for r = 1 we obtain an identity matrix), e_i is the error for equation i (to be defined below), E(ee') is the covariance Ω, θ is the 84×1 vector of stacked parameter sets θ_i for each input, and H_i a T×12 matrix of price terms for each equation. Dropping the subscripts from the equation above we, rewrite (8) as:

$$X = C(X)^{(1-r)}H\theta + e \,. \tag{9}$$

In the case of both the TL and SQ forms, the total cost appears on the right side of equation (9) thus raising the question of consistency. We can define a generalized least squares estimate of θ given the covariance Ω as:

$$\tilde{\theta} = (H'C'\Omega^{-1}CH)^{-1}H'C\Omega^{-1}X \;^{\dagger} \tag{10}$$

To check for consistency we examine the probability limit (Plim) of this estimate:

$$\underset{T\to\infty}{\text{Plim}}\ \tilde{\theta} = \theta + \underset{T\to\infty}{\text{Plim}}\ [(1/T)H'C(X)'^{(1-r)}\Omega^{-1}e] \,. \tag{11}$$

Appendix A demonstrates that the second term on the right hand side does not go to zero, thus indicating the need for an instrumental variable. The implementation of the instrumental variables estimation procedure is quite straightforward. We can show that if CI is the instrument for cost then our estimate is written as:*

$$\tilde{\theta}_{iv} = (H'CI'\Omega^{-1}CH)^{-1}H'CI\Omega^{-1}X \,. \tag{12}$$

If we multiply H by (1-r)(C/CI) and X by (1-r)(CI/C) we obtain the instrumental variables estimator when running the standard GLS procedure on these transformed regressors and input values. The appropriate estimated covariance matrix of $\tilde{\theta}_{iv}$ is computed by

$$\tilde{\text{Cov}}(\tilde{\theta}_{iv}) = (H'CI\tilde{\Omega}^{-1}CH)^{-1}(H'CI\tilde{\Omega}^{-1}CIH)(H'CI\tilde{\Omega}^{-1}CH)^{-1}. \tag{13}$$

---

*We compute the CI instrument from the actual prices and the predicted inputs from and ad hoc time specification.

†We have dropped the 1-r power from C in the following discussion of the estimators.

In another departure from many input demand studies we do not employ cost share equations. The primary reason is to reduce the potential consistency problems caused when prices appear on both sides of the equation, especially in light of the large variations in price caused by seasonal electricity rate shifts in the present case (see Hirschberg and Aigner (1983) for details). Also if prices are considered stochastic we can show consistency problems similar to the case for total cost shown above. Furthermore, using cost shares as dependent variables results in a singular cross equations covariance matrix, which requires the assumptions of a common autoregressive parameter for all equations (see Berndt and Savin (1975)). And in the present study we have reason to believe the inputs have different error structures based on their definition. (See Section 4 for further discussion). In addition, a number of researchers have been concerned with the truncated nature of cost shares in particular, Woodland (1979) proposes the use of a Dirichlet distribution for cost share equation errors, and Rossi (1983) suggest the log of the ratios of cost shares as the appropriate dependent variable in order to obtain a normally distributed error. Both of these solutions involve the use of non-linear estimation procedures of more complexity than is feasible for the individual firm models considered here.

We define each input equation with a different first-order serial correlation parameter ρ_i. That is: $e_{it} = \rho_i e_{it-1} + u_{it}$ (where $i = a,\ldots, g$; $t = 1,\ldots, T$), where u_{it} is an independently distributed normal variate with zero mean, constant variance (σ_{ii}) and covariances (σ_{ij}). Thus we obtain the cross-equation covariance $E(e_{it}e_{jt}) = \sigma_{ij}/(1-\rho_i\rho_j)$, for all i,j.

This specification is the ordinary seemingly unrelated regression (SUR) covariance matrix (Zellner (1962)) with the inclusion of a first order autoregressive parameter for each input equation as proposed by Parks (1967).*

To incorporate the restrictions implied by the linear homogeneity of the cost function in prices and the symmetry of B.** We compute the restricted least squares estimates using the substitution method. In the case where r = 0 we have a constrained minimization of the form: (using P as the matrix of regressors, $P = C(X)^{(1-r)}H$ in terms of equation (9))

$$\min_{\text{wrt } \theta}(X-P\theta)'\tilde{\Omega}^{-1}(X-P\theta); \qquad \text{st.} \quad R\theta = 0 \tag{14}$$

and the solution is given by;

$$\tilde{\theta} = \hat{\theta} - DR'(RDR')^{-1}(R\hat{\theta}) \tag{15}$$

where $D = (P'\tilde{\Omega}^{-1}P)^{-1}$ and $\hat{\theta}$ is the unrestricted GLS estimate. The substitution method provides the following non-restricted equivalent minimization:

$$\min_{\text{wrt } W}(X-PAW)'\tilde{\Omega}^{-1}(X-PAW) \tag{16}$$

where

$$\tilde{\theta} = A\hat{W}, \tag{17}$$

$$\tilde{\theta} = A(A'P'\tilde{\Omega}^{-1}PA)^{-1}A'P'\tilde{\Omega}^{-1}Y \tag{18}$$

and A is the matrix of eigenvectors corresponding to the zero valued eigenvalues of R'R. This equivalence is the result of a proof by Lawson and Hanson (1974).† In general, for efficiency reasons, the application of systems of demand equations to micro data necessitates the use of such methods for constrained least squares.

---

*Some investigation was made for higher order autoregressive error structure — 12th and 13th order to account for annual cycles — but a study of a subset of the establishments found little evidence for such an error structure. Also, because we have a seasonal price shift in the data it is necessary that our error specification does not remove the variance we are trying to explain by our model specification.

**The concavity condition cannot be imposed with linear restrictions.

†The matrix A is not unique up to a linear transformation and a number of authors have proposed alternative methods for similar restricted results (e.g., Dent (1980)).

Although computing the eigenvalues and eigenvectors of R'R is an extra computation, once they are available for a particular set of restrictions they can be applied to all models subject to the same parameter restrictions.

By incorporating the restrictions in this manner we accept the homogeneity and symmetry properties without testing them. Our estimation procedure is performed in three steps: First we compute the restricted OLS estimates using all seven equations in stacked form. Then we use these parameters to estimate the appropriate residuals and compute the autoregressive parameters ($\hat{\rho}_i$) for each equation. Using the estimated autoregressive parameters to transform the data we estimate a second model. Finally, the residuals from the second model are then used to estimate the cross equation covariance matrix ($\hat{\sigma}_{ij}$) which is then used to compute the parameters a third time. The third step was repeated at least twice in order to approximate an iterative Zellner solution.

The value of the log-likelihood evaluated at the estimates is given by (see Appendix C for the derivation):

$$\text{Log}(L) = \text{const.} + \frac{1}{2} \sum_{i=1}^{7} \log\,(1-\hat{\rho}_i^2) - \frac{T}{2} \log|\hat{\Sigma}| \;, \tag{19}$$

where $\hat{\Sigma}$ is the estimated cross equation covariance matrix, $\hat{\rho}_i$ is the estimated 1st order autoregressive parameter for the i^{th} equation, and const is a constant that does not vary by specification.

As T gets larger the estimates of ρ_i have a smaller and smaller impact on the value of the log-likelihood. This occurs because the $(1-\hat{\rho}_i^2)$ term appears as the transformation of the 1st observation of the time series, thus when the series is long it has less and less of an influence. And if we ignore the first observation the log-likelihood would solely be a function of $|\hat{\Sigma}|$.

4. Model Comparisons Using Measures of "Fit" and the Log-Likelihood.

We have made two types of model comparisons: those based on the "fit" to the data (in sample prediction) and those based on the precision of the price elasticity estimates of the inputs using the estimated asymptotic covariance matrix of the parameters. In this section we will describe the comparisons using the fit type measures with the precision based measures defined in Section 5. Under the first type we have a "raw" R^2 in the form of the square of the standard correlation coefficient between the estimated and the actual input levels, the values of the log-likelihood function, and "information" based measures proposed by Theil for the assessment of cost share predictions.

Table 1 shows the R^2 values* with a cross-firm comparison and the mean and median values for the measures. The cross-firm comparisons are made by ranking the R^2 values for each firm then counting the number of times each model has the highest or the lowest value. For example, for input A, 5 establishments have their lowest R^2 value for the TL specification, while 90 establishments have their highest R^2 value for the TL model. And we see that 60 establishments had their lowest R^2 value for the SQ specification applied to input A demand. Note that in general these R^2 values are quite low except for the commodities A-D for the Translog. However, there are a number of firms for which the GL and the SQ provide the greater correlation. Also on this table is the square of the correlation of the total cost computed from the predicted inputs. However, due to the properties of the TL it always predicts the cost used to evaluate the predicted inputs.**

---

*Note that due to the form of our demand equations no free intercept appears in the equations thus we cannot use the F-statistic equivalent measures proposed by McElroy (1977) or Buse (1978).

**(See appendix B for details).

Table 1

Comparisons of the R-Squared Values Computed on Actual Residuals
Across Cost Function Specifications, by Input Equation

			TL	GL	SQ
frequency	lowest	A	5	38	60
	highest		90	6	7
	mean		.684	.327	.251
	median		.764	.257	.185
frequency	lowest	B	5	36	62
	highest		91	6	6
	mean		.723	.352	.265
	median		.759	.301	.235
frequency	lowest	C	8	39	56
	highest		85	11	7
	mean		.684	.328	.288
	median		.742	.279	.269
frequency	lowest	D	11	40	52
	highest		82	12	9
	mean		.647	.331	.293
	median		.758	.296	.273
frequency	lowest	E	37	25	41
	highest		26	43	34
	mean		.059	.077	.070
	median		.033	.045	.037
frequency	lowest	F	42	30	31
	highest		26	34	43
	mean		.072	.081	.088
	median		.031	.032	.033
frequency	lowest	G	43	24	36
	highest		24	46	33
	mean		.152	.203	.164
	median		.067	.131	.108
frequency	lowest	Cost	0	17	86
	highest		103	0	0
	mean		1.0	.779	.570
	median		1.0	.906	.706

The next set of comparisons employs the value of the log-likelihood (LL) definded in equation (19). We find that the Generalized Leontief dominates the comparison of LL in all but 14 cases. This is at odds with the results obtained for the raw R^2 measure given above and is quite probably due to the need for an instrumental variable for cost in the TL and SQ specification.

Table 2 provides a summary of the results of the comparisons of the values of LL with the mean for the set of firms and the median, along with the distribution of these firm by firm comparisons by their standard industrial codes (SIC).

In some cases, especially for the SQ, we find very high cross-equation covariances that result in large $|\hat{\Sigma}|$ indicating very poor fit for some firms and resulting in low values of LL.

The very large LL values that we found in a few cases could be due not only to low variances but also to singularity of the $\hat{\Sigma}$ matrix. To check for singularity we computed the maximum entropy value of $\hat{\Sigma}$ using the method proposed by Theil and Fiebig (1984) which is assured not to be singular. The determinant of these estimates varied little from our other estimates and the rankings in Table 2 differed only slightly when using the maximum entropy value of $\hat{\Sigma}$.

Theil (1972, 1975) has proposed the use of information measures for the comparison of cost-share estimations. Theil made the analogy between cost-shares and the probabilities that a particular dollar spent will be spent on a particular input. The Theil measure will be designated as I and is defined as:

$$I(m,\hat{m}) = \sum_{i=1}^{N} m_i \log(\hat{m}_i/m_i) \ , \qquad (20)$$

where m_i, and $\hat{m}_i$ are the actual and predicted cost shares for inputs i.

Table 2

Comparisons of the Log-Likelihood Function

		TL	GL	SQ	# of firms
Food and Paper Products (2033,2621,2647)	lowest	2	0	3+	5
	highest	0	5*	0	
Chemical Products (2813,2819,2821, 2841,2874)	lowest	5	0	5	10
	highest	0	10*	0	
Petroleum Refining & Rubber (2911,3011)	lowest	7+	0	4	11
	highest	1	8*	2	
Concrete (3221,3241,3295)	lowest	1	0	10+	11
	highest	0	11*	0	
Metals Production (3312,3325,3351)	lowest	2	0	4+	6
	highest	1	4*	1	
Machinery Production (3533,3662)	lowest	4	0	5+	9
	highest	1	6*	2	
Transportation Equipment (3711,3724,3728,3761)	lowest	4	0	11+	15
	highest	0	13*	2	
Oil Fields & Pipelines (1311,1389,4613)	lowest	9	0	7	16
	highest	0	16*	0	
Services & Wholesale Trade (5141,7391,7813,8221, 8223,8911)	lowest	7+	0	4	11
	highest	0	10*	1	
National Security (9711)	lowest	3	0	5+	8
	highest	1	6*	1	
Total	lowest	44	0	59+	103
	highest	4	90*	9	
	mean	-20.1	336.8	-268.7	
	median	52.7	367.1	-154.5	

*The best model performance.
+The poorest model performance.

There are two problems with this measure. First $\hat{m}_i$ can be negative — thus $\log(m_i/\hat{m}_i)$ is not defined, and second, negative and positive values may cancel each other out, $\log(\hat{m}_i/m_i) > 0$ if $\hat{m}_i < m_i$ or $\log(\hat{m}/m_i) < 0$ if $\hat{m}_i > m_i$. For this reason we computed two alternative measures as well as $I(\cdot,\cdot)$.

First, to avoid the possibility of negative and positive information values canceling each other we use the absolute value of the log ratio:

$$IA(m,\hat{m}) = \sum_{i=1}^{N} m_i |\log(m_i/\hat{m}_i)| \tag{21}$$

and second we use a Taylor Series approximation to $I(\cdot,\cdot)$ of the form:

$$K(m,\hat{m}) = \sum_{i=1}^{N} (\hat{m}_i - m_i)^2/m_i \tag{22}$$

$K(\cdot,\cdot)$ as well as $IA(\cdot,\cdot)$ take equal stock of over and under estimation and K can also evaluate the influence of negative predicted values.*

Table 3 provides the rankings for the information measures. Note for $I(\cdot,\cdot)$ the higher values seem to indicate the better models because negative values seem to predominate. However Theil's $I(\cdot,\cdot)$ measure does not furnish a criteria to evaluate positive versus negative values. Positive values of IA and K are, however, nonambiguous indicators of poor models. Also, note how large the mean values for $K(\cdot,\cdot)$ become while the median values are much closer in magnitude. This is due to the very poor fit obtained for a few outlier firms. These criteria demonstrate that the TL no longer dominates the other specifications with the firm by firm ranks showing the SQ as the most successful for the $K(\cdot,\cdot)$ criteria but the GL results in lower median value. On the whole these results confirm the TL's superiority but also show that the other specifications are not as dramatically poor as the R^2 measures seem to indicate.

---

*Actually very few predicted values are negative less than 8% for the worst case (SQ) and less than 1% for the best case (TL).

Table 3

Mean and Median of Average Information Measures

	TL	GL	SQ	
I(·,·)	-.0378	-.0916	-.1631	mean
	-.0012	-.0068	-.0029	median
K(·,·)	46.4	37512.0	7640.0	mean
	.1003	.0653	.0768	median
IA(·,·)	.1839	.2923	.3723	mean
	.1063	.1621	.1681	median

Cross Specification Rankings of the Information Measures

	TL	GL	SQ
I(·,·)			
lowest	31	28	44
highest*	37	43	23
K(·,·)			
lowest*	30	22	51
highest	44	23	36
IA(·,·)			
lowest*	42	20	41
highest	24	42	37

*The best model performance.

Another method for evaluating these specification comparisons is the way in which they demonstrate the difference between various firm's technologies. Table 4 provides a comparison of K(·,·) by SIC group.† The *'s denote lower entropy or better predictions, and the +'s show the poorer performing models. Note the advantages for the modelling of Metals Production and Transportation Equipment industries for the translog specification while the square root quadratic form provides superior results for Machinery Production, Services and Wholesale Trade and National Security. The other firm groups do not imply such dominant results for any particular specification. These results provide an indication that the use of different technological modes is dictated by differences in industrial processes.

Table 5 is the equivalent table to Table 4 only prepared for IA(·,·) and it does not show such strong tendencies by firm type; in fact for National Security establishments we actually get a poorer fit for the quadratic square root than indicated by the K(·,·) criteria.

5. Precision Based Comparison Measures

The following measures are based on the precision of the elasticity estimates which depend on estimated asymptotic covariance matrix of the parameter estimates. Kumm (1981) proposes the use of the expected squared deviations of the estimated elasticities from their actual which we can show reduces to the trace of the estimated covariance matrix (where N is the vectorization of the elasticity matrix evaluated at the mean price and input values*) or the sum of the elasticity variances.

---

†One firm is in its own "group" so we report results for only 102 firms on this chart.

*The formulas for the estimation of the price elasticities are given in Hirschberg (1984) Appendix C.

Table 4

Comparison of Information Measure K(·,·) by SIC Group and by Cost Function

		TL	GL	SQ	# of firms
Food and Paper Products (2033,2621,2647)	lowest	2	2	1	5
	highest	1	0	4+	
Chemical Products (2813,2819,2821, 2841,2874)	lowest	2	3	5*	10
	highest	4	2	4	
Petroleum Refining & Rubber (2911,3011)	lowest	3	2	6*	11
	highest	6+	3	2	
Concrete (3221,3241,3295)	lowest	5*	3	3	11
	highest	4	1	6+	
Metals Production (3312,3325,3351)	lowest	5*	0	1	6
	highest	0	2	4+	
Machinery Production (3533,3662)	lowest	1	0	8*	9
	highest	7+	1	1	
Transportation Equipment (3711,3724,3728,3761)	lowest	9*	1	5	15
	highest	4	6+	5	
Oil Fields & Pipelines (1311,1389,4613)	lowest	1	7	8*	16
	highest	7+	3	6	
Services & Wholesale Trade (5141,7391,7813,8221, 8223,8911)	lowest	2	2	7*	11
	highest	6+	2	3	
National Security (9711)	lowest	0	1	7*	8
	highest	5+	3	0	
Total	lowest	30	22	51*	103
	highest	44+	23	36	

*The best model performance.
+The poorest model performance.

Table 5

Comparison of Information Measure IA(·,·) by
SIC Group and by Cost Function

		TL	GL	SQ	# of firms
Food and Paper Products (2033,2621,2647)	lowest	3*	2	0	5
	highest	0	2	3+	
Chemical Products (2813,2819,2821, 2841,2874)	lowest	2	4	4	10
	highest	0	4	6+	
Petroleum Refining & Rubber (2911,3011)	lowest	5	0	6*	11
	highest	4	5+	2	
Concrete (3221,3241,3295)	lowest	6*	2	3	11
	highest	3	4	4	
Metals Production (3312,3325,3351)	lowest	5*	0	1	6
	highest	0	3	3	
Machinery Production (3533,3662)	lowest	2	2	5*	9
	highest	3	4+	2	
Transportation Equipment (3711,3724,3728,3761)	lowest	8*	1	6	15
	highest	2	8+	5	
Oil Fields & Pipelines (1311,1389,4613)	lowest	4	6	6	16
	highest	6+	4	6+	
Services & Wholesale Trade (5141,7391,7813,8221, 8223,8911)	lowest	3	1	7*	11
	highest	5+	2	4	
National Security (9711)	lowest	4*	1	3	8
	highest	1	6+	1	
Total	lowest	42*	20	41	
	highest	24	42+	37	

*The best model performance.
+The poorest model performance.

$$E[(\hat{N}-N)'(\hat{N}-N)] = \sum_i \sum_j \hat{var}(\hat{n}_{ij}). \tag{23}$$

where $\hat{n}_{ij}$ is the estimated elasticity of input i with respect to price j. Because the elasticities are unitless measures a comparison of a function of the variances is a viable measure. (The elasticities are evaluated at the average input and price levels).

Next we consider the construction of a χ^2 statistic to test the null hypothesis that all elasticities are zero but we find that due to the nature of the symmetry conditions and the formulae for the elasticity matrices the covariance matrix is not of full rank so $\chi^2 = \hat{N}'\hat{cov}(\hat{N})\hat{N}$ cannot be computed. However if we only use the own-price elasticities we can make such comparisons. From the comparisons in Table 5 we find that the SQ specification results in the greatest value of the χ^2 in the most number of firms and in most of these values of χ^2 were significantly greater than zero. And in almost all cases under all three specifications, we can reject the null hypothesis that all the own-price elasticities are equal to zero.

From this comparison in table 6 we find a definite advantage for the GL model. This could be due to the absence of the use of instrumental variable covariance for this case. (Equation (13) need not be used when r = 1.)

Tables 6 and 7 provide the distribution of these comparisons by the same SIC designations given in Tables 3 and 4. From Table 7 we find a much higher prevalence of preferred models that employ the GL specification. All the industries except for Services and Wholesale Trade have more firms with lower values for the trace of the price elasticity matrix using the GL specification than either of the other two models.

From Table 7 we note that the SQ specification dominates the comparisons of the χ^2 of the estimated own-price elasticities in all but three sectors. Table 7 also provides the number of firms under each specification that reject

Table 6

Comparison of Trace of the Elasticity Covariance Matrix SIC Group and by Cost Function

		TL	GL	SQ	# of firms
Food and Paper Products (2033,2621,2647)	lowest	0	4*	1	5
	highest	2	0	3+	
Chemical Products (2813,2819,2821, 2841,2874)	lowest	0	8*	2	10
	highest	4	0	6+	
Petroleum Refining & Rubber (2911,3011)	lowest	2	8*	1	11
	highest	4	1	6+	
Concrete (3221,3241,3295)	lowest	0	9*	2	11
	highest	3	0	8+	
Metals Production (3312,3325,3351)	lowest	0	4*	2	6
	highest	2	1	3+	
Machinery Production (3533,3662)	lowest	0	5*	4	9
	highest	6+	0	3	
Transportation Equipment (3711,3724,3728,3761)	lowest	0	8*	7	15
	highest	8+	0	7	
Oil Fields & Pipelines (1311,1389,4613)	lowest	0	10*	6	16
	highest	13+	0	3	
Services & Wholesale Trade (5141,7391,7813,8221, 8223,8911)	lowest	1	4	6*	11
	highest	6+	1	4	
National Security (9711)	lowest	0	5*	3	8
	highest	7+	0	1	
Total	lowest	3	66*	34	103
	highest	55+	3	45	
	median	69.9	6.8	89.7	

*The best model performance.
+The poorest model performance.

Table 7

Comparison of χ^2 for Own-Price Elasticities by SIC Group and by Cost Function

		TL	GL	SQ	# of firms
Food and Paper Products (2033,2621,2647)	lowest	3+	0	2	5
	highest	2	1	2	
Chemical Products (2813,2819,2821, 2841,2874)	lowest	4	2	4	10
	highest	4	1	5*	
Petroleum Refining & Rubber (2911,3011)	lowest	6+	0	5	11
	highest	2	5*	4	
Concrete (3221,3241,3295)	lowest	4	2	5+	11
	highest	2	5*	4	
Metals Production (3312,3325,3351)	lowest	1	2	3+	6
	highest	1	4*	1	
Machinery Production (3533,3662)	lowest	7+	2	0	9
	highest	1	1	7*	
Transportation Equipment (3711,3724,3728,3761)	lowest	10+	1	4	15
	highest	2	4	9*	
Oil Fields & Pipelines (1311,1389,4613)	lowest	12+	0	4	16
	highest	3	4	9*	
Services & Wholesale Trade (5141,7391,7813,8221, 8223,8911)	lowest	8+	2	1	11
	highest	1	2	8*	
National Security (9711)	lowest	7+	0	1	8
	highest	1	0	7*	
Total	lowest	62+	11	20	103
	highest	19	28	56*	
Reject Ho at 5% significance	# of Firms	38	68	66	

*The best model performance.
+The poorest model performance.

the hypothesis that the own-price elasticity vector is equal to zero. It is interesting to note that although in one to one comparisons of each firm's χ^2 values the SQ is superior, the GL actually results in more cases where the elasticity estimates are significantly different from zero.

6. Summary

Table 8 summarizes Tables 2, 4, 5, and 7 by reporting the specification by industrial group that had the largest proportion of firms with the best performance. In addition we have the results from the log-likelihood comparison. We find that Services and Wholesale Trade is the one firm type with four out of five measures promoting the SQ specification, and there are a number of cases whereby three criteria we could agree on the specification of choice (Petroleum Refining Concrete, Metals, and Machinery). Interestingly we find that the specification that leads in many cases is the SQ the least used in empirical work of the three considered, and the TL, which is used widely, performs quite poorly.

In other studies comparing similar cost function specifications Applebaum (1979) found the models ranked as: 1) SQ, 2) GL, 3) TL and Berndt and Khaled (1979) found they could not reject the GL but could reject the TL and the SQ. However these were studies of aggregate data with less that half the number of inputs we consider here. In the above mentioned cases the models were estimated with a free r parameter, however this required the use of a more complex nonlinear estimation process.

A few caveats are in order with respect to the particular sample of firms used in this comparison. In general, earlier studies of this data (Hirschberg and Aigner (1982, 1983), Jazayeri (1984), and Tishler (1983)) found that the aggregate own-price elasticities of TOU energy (inputs A, B, C and D) for these firms are estimated to be very low. Thus the results of the present

Table 8

The Comparison of Best Model Indicators by Firm Type
(= the number of firms for which the specification is the best)

	LogL	K(·,·)	IA(·,·)	χ^2	tr(COV)	# of firms
Food and Paper Products (2033,2621,2647)	GL=5	--	TL=3	--	GL=4	5
Chemical Products (2813,2819,2821, 2841,2874)	GL=10	SQ=5	--	SQ=5	GL=8	10
Petroleum Refining & Rubber (2911,3011)	GL=8	SQ=6	SQ=6	GL=5	GL=8	11
Concrete (3221,3241,3295)	GL=11	TL=5	TL=6	GL=5	GL=9	11
Metals Production (3312,3325,3351)	GL=4	TL=5	TL=5	GL=4	GL=4	6
Machinery Production (3533,3662)	GL=6	SQ=8	SQ=5	SQ=7	GL=5	9
Transportation Equipment (3711,3724,3728,3761)	GL=13	TL=9	TL=8	SQ=9	GL=8	15
Oil Fields & Pipelines (1311,1389,4613)	GL=16	SQ=8	--	SQ=9	GL=10	16
Services & Wholesale Trade (5141,7391,7813,8221, 8223,8911)	GL=10	SQ=7	SQ=7	SQ=8	SQ=6	11
National Security (9711)	GL=6	SQ=7	TL=4	SQ=7	GL=5	8
TOTAL	GL=90	SQ=51	TL=42	SQ=56	GL=66	103

analysis may be due to the inherent "flexibility properties" of the cost functions as investigated by Caves and Christensen (1980). Except that in this case we are not investigating the ability of these cost function specifications to conform to the properties of an optimal solution.*

In conclusion we find all three specifications provide superior results for the majority of establishments' data under some criteria. The TL seems to dominate on the basis of the individual equation R^2 criteria, a result that is probably due to the TL's special ability to always predict the total cost used to evaluate the inputs (see Appendix B). The GL dominates all the other specifications when we use the log-likelihood as our measure of comparison. Furthermore, the GL has the advantage that total cost does not appear in the input equations thus no separability assumptions need be made and the input demand equation estimation procedure does not require the use of an instrumental variable for total cost. The SQ specification, although little used in empirical applications, does perform well in a number of cases under various criteria and may serve as a reasonable alternative.

---

*The concavity of the cost functions estimated was checked using the eigenvalues of the Hessian computed at the mean price values. However, because these are stochastic variables with no readily available way of estimating their second moments we have not reported them as a criteria for comparison.

Appendix A

The Consistency Problems of Including Actual Costs in Demand Equation Specifications

To show the inconsistency inherent in GLS estimation of the input demand equations (4) we take a first order Taylor series approximation of $C(X)^{(1-r)}$ around the mean of X, Xm = mean(X):

$$C(X)^{(1-r)} \cong Q1 + Q2\ e^*, \tag{A-1}$$

where:

$$Q1 = (rC(X_m)^{(1-r)})I_{mT} + ((1-r)C(X)^{(1-2r)})P^*\ \mathrm{diag}(\pi(P)B),$$

$$Q2 = ((1-r)C(X)^{-r})P^*,$$

$$P^* = (P':P':\ \ldots\ :P')'_{7T\times 7T}\ ,$$

$$P = [\mathrm{diag}(P_1):\mathrm{diag}(P_2):\ \ldots\ :\mathrm{diag}(P_7)]_{(T\times 7T)}\ ,$$

P_i = the T×1 vector of prices for input i etc.,

$\pi(P)$ = 7Txk block diagonal matrix with P_is on the diagonal,

B = kx1 vector of parameters,

$$e^* = (eo':eo':\ \ldots\ :eo'),$$

$$eo = [\mathrm{diag}(e_1):\mathrm{diag}(e_2):\ \ldots\ :\mathrm{diag}(e_7),]_{(T\times 7T)}\ ,$$

e_i = the T×1 vector of errors for equation i.

Consistency requires the right most expression from equation (11) to go to zero. Inserting the approximation to $C(X)^{(1-r)}$ we obtain:

$$\underset{T\to\infty}{\mathrm{Plim}}(1/T)(H'(Q1+Q2e^*)^{-1}\Omega e) = \Omega^{-1}\underset{T\to\infty}{\lim}(1/T)(H'Q1)\underset{T\to\infty}{\mathrm{Plim}}(1/T)e$$
$$+ \underset{T\to\infty}{\lim}(1/T)(H'Q2)\underset{T\to\infty}{\mathrm{Plim}}(1/T)(e^*\Omega^{-1}e). \tag{A-2}$$

By definition we have $\underset{T\to\infty}{\mathrm{Plim}}(1/T)e = 0$ so the first term goes to zero but the second term is non-zero and the term of interest is:

$$\underset{T\to\infty}{\mathrm{Plim}}(1/T)e^{*}\Omega^{-1}e, \qquad \text{(A-3)}$$

where we define the error covariance matrix in the form of the SUR type: $\Omega = (\Sigma \otimes I_T)$ ($\otimes$ denotes a Kronecker product). Therefore we can rewrite this term in another form as $F(\Omega) = e^{*}\Omega^{-1}e$ where $F(\Omega) = (f_1', f_2', \ldots, f_7')'$ T7×1, and: $f_i = \Sigma_{j=1}^{7} s_{ij} d_{ij}(LT)$, i = 1 to 7, s_{ij} = ij^{th} element of the Σ matrix, d_{ij} = ij^{th} element of the inverse of Σ matrix. Note that if we only had 2 commodities then e^{*} is symmetric and we get $F(\Omega)$ = LT2 (a column vector of all ones with length T2).

Appendix B

The Adding-Up Property of the Translog Specification

This appendix demonstrates that the predicted costs from a set of translog (TL) input demand equations will always be exactly equal to the cost used to evaluate the predicted input.*

We can write the typical TL input demand equation in the form,

$$X_{it} = \frac{C_t}{P_{it}}[\ell n(P_t)B_i + a_i] + e_{it} \qquad \text{(B-1)}$$

where: C_t = total costs at time t, P_{it} = price of input i at time t, $\ell n(P_t)$ = row vector of the log of all input prices at time t, B_i = the column vector of price coefficients for the i^{th} input, e_{it} = an additive error.

Once we have computed estimates of B_i and a_i we have the following expression for the estimated inputs: $\hat{X}_{it} = C_t/P_{it}[\ell n(P_t)\hat{B}_i + \hat{a}_i]$ evaluated at a particu-

---

*In the estimation we performed we use an instrumental variable for total cost but the predicted values are computed using actual cost.

lar cost C_t.* Thus our expression for the estimated total costs becomes: $\hat{C}_t = \Sigma_{i=1}^{n} P_{it}\hat{X}_{it}$.

On substituting the value of $\hat{X}_{it}$ from equation B-2 we find: $\hat{C}_t = C_t\Sigma_{i=1}^{n}$ $[\ell n(P_t)\hat{B}i + \hat{a}_i]$. Now we can rewrite the summed terms in a different form, using the summations implicit in vector multiplication,

$$\sum_{i=1}^{n}[\ell n(P_t)\hat{B}_i + \hat{a}_i] = \sum_{k=1}^{n} \ell n(P_{kt}) \sum_{i=1}^{n} \hat{b}_{ik} + \sum_{i=1}^{n} \hat{a}_i \qquad \text{(B-2)}$$

where b_{ik} is the k^{th} row of the coefficient vector B_i, and $\ell n(P_{kt})$ is the log of the k^{th} price at time t. We can now use the restrictions imposed on the parameters of a TL cost function to insure the homogeneity of degree one in the prices that provide: $\Sigma_{i=1}^{n}\hat{b}_{ik} = 0$, for all k, $\Sigma_{i=1}^{n}\hat{a}_i = 1$. Thus, we can show that $\hat{C}_t = C_t$, for all t.* This result only holds for the translog specification and may explain the superiority of this specification in some of our information based comparisons.

The ability of the translog specification to exactly predict the total cost can be shown to influence Theil's entropy based measure as well as the R^2 measure. We can decompose $I(\cdot,\cdot)$ as:

$$I(\cdot,\cdot) = \log(C/\hat{C}) + \sum_{i=1}^{N} m_i \log(\hat{X}_i/X_i) \qquad \text{(B-3)}$$

Thus for the translog case the first term always equals zero.

---

*If we had evaluated the predicted inputs using the instrumental variable CI for total cost the predicted cost would be equal to CI.

Appendix C

Derivation of the Log-Likelihood for a Set of SUR with a AR1 by Equation

The log Likelihood of the three specifications can be written in terms of the likelihood for the stacked vector of input errors E as;

$$\log(L) \equiv \frac{1}{2}\log(2\pi) + \frac{1}{2}\log(|\Omega^{-1}|) - \frac{1}{2}(E'\Omega^{-1}E), \qquad \text{(C-1)}$$

or

$$\log(L) = \text{const.} + \frac{1}{2}\sum_{i=1}^{7} \log(1-\hat{\rho}_i^2) - \frac{T}{2}\log|\hat{\Sigma}|, \qquad \text{(C-2)}$$

when evaluated at the estimates $\hat{P}_i$ and $\hat{\Sigma}$.

Where:

$\hat{E} = X - PA\hat{W}$, (from equation 23)

$\hat{\Omega}^{-1} = \hat{Q}'(\hat{\Sigma}^{-1} \otimes I_T)\,\hat{Q}$, (SUR with 1st order autoregressive parameters)

$\text{const.} = -\frac{1}{2}\log(2\pi) - \frac{1}{2}(T-q)$,

$\tilde{Q}$ = a 7Tx7T block diagonal matrix with matrices A_i on the diagonal,

$$A_i = \begin{bmatrix} \sqrt{1-\hat{\rho}_i^2} & 0 & \cdots & 0 & 0 \\ -\hat{\rho}_i & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & & -\hat{\rho}_i & 1 \end{bmatrix}_{TxT} ,$$

$\hat{\rho}_i$ = the 1st order autoregressive parameter estimated for equation i ,

$\hat{\Sigma}$ = the estimated cross equation covariance matrix.

To demonstrate this equivalence we first derive the determinant of $\hat{\Omega}^{-1}$. By properties of determinants and kronecker products (Graybill 1983, p. 184) we can write $|\Omega^{-1}| = |Q|^2|\hat{\Sigma}|^{-T}$. Furthermore, the determinant of a block diagonal matrix is the product of the determinants of the matrices on the diagonal, (Graybill 1983, p. 202) thus $|Q| = \Pi_{i=1}^{7}|A_i|$. And because the A_i's are triangular matrices their determinants are the product of the diagonal elements (Graybill 1983, p. 224) so $|A_i| = \sqrt{1-\hat{\rho}_i^2}$ and consequently we have $|Q|^2 = \Pi_{i=1}^{7}(1-\hat{\rho}_i^2)$. And the log $|\hat{\Omega}^{-1}|$ can be expressed as:

$$\log |\hat{\Omega}^{-1}| = \sum_{i=1}^{7} \log(1-\hat{\rho}_i^2) - T\log|\hat{\Sigma}| \qquad \text{(C-3)}$$

Now we show that because $\hat{\Sigma}$ is computed from $\hat{E}'Q'Q\hat{E}$ the $\hat{E}'\hat{\Omega}^{-1}\hat{E}$ term becomes a constant and does not vary by specification*.

Define u as a Tx7 matrix of the residuals of the regression transformed to account for the autoregressive parameters:

$$u = (A_1\hat{E}_1 \vdots A_2\hat{E}_2 \vdots \cdots\cdots\cdots \vdots A_7\hat{E}_7)_{Tx7} \qquad \text{(C-4)}$$

The estimate of the cross-equation covariance $\hat{\Sigma}$ is computed as: $u'u(1/T-q)$ where q is the average number of parameters in each equation. The right most term in equation (C-1) is of the form $(\hat{E}'\hat{\Omega}^{-1}\hat{E})$ which we can rewrite as:

$$(\hat{E}'\hat{\Omega}^{-1}\hat{E}) = tr\ [vec(u')'(\Sigma^{-1} \otimes I)\ vec(u)] \qquad \text{(C-5)}$$

Via properties of vectorization of matrices (Graybill 1983, p. 311) and of traces we can rewrite this as $(\hat{E}'\hat{\Omega}^{-1}\hat{E}) = tr[u'u\ \hat{\Sigma}^{-1}]$ and thus we find that $(\hat{E}'\hat{\Omega}^{-1}\hat{E}) = t - q$, or the number of degrees of freedom used in estimating the cross equation covariance matrix.

---

*It will vary only by the degrees of freedom used in computing $\hat{\Sigma}$.

References

Appelbaum, E., 1979, "On the Choice of Functional Forms," International Economic Review 20, 449-458.

Berndt, E.R. and M.S. Khaled, 1979, "Parametric Productivity Measurement and Choice Among Flexible Functional Form," Journal of Political Economy 86, 1220-1246.

Berndt, E.R. and N.E. Savin, 1975, "Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances," Econometrica 43, 937-957.

Blackorby, C., D. Primont and R.R. Russell, 1978, Duality, Separability and Functional Structure: Theory and Economic Applications, North-Holland Publishing: Amsterdam.

Buse, A., ????, "Goodness of Fit in the Seemingly Unrelated Regressions Model: A Generalization," Journal of Econometrics 10, 109.

Caves, D.W. and L.R. Christensen, 1980, "Global Properties of Flexible Functional Forms," American Economic Review 70, 422-432.

Christensen, L.R., D.W. Jorgenson and L.J. Lau, 1971, "Conjugate Duality and the Transcendental Logarithmic Function," Econometrica 39, 255-256.

Chung, C. and D.J. Aigner, 1981, "Industrial and Commercial Demand for Electricity by Time-of-Day: A California Case Study," The Energy Journal 2, 91-110.

Dent, W.T., 1980, "On Restricted Estimation in Linear Models," Journal of Econometrics 12, 49-58.

Diewert, W.E., 1971, "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function," Journal of Political Economy 79, 481-507.

Graybill, F.A., 1983, Matricies with Applications in Statistics, 2nd ed., Wadsworth: Belmont, California.

Hirschberg, J.G., 1984, "The Impact of Time-of-Use Electircity Pricing: A Firm Level Econometric Approach," unpublished Ph.D. dissertation, University of Southern California: Los Angeles.

Hirschberg, J.G. and D.J. Aigner, 1983, "An Analysis of Commercial and Industrial Customer Response to Time-of-Use Rates," The Energy Journal 4 (Special Electricity Issue), 103-126.

Hirschberg, J.G. and D.J. Aigner, 1982, "Econometric Analysis of Southern California Edison Company's TOU-8 Rate Class, 1976-1980," Revenue Requirements Department, Southern California Edison Company: Rosemead, California.

Jazayeri, A., 1984, "A New Approach to Modelling Peak Load Pricing Policies: The Case of Electricity Demand by Large Industrial Customers," unpublished Ph.D. dissertation, University of Southern California: Los Angeles.

Kumm, M.S., 1981, "A Comparison of the Generlized Box-Cox and Fourier Functional Forms: An Application to the North Carolina Residential Time-of-Use Electricity Pricing Experiment," unpublished Ph.D. dissertation, Duke University: Durham, North Carolina.

Lawson, C.L. and R.J. Hanson, 1974, Solving Least Squares Problems, Prentice Hall: Englewood Cliffs, New Jersey.

McElroy, M.B., 1977, "Goodness of Fit for Seemingly Unrelated Regressions," Journal of Econometrics 6, 381-387.

Parks, R.W., 1967, "Efficient Estimation of a System of Regressions Equations When Disturbances are Both Serially and Contemporaneously Correlated," Journal of the American Statistical Association 62, 500-509.

Rossi, P.E., 1983, "Specification and Analysis of Econometric Production Models," unpublished Ph.D. dissertation, University of Chicago: Chicago.

Shephard, R.W., 1953, Cost and Production Functions, Princeton University Press: Princeton, New Jersey.

Theil, H. 1972, Statistical Decomposition Analysis, North-Holland Publishing: Amsterdam.

Theil, H., 1975, Theory and Measurement of Consumer Demand, Volume 1, North-Holland Publishing: Amsterdam.

Theil, H. and D.G. Fiebig, 1984, "Exploiting Continuity: Maximum Entropy Estimation of Continuous Distributions," Ballinger: Cambridge, Massachusetts.

Tishler, A., 1983, "The Demand for Electricity Under Time-of-Day Pricing," Journal of Econometrics 23, 369-384.

Woodland, A.D., 1979, "Stochastic Specification and the Estimator of Share Equations," Journal of Econometrics 10, 361-383.

Zellner, A. 1962, "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," Journal of the American Statistical Association 57, 348-368.