A Note on a Method for Estimating the Linear Expenditure System with One Restriction

Seok Koo Lee*

1. A Brief Description on Classical Consumer Theory

We start with a single consumer with given money income μ, who purchases n commodities represented by the vector $\boldsymbol{q}$ at prices $\boldsymbol{p}$. He chooses $\boldsymbol{q}$ so as to maximize the utility index u which is a function of $\boldsymbol{q}$. Thus the consumer's objective is to maximize

$$u(\boldsymbol{q}) \text{ subject to } \boldsymbol{q} \geq 0 \text{ and } \boldsymbol{p}'\boldsymbol{q} \leqslant \mu.$$

In the above calculus-based treatment, the non-negativity constraint can be ignored, and perfect divisibility is assumed in order to allow the second inequality to be replaced by an equality. If in addition u is allowed to be twice differentiable, we may write the first-order maximization conditions as:

$$(1)\quad \boldsymbol{u}_1 - \lambda\boldsymbol{p} = 0$$

$$(2)\quad \boldsymbol{p}'\boldsymbol{q} = \mu,$$

where $\boldsymbol{u}_1$ is the vector of partial differentials of u with respect to $\boldsymbol{q}$. The Lagrange multiplier, λ, may be interpreted as the marginal utility of income (for the purpose of this paper, expenditure is substituted for income).

The n equations (1), stating that relative marginal utilities must equal

* Researcher, Korea Development Institute.

relative prices, together with the budget constraint (2), may be used to eliminate λ and thus to give the quantities, q, in terms of the known prices, $\boldsymbol{p}$, and income μ. Formally

$$(3)\quad \boldsymbol{q} = \boldsymbol{q}(\mu, \boldsymbol{p})$$

represent the n demand equations. At this point, we may remark that if we now replaced u by some monotone increasing function of u, $f(u)$ say, then equations (3) would be unchanged, though the value of λ in (1) would not remain constant. This justifies the use of u as one cardinal representation of an ordinal preference ordering or indifference mapping. These demand functions have three properties:

(i) They satisfy budget constraint.

(ii) They are homogeneous of degree zero in all prices and total income.

(iii) The implied Slutsky substitution matrix is symmetric and negative semidefinite with rank $n-1$.

The latter two are properties of the classical demand function. Since any set of demand functions that satisfies these three conditions is derivable from a well behaved utility function, we call such a set a "complete system of theoretically plausible demand functions."

2. Linear Expenditure System and Its Constraints

Over twenty-five years ago, Professor Klein and Rubin [1] presented a complete set of demand relations which have, over time, come to be known as the linear expenditure system (the name, linear expenditure system, is associated with the work Stone [6]). In matrix form, this system can be written as

$$(4)\quad \hat{\boldsymbol{p}}\boldsymbol{q} = \hat{\boldsymbol{p}}\boldsymbol{b} + \boldsymbol{a}(\mu - \boldsymbol{p}'\boldsymbol{b}),$$

$$\text{where}\quad \hat{p}=\begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \cdots & & p_n \end{pmatrix},\quad q=\begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix},\quad p=\begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix},$$

$$b=\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},\ \text{and}\quad a=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

In (4), $\boldsymbol{b}$ may be identified with a vector of quantities to which consumers are in some sense committed and $\boldsymbol{a}$ denotes a vector of constants which sum to unity. Then, on the hypothesis (4), the expenditure on the i-th commodity (here, category of commodities) is equal to a certain basis consumption, b_i valued at current prices plus a certain proportion, a_i of supernumerary income, measured here by total expenditure, μ less total committed expenditure, $\boldsymbol{p}'\boldsymbol{b}$. On the assumption that $\mu > \boldsymbol{p}'\boldsymbol{b}$, the system (4) entails that consumers first use up a certain amount of their total expenditure in acquiring the consumption vector $\boldsymbol{b}$ at current prices, whatever they may be, and then distribute their remaining income over the set of available commodities in fixed proportions given by the elements of $\boldsymbol{a}$. The linear expenditure system (4), under the assumption $\mu > \boldsymbol{p}'\boldsymbol{q}$, which is compatible with three conditions (that is, (i) budget constraint, (ii) homogeneity, and (iii) the implied Slutsky substitution matrix is symmetric and negative semidefinite) has the following constraints[1]:

i) $\sum_{i=1}^{n} a_i = 1.$

ii) $0 < a_i < 1$: this condition rules out inferior goods and complementary goods.

iii) The assumption $\mu > \boldsymbol{p}'\boldsymbol{b}$ means $q_i - b_i > 0$.

Geary and Samuelson [3], [4] demonstrated that the Klein-Rubin demand

1) See Stone [6] for a detailed explanation.

relations implied a utility function of the form:

$$(5)\quad U=\Phi \prod_{i=1}^{n} (q_i - b_i)^{a_i},$$

$$\text{where } a_i > 0,\ \sum_{i=1}^{n} a_i = 1.0,\ \text{and } q_i - b_i > 0.$$

Maximization of (5) subject to budget constraint yields the set of demand relations:

$$(6)\quad \boldsymbol{q} = (\boldsymbol{I} - \gamma \boldsymbol{p}')\boldsymbol{b} + \gamma \mu,$$

where γ is a $n \times 1$ vector with elements a_i/p_i, where the a_i are parameters to be estimated. And $\boldsymbol{I}$ is unit matrix. The relation (6) corresponds to the system (4). Here, after ruling out inferior goods and complementary goods (or, after grouping them together) we can avoid the constraint $0 < a_i < 1$ for any empirical research. Econometrics does not deal with exact relationship. Probabilistic considerations are fundamental throughout. The pure theory of consumer demand shows how a relationship may be derived, via the route of utility maximization, expressing quantity demanded as a function of relative prices and real income. There may be an epidemic, a disturbance of international relations, a sudden and temporary change of fashion, or, in fact, a veritable multitude of factors which may come to have a direct effect on demand, but which are not accounted for explicitly in the theoretical derivation of the relationship. These extra factors are not permanent, regular, or measurable, yet they are definitely present and are not necessarily negligible. If there is found to be a major systematic factor which strongly influences demand, and which is not included in the theory, then that theory must be altered to encompass this variable. Let us write the typical demand function, into which the disturbance terms enter additively, as

$$(7)\quad \boldsymbol{q} = (\boldsymbol{I} - \gamma \boldsymbol{p}')\boldsymbol{b} + \gamma \mu + \boldsymbol{v},$$

where $\boldsymbol{v}$ is a $n \times 1$ disturbance vector for estimating parameters ($\boldsymbol{a}$ and $\boldsymbol{b}$).

We use time-series data or cross-section data for our empirical research according to the related purpose. However there are difficulties to be considered in using two kinds of data.

3. Dynamic Specification and Stochastic Specification[1]

The linear expenditure system can be made more flexible by allowing the parameters (**a** and **b**) to vary systematically with variables that are exogenous to the system. This device would allow various dynamic specifications. In this paper, only the dynamic form of b_i's is discussed for the sake of a little more simplicity. We adopt the so-called habit forming model by considering the dynamic form of b_i's as a partial adjustment form.

$$(8)\quad b_{it}=\alpha_i+\beta_i z_{it-1},$$

where z_{it-1} is a variable representing consumption of the i-th good prior to period t. The two assumptions made based upon the level of consumption take z_{it-1} to be: (a) the highest level of consumption of the i-th good during the three years prior to period t, and (b) the average level of consumption of the i-th good during the three years prior to period t, by taking into account the habitual consumption pattern. Then, we can write the demand functions with varying b_i's as

$$(9)\quad q_{it}=b_{it}-\frac{a_i}{p_{it}}\sum_{k=1}^{n}p_{kt}\,b_{kt}+\frac{a_i}{p_{it}}\mu_i+v_{it},$$

where b_{it} is the necessary quantity of goods in period t. In matrix form this system is

$$(10)\quad \boldsymbol{q}=(\boldsymbol{I}-\gamma\boldsymbol{p}')\boldsymbol{b}+\gamma\mu+\boldsymbol{v}_t$$

for each period t. It would be convenient to assume that the $\boldsymbol{v}$'s for each time t are mutually independent. However, this assumption is inconsistent

1) Many points of this section follows Pollak and Wales [8].

with the budget constraint, which requires that

$$(11)\ \ \boldsymbol{i}'\boldsymbol{v}_i = 0, \qquad t=1, \cdots\cdots, T$$

where $\boldsymbol{i}$ is the $n \times 1$ column vector $(1, 1, \cdots\cdots 1)'$. That is, $\boldsymbol{v}$'s for each time are dependent upon each other. To satisfy (11), the covariance matrix of disturbance terms for each other must be singular.

In specifying the structure of disturbance terms, this paper considers the method which Pollak and Wales [8] feel is superior among three methods presented by them. This method is based on replacing b_{it} in each demand equation by $b_{it} + u_{it}$ where $\boldsymbol{u}_t$ is a random variable for each period. The implied stochastic demand functions of the form (10) where $\boldsymbol{v}_t$ is given by $\boldsymbol{v}_t = (\boldsymbol{I} - \gamma \boldsymbol{p}')\boldsymbol{u}_t$. The $\boldsymbol{u}$'s can be interpreted as random variations in the necessary baskets. Pollak and Wales demonstrated five implications of this kind of error structure. That is, they are: (i) $\boldsymbol{u}_t$ is directly related to $\boldsymbol{v}_t$, (ii) the adding-up condition (11) is automatically satisfied, (iii) proportional changes in all prices and income (providing these changes do not affect the distribution of the $\boldsymbol{v}$'s) will not affect the distribution of the $\boldsymbol{u}$'s, (iv) the Slutsky substitution matrix is symmetric and negative semidefinite regardless of the values assumed by the $\boldsymbol{v}$'s, and (v) this method of specifying the error structure treats all goods in a symmetric manner.

In order to proceed to the estimation of parameters, let us consider the general assumptions on the distributions of the $\boldsymbol{u}$'s;

$$(12)\ \ E(u_{it}) = 0$$

$$(13)\ \ E(u_{it},\ u_{jt+T}) = \textstyle\sum_t, \ \text{when}\ \ T=0$$
$$\qquad\qquad = 0, \quad \text{when}\ \ T \neq 0$$

(14) The $\boldsymbol{u}$'s have a multivariate normal distribution.

The implications for the distribution of the $\boldsymbol{v}$'s of our assumptions on the distributions of $\boldsymbol{u}$'s are easily derived. Since the $\boldsymbol{v}$'s are linear combinations of the $\boldsymbol{u}$'s, they also have a multivariate normal distribution. The variance of the disturbance term in the i-th demand equation is independent of income, and it is not directly related to consumption of the i-th good, although it is inversely related to the i-th good. One would expect autocorrelation of the $\boldsymbol{v}$'s of a higher level of consumption of the i-th good yesterday is associated with a higher level of consumption of the i-th good today. But in the habit forming models which depend on lagged consumption, this relationship has already been taken into account. In all of these models, a higher level of v_{it-1} implies a higher level of q_{it-1} which in turn implies a higher level of b_{it} and q_{it}. Thus $\boldsymbol{v}$'s from different periods are mutually independent in the habit forming models which depend on consumption in the preiod. Finally, inter-correlation or (multi-)collinearity between relative prices and real income (explanatory variables) is a relative matter. The sampling error of an individual coefficient depends on both the inter-correlation with other explanatory factors and upon the over-all correlation of whole equation. Thus, the inter-correlation is not necessarily a problem unless it is high relative to the over-all degree of multiple correlation among all variables simultaneously.

For more simplicity, $E(u_{it},\ u_{jt})=\sum_t$ is assumed to be the following two kinds of heteroscedasticity.

$$(15)\quad \textstyle\sum_t=\text{Diag }(\sigma_1^2 q_{1t}^2,\cdots\cdots\cdots,\sigma_n^2 q_{nt}^2)$$

$$(16)\quad \textstyle\sum_t=\text{Diag }(\sigma_1^2 q_{1t},\cdots\cdots\cdots,\sigma_n^2 q_{nt}).$$

In this fashion, we can compare the relative efficiency of the related estimators between two kinds of assumed heteroscedasticity.

4. Estimation Method

The Klein-Rubin linear expenditure system was first estimated by Stone [6]. We described the structure of disturbance terms in previous section. We use the maximum likelihood method, which we believe to be most appropriate, for estimating parameters. The stochastic demand equations are written as

$$(17)\quad q = (I-\gamma p')b+\gamma\mu+v_t$$
$$= (I-\gamma p')\,b+\gamma\mu+(I-\gamma p')u_t$$
$$= (I-\gamma p')b+\gamma\mu+M_tu_t$$

where $M_t=(I-\gamma p')$. The disturbance vector of the stochastic demand equation is given by $v_t=(I-\gamma p')u_t=M_tu_t$. Herein, consider the following useful theorems:

Theorem 1[1] If any $n\times 1$ vector X is distributed according to $N(\mu,\sum)$, then $Z=DX$ is distributed according to $N(D\mu,\ D\sum D')$.

This theorem includes the cases where X may have a nonsingular or a singular distribution and D may be nonsingular or singular.

Theorem 2[2] Since u_t is assumed to be multivariate normal with covariance matrix $\sum_t$, then $v_t=M_tu_t$ is multivariate normal with covariance matrix $\Omega_t=M_t\sum_tM_t'$ and $W_t=\hat{p}_tv_t$ is multivariate normal with covariance matrix $S_t=\hat{p}_t\Omega_t\hat{p}_t$, where $\hat{p}_t=\text{Diag}(p_{it},\cdots\cdots,p_{nt})$. As both Ω_t and S_t are singular, the densities of v_t and w_t cannot be expressed directly in terms of Ω_t and S_t. Barten [7] has shown, however, that in this case the density of w_t (ignoring a factor of proportionality) is given by

$$f(w_t)=|S_t+ii'|^{-\frac{1}{2}}\ e^{-\frac{1}{2}w_t'[S_t+ii']^{-1}w_t}$$

where i is the $n\times 1$ column vector $(1,\cdots\cdots,1)$. Hence, the density of v_t (again ignoring a factor of proportionality) is given by

$$g(v_t)=|\Omega_t+\hat{p}_t^{-1}ii'\hat{p}_t^{-1}|^{-\frac{1}{2}}\ e^{-\frac{1}{2}v_t'(\Omega_t+p_t^{-1}ii'p_t^{-1})^{-1}v_t}$$

1) T.W. Anderson, "Introduction to Multivariate Analysis," New York, John Wiley and Sons, pp. 25-27.

2) This theorem comes form Pollak and Wales[8].

Applying the Theorem 1, $\boldsymbol{v}_t$ is a linear transformation of $\boldsymbol{u}_t$, with the covariance matrix, $\Omega_t = \boldsymbol{M}_t \sum_t \boldsymbol{M}_t'$. Since the matrix $\boldsymbol{M}_t$ is singular, as shown by the previous section, the covariance matrix Ω_t is also singular. However, there is the estimation theorem (Theorem 2) shown by Barten [7] with the singular matrix. Here, applying Theorem 2 to the estimation of parameters, the likelihood function of the sample $(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots\cdots, \boldsymbol{v}_T)$ is:

$$L(\boldsymbol{v}_1, \cdots\cdots, \boldsymbol{v}_T) = \prod_{t=1}^{T} g(\boldsymbol{v}_t)$$

$$= \prod_{t=1}^{T} |A|^{-\frac{1}{2}} e^{-\frac{1}{2}v_t' A^{-1} v_t}$$

where $\boldsymbol{A} = \Omega_t + p_t^{-1} \boldsymbol{i}\boldsymbol{i}' \boldsymbol{p}_t^{-1}$ and the logarithm form of the above likelihood function is the following:

$$L(\boldsymbol{v}_1, \cdots\cdots, \boldsymbol{v}_T) = \sum_{t=1}^{T} \log\ g(\boldsymbol{v}_t)$$

$$= \frac{1}{2} \sum_{t=1}^{T} (\log |\boldsymbol{A}| + \boldsymbol{v}_t' \boldsymbol{A}^{-1} \boldsymbol{v}_t).$$

Our purposes to maximize L with respect to $\boldsymbol{a}$, $\boldsymbol{b}$, and σ subject to the linear constraint, $\boldsymbol{i}'\boldsymbol{a} = 1$. Let $G = \sum_{t=1}^{T} (\log |\boldsymbol{A}| - \boldsymbol{v}_t' \boldsymbol{A}^{-1} \boldsymbol{v}_t)$; then, maximizing L is equivalent to minimizing $\boldsymbol{G}$. Let us solve this problem through the well-known Lagrange multiplier methcd. Introducing the Lagrange multiplier, λ,

$$\varphi = \sum_{t=1}^{T} (\log |\boldsymbol{A}|^{-\frac{1}{2}} + \boldsymbol{v}_t' \boldsymbol{A}^{-1} \boldsymbol{v}_t) - \lambda(\boldsymbol{i}'\boldsymbol{a} - 1),$$

where $\boldsymbol{v}_t = \boldsymbol{q} - (\boldsymbol{I} - \gamma \boldsymbol{p}')\boldsymbol{b} - \gamma \mu$ for time t. For the model which is given by $b_{it} = \alpha_i + \beta_i z_{it-1}$, where z_{it-1} is any predetermined variable, φ is minimized with respect to a_i's, α_i's, β_i's, and σ_i's. Then we can get the best linear unbiased estimators.

Remarks on Theoretical Background of the Restricted Estimator[1]

In the classical general model $\boldsymbol{y}_{(n\times 1)}=\boldsymbol{X}\beta+\boldsymbol{u}$ with the appropriate assumptions about the distribution of $\boldsymbol{u}$'s we get the estimator $\hat{\beta}=(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ by the maximum likelihood method or the least-squares method, and the covariance matrix of $\hat{\beta}$, $\mathrm{Var}(\hat{\beta})=\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$.

Let us consider the linear restriction model. Linear restrictions on parameters to be estimated may be expressed in the following form: $\gamma=\boldsymbol{R}\beta$ where γ is a known column vector of g being the number of restrictions, and $\boldsymbol{R}$ is a known matrix of order $g\times n$. We now require the estimated coefficient vector $\boldsymbol{b}$, to satisfy the restrictions, so we must choose $\boldsymbol{b}$ to minimize $(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{b})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{b})$ subject to the linear restrictions, $\boldsymbol{R}\boldsymbol{b}=\gamma$. Through the Lagrange multiplier, we get the result $\boldsymbol{b}=\hat{\beta}+(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}^{-1}(\gamma-\boldsymbol{R}\hat{\beta})]$, where $\hat{\beta}=(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$, and the covariance matrix of $\boldsymbol{b}$, $\mathrm{Var}(\boldsymbol{b})=\boldsymbol{A}-\boldsymbol{R}'(\boldsymbol{R}\boldsymbol{A}\boldsymbol{R}')^{-1}\boldsymbol{R}\boldsymbol{A}$, where $\boldsymbol{A}=\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$ being the covariance matrix of the unrestricted least-squares estimator. It is seen that the covariance matrix of is obtained from $\boldsymbol{A}$ by deucting a positive semi-definite matrix, $\boldsymbol{R}'(\boldsymbol{R}\boldsymbol{A}\boldsymbol{R}')^{-1}\boldsymbol{R}\boldsymbol{A}$. Therefore we can say that the restricted estimator $\boldsymbol{b}$ is more efficient than the unrestricted one, $\hat{\beta}$.

5. Conclusion

Pollak and Wales [8] estimated the parameter: $b_1,\cdots\cdots,b_n$; and $\sigma_1,\cdots\cdots,\sigma_n$; and $a_1,\cdots\cdots,a_{n-1}$ and the last a_n obtained by deducting $\sum_{i=1}^{n-1}a_i$ from 1, through the maximum likelihood method.

Theoretically we know that the restricted estimator obtained through the Lagrange multiplier is more efficient than the unrestricted one, when any linear restriction on parameters is given.

We have the estimation theorem shown by Barten [7], with its singular covariance matrix. Furthermore, we have the linear restriction on a_i. Now

1) H. Theil, *Econometric Forecasts and Policy*, Second Revised Edition, North-Holland, 1961, pp. 331-333.

we can use the Lagrange multiplier so as to maximize the associated likelihood function subject to that linear restriction.

SUMMARY

Over twenty-five years ago, Professor Klein and Rubin [1] presented the linear expenditure system. That system was first estimated by Stone [6]. Subsequently many investigators have estimated that system.

In this paper, many points of the error structure shown by Pollak and Wales [8] are referred to. Barten [7] presented an estimation theorem on a singular covariance matrix. In order to estimate parameters, we place an emphasis on the maximum likelihood method which we believe to be most appropriate. As we have one linear restriction on parameters to be estimated, we maximized the associated likelihood function subject to that linear restriction through the well-known Lagrange multiplier method.

This paper is organized in the following fashion: (1) a brief description on classical consumer theory, (2) a linear expenditure system and its constraint, (3) dynamic specification and stochastic specification, (4) estimation method, and (5) conclusion.

REFERENCES

[1] Klein, L. R, and Rubin, H. "A Constant Utility Index of the Cost of Living," *Review of Economic Studies*, Vol. 15(1947-48).

[2] Klein, L. R., *An Introduction to Econometrics*, Prentice-Hall, 1962.

[3] Geary, R. C., "A Note on 'A Constant Utility Index of the Cost of Living'," *Review of Economic Studies*, Vol. 17(1950-51).

[4] Samuelson, P. A., "Some Implications of Linearity," *Review of Economic Studies*, Vol. 15(1947-48).

[5] Samuelson, P. A., *Foundations of Economic Analysis,* Harvard University Press, 1947.

[6] Stone, R., "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of Britich Demand," *Economic Journal,* September, 1954.

[7] Barten, A. P., "Estimating Demand Equation," *Econometrica,* Vol. 36, April, (1968).

[8] Pollak, R. A., and T. J. Wales, "Estimation of the Linear Expenditure System," *Econometrica,* Vol. 37, October, (1969).

[9] Parks, R. W., "Systems of Demand Equactions: An Empirical Comparison of Alternative Functional Forms," *Econometrica,* Vol. 37, October, 1969.

[10] Brown, A. and Deaton, A., "Surveys in Applied Economics: Model of Consumer Behaviour," *Economic Jonrnal,* December, (1972).

[11] Brown M. and Heien, D. M., "The S Branch Utility Tree: A Generalization of the Linear Expenditure System," *Econometrica,* Vol. 40, July, (1972).

[12] Theil, H., *Econometric Forecasts and Policy,* Second Revised Edition, North-Holland, 1961.

[13] Anderson, T. W., *Introduction to Multivariate Statistical Analysis,* NewYork: John Wiley and Sons, 1958.

[14] Morrison, D. F., *Multivariate Statistical Methods,* New York: McGraw-Hill, 1967.