Econometrica, Vol. 59, No. 3 (May, 1991), 711–730

THE RANK OF DEMAND SYSTEMS: THEORY AND NONPARAMETRIC ESTIMATION[1]

BY ARTHUR LEWBEL

Gorman's (1981) concept of Engel curve "rank" is extended to apply to any demand system. Rank is shown to have implications for specification, separability, and aggregation of demands. A simple nonparametric test of rank using Engel curve data is described and applied to U.S. and U.K. consumer survey data. The test employs a new general method for testing the rank of estimated matrices. The results are used to assess theoretical and empirical aggregation error in representative consumer models, and to explain a representative consumer paradox.

KEYWORDS: Rank, Engel curves, demand systems, aggregation, separability, representative consumer, Gaussian elimination, nonparametric tests.

1. INTRODUCTION

DEFINE THE RANK M of any demand system to be the maximum dimension of the function space spanned by the Engel curves of the demand system. This definition of rank extends the definitions of rank in Gorman (1981) and Lewbel (1989a) to encompass all demand systems. This paper describes a simple nonparametric procedure for estimating the rank of observed demands using ordinary survey data, without price variation. The technique first uses budget share and total expenditure data to construct a matrix $\hat{Y}$ having the property that the rank of the expected value of $\hat{Y}$ equals M, then a test for the rank of any estimated matrix is applied. This test is described in the Appendix.

Many known theoretical results concerning aggregation, separability, and functional structure of demands will be shown to be implications of rank. The definition and nonparametric test of rank described in this paper provide a way of simultaneously unifying and empirically applying a largely theoretical literature on aggregation and functional form specification.

The nonparametric rank test will be illustrated with an application to U.S. and U.K. data. The results will be used to investigate the paradox of why representative consumer models fit aggregate consumption data reasonably well, even though empirical Engel curves do not possess the linearity that standard aggregation theory says is required for a good fit. This includes a theoretical and empirical analysis of the size of aggregation errors in representative consumer specifications.

The rank test is a prespecification test, providing information about the degree of separability, aggregate structure, and cost function structure that are

[1]Portions of this work were done while the author was a visitor at the MIT Sloan School. This research was supported in part by the National Science Foundation, through Grant SES-8712787. I would like to thank Jerry Hausman and Gregory Leonard for providing U.S. data, Richard Blundell, Paul Baker, the U.K. Department of Employment, and the Institute for Fiscal Studies for providing U.K. data, and Leonard Gill, Thomas Stoker, Jerry Hausman, Dale Jorgenson, Terence Gorman, Richard Blundell, Angus Deaton, Robert Porter, participants of Harvard-MIT, Econometric Society, and other seminars, and two anonymous referees for many helpful comments and suggestions.

consistent with a given data set. It is standard practice to assume separability, aggregability, and a parametric or semiparametric functional form before doing any empirical analysis. The rank test thus provides information on precisely these aspects of demands that are either weakly tested or not tested in most empirical demand analyses.

Very little empirical work on rank exists. Hausman, Newey, and Powell (1988) estimate rank within the context of polynomial Engel curves, though they examine a few other related functional forms as well. Using U.S. consumer expenditure survey data, they find evidence for a rank of three. Although Leser's (1963) examination of different parametric Engel curve models is often cited to rationalize Engel curves that are linear in log income (a rank two specification), Leser in fact found a significantly better fit by regressing budget shares on more functions of income, which is indicative of a rank greater than two. Hardle and Jerison (1986) obtain kernel estimates of Engel curves and derive some rank implications of their results.

While this paper exclusively analyzes consumer demand, all the techniques described here could be applied to factor demands in a production context. The results would provide information on the empirical reasonableness of industry aggregate models, and of common assumptions about the separability of factors.

2. THE THEORY OF RANK CONDITIONS

Let $z = C(u, R)$ be the log of the cost function describing any demand system for N goods, where u is utility, R is the N vector of the log of prices of goods, x is total expenditures (income or cost for short), and $z = \ln x$ is log income. All that is being analyzed is the allocation of consumption expenditures within a period, so I will follow the common practice of using the term income to refer to x, even though formally x is total consumption expenditures. Let $w = d(z, R)$ be the N vector of budget shares derived from the cost function C. Standard cost function continuity and differentiability is assumed. Let $\zeta(R)$ denote the space spanned by $d(z, R)$ for a given R over all real z. Define the rank of the demand system local to R to be $\operatorname{rank}[\zeta(R)]$, and define the (global) rank M of the demand system by

$$M = \sup\{\operatorname{rank}[\zeta(R)] \mid R \text{ real}\}. \tag{2.1}$$

By the definition of rank,

$$w_i = d_i(z, R) = \sum_{m=1}^{M} a_{mi}(R) g_m(z, R) \qquad (i = 1, \ldots, N), \tag{2.2}$$

or $w = a(R)g(z, R)$, where $a(R)$ is an N by M matrix that is rank M for some R, and the M vector $g(z, R)$ is a basis for $\zeta(R)$. Note that all the budget shares are linear in the same M functions $g(z, R)$. By construction, any system d can be written in the form of equation (2.2) with $M \leqslant N$, and by definition, the rank of d is the minimum M for which (2.2) can be satisfied for all R and z. The space spanned by g includes the constant function, because budget shares sum to one.

For a given price regime R^*, budget share Engel curves are

$$w_i = \sum_{m=1}^{M} A_{mi} G_m(z) \qquad (i = 1, \ldots, N), \tag{2.3}$$

or simply $w = AG(z)$, where $A_{mi} = a_{mi}(R^*)$ and $G_m(z) = g_m(z, R^*)$. These Engel curves are linear in M functions of income, though the precise form of the income functions G typically depends on the prevailing price regime R^*. Note $\operatorname{rank}(A) = \operatorname{rank}[\zeta(R^*)]$.

THEOREM 1: *A demand system has rank M if and only if M is the smallest integer such that the cost function is of the form*

$$C(u, R) = H(u, \theta_1(R), \ldots, \theta_M(R)) \tag{2.4}$$

for some functions $\theta_1, \ldots, \theta_M$, and H.

PROOF: To see that equation (2.4) implies rank M, apply Shepard's Lemma to Equation (2.4) to obtain

$$w_i = \frac{\partial C}{\partial R_i} = \sum_{m=1}^{M} \frac{\partial \theta_m}{\partial R_i} \frac{\partial H}{\partial \theta_m}. \tag{2.5}$$

Let g_m for $m = 1, \ldots, M$ be given by

$$g_m(z, R) = \partial H[v(z, R), \theta_1(R), \ldots, \theta_M(R)] / \partial \theta_m(R) \tag{2.6}$$

where $v(z, R)$ is the indirect utility function of these demands (v is the inverse function of C with respect to u), and let a_{mi} for $m = 1, \ldots, M$ and $i = 1, \ldots, N$ be given by

$$a_{mi}(R) = \partial \theta_m(R) / \partial R_i. \tag{2.7}$$

Substituting equations (2.7) and (2.6) into (2.5) yields equation (2.2).

Now assume demands given by equation (2.2). To show that rank M implies equation (2.4), we must find M functions $\theta_1(R), \ldots, \theta_M(R)$ such that $\partial C / \partial R$ is collinear with $\partial[\theta_1(R), \ldots, \theta_M(R)] / \partial R$ for all R. By Shepard's Lemma

$$\partial C(u, R) / \partial R_i = \sum_{m=1}^{M} a_{mi}(R) f_m(u, R) \tag{2.8}$$

where for $m = 1, \ldots, M$, f_m is defined by

$$f_m(u, R) = g_m[C(u, R), R]. \tag{2.9}$$

Choose utility levels $u_1, \ldots, u_M$ such that the matrix with elements given by $f_m(u_\lambda, R)$ for $m = 1, \ldots, M$ and $\lambda = 1, \ldots, M$ is nonsingular at some value of R for which the local rank is M. Some such set of utility levels must exist, or else the space spanned by f, and hence the space spanned by g, can be spanned with fewer than M functions, which contradicts the assumption that the rank is M. (Note that $z = C(u, R)$ is monotonic in u.)

Define $\theta_m(R)$ for $m = 1, \ldots, M$ by

$$\theta_m(R) = C(u_m, R). \tag{2.10}$$

By equations (2.8) and (2.10), we have for $m = 1, \ldots, M$ and $i = 1, \ldots, N$ that

$$\partial\theta_m(R)/\partial R_i = \sum_{\lambda=1}^{M} a_{mi}(R) f_m(u_\lambda, R). \tag{2.11}$$

For each i from 1 to N, stack equation (2.11) for $m = 1, \ldots, M$ and solve for the vector $(a_{1i}(R), \ldots, a_{Mi}(R))$ in terms of the vector $(\partial\theta_1(R)/\partial R_i, \ldots, \partial\theta_M(R)/\partial R_i)$ by inverting the matrix of functions $f_m(u_\lambda, R)$. Substitute the result into equation (2.8) to get

$$\partial C(u, R)/\partial R_i = \sum_{m=1}^{M} h_m(u, R)\partial\theta_m(R)/\partial R_i \qquad (i = 1, \ldots, N) \tag{2.12}$$

for some functions $h_m(u, R)$, $m = 1, \ldots, M$. This shows that $\partial C/\partial R$ and $\partial[\theta_1(R), \ldots, \theta_M(R)]/\partial R$ have the necessary collinearity to ensure the existence of a function H satisfying equation (2.4). *Q.E.D.*

Theorem 1 says that rank M demands have cost functions that can be written in terms of M price indices, $\theta_1, \ldots, \theta_M$. Since cost functions are linearly homogeneous in prices, by the proof of Theorem 1 each function $\theta_m(R)$ equals the log of a homogeneous of degree one function of prices, and so can be interpreted as the expenditures on some composite commodity produced under constant returns. The function H is then interpretable as the log cost function for the composite commodities (the budget share of each composite commodity m is given by $\partial H/\partial\theta_m$), though in general H need not satisfy concavity conditions in terms of these composites.

3. IMPLICATIONS OF RANK

This section describes many implications of rank.

1. A demand system has rank $M = 1$ if and only if the demands are homothetic, that is, budget shares independent of the level of income. If demands are of rank one, then by equation (2.3) the Engel curves in any price regime are $w_i = A_{1i}G_1(z)$. Because budget shares sum to one for all z, we require that $G_1(z)$ not depend on z.

2. A demand system has rank $M = 2$ if and only if the demands are generalized linear (GL; see Muellbauer (1975)), since the definition of GL requires that $C(u, R) = H(u, \theta_1(R), \theta_2(R))$. Also, the indirect subutility functions in Gorman polar form (see Gorman (1959)) demand systems have rank $M = 2$. By definition, Gorman polar form indirect subutility functions are $h^*(\alpha(R) + \beta(R)x)$, which only depend on two functions of prices.

GL is a necessary and sufficient condition for aggregate demands to resemble representative agent models in certain ways (see Muellbauer (1975)), and is necessary for aggregate demands to exhibit the weak axiom of revealed preference (see Freixas and Mas-Colell (1987), who refer to GL as the "no torsion" condition). The popular AIDS, translog, and linear expenditure systems are all rank two models. More generally, rank two demands include the PIGL (including quasihomothetic), PIGLOG (see Muellbauer (1975)) and fractional demand systems (Lewbel (1987a)). See also Lewbel (1987b, 1990). Virtually all these classes of demands have been proposed in part because of their exact aggregation or representative agent properties.

3. Consider the "exactly aggregable" class of demands

$$w_i = \sum_{k=1}^{K} b_{ki}(R)G_k^*(z) \qquad (i=1,\dots,N) \tag{3.1}$$

for some functions $b_{ki}(R)$ and G_k^*. (Studies of exactly aggregable demands include Muellbauer (1975), (1976), Lau (1977a, 1977b), Jorgenson, Lau, and Stoker (1982), Lewbel (1987b, 1988), and Gorman (1981).) Let $b(R)$ be the N by K matrix of functions $b_{ki}(R)$. Gorman (1981) proved that for all equation (3.1) demands, the maximum possible rank of $b(R)$ is three. This implies that $M \leqslant 3$, because we may define $a(R)$ as a matrix whose columns form a basis for the columns of $b(R)$, and write equation (3.1) in (2.2) form with $g_m(R,z) = \sum_{k=1}^{K} b_{km}^*(R)G^*(z)$ for some functions b_{km}^*. An example of a rank three class of demands that is not in equation (3.1) form, but is still convenient for aggregation, is given in Lewbel (1989d).

4. Consider demands in the "deflated income" class

$$w_i = \sum_{k=1}^{K} b_{ki}(R)G_k^*(z-\theta^*(R)) \qquad (i=1,\dots,N) \tag{3.2}$$

for some functions $b_{ki}(R)$, G_k^*, and $\theta^*(R)$. These demands are analogous to those of equation (3.1), the only difference being that (3.2) demands are linear in functions of deflated income $x/\alpha(R)$ instead of nominal income x (here $z-\theta^*(R) = \ln[x/\alpha(R)]$). The usefulness of deflated income models is that they permit far more general Engel curves than equation (3.1) type demands, yet can still be exactly aggregated across agents using data on the distribution of deflated income instead of nominal income. Lewbel (1989a) proved that for all demands in the form of equation (3.2), the maximum rank of $b(R)$ is four. By the reasoning of point 3 above, $M \leqslant 4$.

5. When preferences are identical across agents, the results of Stoker (1984) imply that aggregate demands in an economy depend on M statistics of the income distribution of agents if and only if demands are in the form of equation (2.2), and aggregate demands satisfy generalized Slutsky symmetry (see Diewert (1977)) if and only if demands are in the form of equation (2.2) with $M \leqslant N$. Of course, equation (2.2) is equivalent to rank equaling M.

In general, the above results show that the lower the rank, the greater is the degree of utility related structure possessed by aggregate demands. It is therefore important to know the rank of demands to specify aggregate demand equations appropriately.

Now consider the implications of rank for utility production models and separability. Household production models either directly posit cost functions of the form

$$(3.3) \qquad C(u,R)=H(u,\phi_1(R),\dots,\phi_L(R))$$

for some L, or assume that direct utility can be written in terms of L homogeneous functions of goods, from which equation (3.3) follows directly. Here L is the number of "produced" intermediate goods from which agents derive utility, and which are in turn constructed from the raw, purchased goods.

Demands that are homothetically separable (either weakly or strongly) into L groups of goods also have cost functions given by equation (3.3), where in this case the price vector of purchased goods is partitioned among the ϕ functions.

Cost functions in the form of equation (3.3) may also arise from household welfare functions, in which ϕ_i is the cost of attaining a given level of utility for household member i, and H is dual to a welfare function that combines the utility of each household member.

Since the rank M is the minimum number of price indices in terms of which cost functions can be written, M cannot exceed L. Therefore, rank places a lower bound on the degree of separability and household production function structure.

Rank has additional implications that are summarized elsewhere. For example, implications of rank for welfare comparisons and equivalence scales are described in Lewbel (1991, 1989e). Many popular functional forms for directly additive utility functions are low rank (see Lewbel (1989c)). Finally, in demands for risky assets, the degree of portfolio separation (also called mutual fund separation, e.g., Cass and Stiglitz (1970)) equals the rank of the demand system, and in fact the definition of mutual fund separation is almost synonymous with rank (see Lewbel (1989b)). For example, the mean-variance utility functions used to derive the standard Sharpe-Lintner Capital Asset Pricing Model are rank two.

4. ESTIMATING THE RANK OF ENGEL CURVES: THEORY

Let $t=1,\dots,T$ index agents observed in a single price regime. Assume for now that agents have identical preferences, up to a noise component. (This assumption will be relaxed later.) In matrix form, equation (2.3) with an error added is

$$(4.1) \qquad w_t=AG(z_t)+e_t$$

where e_t is an N vector of mean zero errors that are assumed to be independent of z. No distributional assumptions concerning e will be made except that $e'\iota = 0$ (where $'$ denotes transpose and ι is the N vector of ones), which is required for $w'\iota = 1$. In particular, no restrictions are placed on the rank of $E(ee')$ except for those implied by $e'\iota = 0$.

Without functional form assumptions $G(z)$ is unknown. The structure of equation (4.1) therefore resembles that of factor analysis, where G is the set of unknown factors that "explain" w, and the rank M equals the unknown number of factors. Unfortunately, factor analysis requires restrictions on $E(ee')$ that are unreasonable to impose on equation (4.1). In particular, the usual factor analysis assumption that $E(ee')$ be diagonal cannot hold in equation (4.1), because $e'\iota = 0$ requires the existence of nonzero covariances. Therefore, tests for the number of factors in factor analysis (e.g., the likelihood ratio test described by Lawley and Maxwell (1971)) cannot be used to identify M.

To estimate M nonparametrically, the unique structure of the present application can be exploited. In particular, the information that the "factors" $G(z)$ are all functions of z can be used. Let $Q(z)$ be a vector of N or more functions having finite means. Postmultiply equation (4.1) by $Q(z_t)'$ to get

$$w_t Q(z_t)' = AG(z_t)Q(z_t)' + e_t Q(z_t)'. \tag{4.2}$$

Denote expected value across agents by E, and let $Y = E[wQ(z)']$, which is assumed to exist. Now, $E[eQ(z)'] = 0$ because e is independent of z. Therefore, $Y = AE[G(z)Q(z)']$, so $\operatorname{rank}(Y) = M$, unless by coincidence some component of G is orthogonal to all the elements of Q, or the price regime locally has rank less than M. Either case would then yield $\operatorname{rank}(Y) < M$.

Let $\hat{Y} = \Sigma_t[w_t Q(z_t)']/T$. By ordinary central limit theorems (assuming the existence of some moments), $\hat{Y}$ is a consistent and "root-T" asymptotically normal estimate of the corresponding population cross product matrix Y. If G were known, equation (4.2) would yield an instrumental variables estimate of A. Here G and A are unknown and not estimated. Instead, we estimate $\operatorname{rank}(Y)$ to obtain an estimate of $\operatorname{rank}(A)$.

The problem has been reduced to estimating $\operatorname{rank}(Y)$ given $\hat{Y}$, but this is difficult. For any finite T, $\operatorname{rank}(\hat{Y})$ will generally equal N, since it includes the error component. The obvious approach is to estimate sample eigenvalues and their asymptotic distribution, the estimate of rank being the number of eigenvalues of $\hat{Y}$ that are significantly different from zero. Unfortunately, the asymptotic variances of eigenvalues depends in a nontrivial way on $\operatorname{rank}(Y)$ itself, and on the unknown multiple eigenvalue structure of Y (see, e.g., Anderson (1963) and Magnus (1984)). Also, eigenvalues can be complex valued, which is a nuisance, and eigenvalues are highly nonlinear functions of the data, which can cause computational problems.

The alternative strategy proposed here is to consider the Gaussian elimination based LDU decomposition. The rank of any matrix equals the number of nonzero elements of the diagonal matrix D in this decomposition. Here again the asymptotic distribution of D will depend on rank(Y), but the explicit calculation of the distribution for alternative possible ranks is rendered tractable by the fact that, unlike eigenvalues, D is a simple function of Y. This is true even when Y is nonsquare and asymmetric.

Gill and Lewbel (1988) derive an explicit expression for the asymptotic covariance matrix of D for arbitrary random matrices, and use it to construct a chi-squared test of the hypothesis that sets of elements of D are zero, which is identical to a test of rank. This test is summarized in the Appendix.

5. HETEROGENEITY OF PREFERENCES

The nonparametric rank estimator described in Section 4 assumed agents have identical preferences up to an additive error. However, the procedure remains consistent with heterogeneous preferences, as long as each agent's Engel curves span the same space $\zeta(R)$ and the distribution of preferences is independent of the distribution of income (total expenditures). The first assumption requires that all agents have budget shares that are linear in the same functions of income, as is almost always assumed in practice. In contrast, the second assumption is not likely to hold for all agents (e.g., preferences and income may both be related to household size). However, it is not an unreasonable assumption to make about collections of households that are broadly similar demographically, especially if demands for consumption goods are assumed to be separable from the labor-leisure choice.

To see why the rank test remains consistent under these assumptions, observe that if preferences vary across households while $\zeta(R)$ does not, then equation (4.1) is replaced by $w_t = A_t G(z_t) + e_t$. If the distribution of A across agents is independent of z_t, then $Y = E[wQ(z)'] = AE[G(z)Q(z)']$, where now $A = E[A_t]$. Since each A_t has full column rank M, rank$(A) = M$ unless there is a coincidence across preferences that makes rank$(A) < M$. As before, rank(Y) is a lower bound for M, and will typically equal M.

Instead of controlling for preference heterogeneity by selecting samples for which the above assumptions may reasonably hold, one could instead define an extended notion of rank as the dimension of the space spanned by demands as functions of demographic attributes as well as income. This "extended" rank of demands equals the number of aggregate statistics, or indices, in the classes of demands analyzed by Lau (1982) and Section 1 of Jorgenson, Lau, and Stoker (1982), and is relevant to Hardle and Jerison (1988). Stoker's (1984) theorem can be expressed as, "aggregate demands have generalized Slutsky symmetry of order M if and only if agents' demands have extended rank M." The usefulness of this extended notion of rank is limited by the fact that none of the other attributes of rank described in Section 3 carry over to extended rank, and because extended rank will often equal N. For example, any demand system having unrestricted Barten style commodity specific equivalence scales will have extended rank equal to N.

6. ESTIMATING THE RANK OF ENGEL CURVES: EMPIRICAL RESULTS

The nonparametric rank estimation method described in the previous sections was applied to individual household expenditures data from the 1970 to 1984 United Kingdom Department of Employment Family Expenditure Surveys (FES) and the 1982 United States Bureau of Labor Statistics Consumer Expenditure Survey (CES).

Since the rank estimator applies to cross section data (with no price variation), the FES data were analyzed separately for each year from 1970 to 1984. To minimize income correlated demographic variation, the selection criteria were: married couples with two children, head of household working full time, and no relevant variables missing. The resulting sample sizes range from 216 to 340 households per year. Total consumption expenditures of the households were divided into eight groups of goods: food (including alcohol and tobacco), housing, fuel and light, durables, clothing, transportation and vehicles, services, and other.

For the larger U.S. CES, the data analyzed were from the first quarter of 1982, with the selection criteria being married couples, age of head between 30 and 60 years old, nominal annual income between \$2000 and \$55000, and no relevant variables missing. A total of 881 households met these criteria. In the U.S. data, total consumption expenditures of the selected households were divided into the seven groups: food, clothing, recreation, furnishings, health care, transportation, and other.

In both data sets total expenditures x equal the sum of expenditures in the groups of goods. Budget shares w are group expenditures divided by x.

By definition, the budget share for each group of goods is the sum of the budget shares of each raw good comprising the group. Thus the rows of A in the "grouped goods" model are sums of rows in the underlying model of individual goods. This summing may result in an underestimate of the true rank, since the rank of A in the grouped goods model is at most equal to the rank of the corresponding matrix of Engel curves for the raw goods. The rank of A will not be reduced if demands happen to be separable in the chosen groups.

The "instrument" list Q consists of the functions 1, x, z, x^2, z^2, $1/x$, $1/x^2$, and xz, where $z = \ln x$. The error in the rank estimation procedure is minimized if Q contains the actual G functions, so the above list was chosen to include the functions of income contained in commonly used Engel curve specifications. Each element of Q was divided by its mean in each sample, to ensure that Y was not ill conditioned as a result of the enormous range of magnitudes in the functions comprising Q.

By the results in Section 4 and 5, $\text{Rank}(E(\hat{Y})) \leqslant M$, and typically will equal M. Table 1 gives estimated pivots (diagonal elements of D in the LDU decomposition) of $\hat{Y}$ for the CES and for each year of the FES. Table II summarizes the results of applying the LDU rank test to the estimated pivots.

For a few of the years of FES data, the results in Table II are ambiguous about whether the rank is two, three, or four. However, the U.S. data and eleven out of fifteen years of the U.K. data show a less than 1% probability that

TABLE I

ESTIMATED PIVOTS

DATA		d_1	d_2	d_3	d_4	d_5
CES	'82	0.326	0.069	0.045	−0.001	0.000
FES	'70	0.377	0.241	0.028	−0.006	−0.001
	'71	0.369	0.090	−0.019	0.000	0.000
	'72	0.370	0.253	0.063	−0.012	0.001
	'73	0.342	0.056	0.023	0.000	−0.000
	'74	0.327	0.077	−0.008	−0.002	−0.000
	'75	0.346	0.092	0.020	−0.000	0.000
	'76	0.340	0.086	0.017	0.001	0.000
	'77	0.338	0.067	0.012	−0.000	−0.000
	'78	0.333	0.083	0.016	−0.001	−0.000
	'79	0.344	0.074	0.013	−0.002	0.001
	'80	0.324	0.138	0.021	−0.003	0.000
	'81	0.311	0.122	0.023	−0.002	−0.000
	'82	0.295	0.104	0.020	0.002	0.000
	'83	0.306	0.095	0.020	0.002	0.000
	'84	0.310	0.127	0.029	−0.003	0.000

Each row shows the five largest pivots, in descending order, for the U.S. CES data and each year of the U.K. FES data described in the text. All remaining pivots were zero to at least three decimal places.

TABLE II

LDU RANK TEST CHI-SQUARED STATISTICS

DATA		$r=1$	$r=2$	$r=3$	$r=4$
CES	'82	194.733	190.503	0.123	0.000
FES	'70	2586.954	19.171	0.426	0.025
	'71	518.773	28.736	0.006	0.000
	'72	4054.107	469.567	3.418	0.000
	'73	429.056	12.248	0.010	0.001
	'74	254.524	2.320	1.390	0.000
	'75	558.246	127.291	0.027	0.000
	'76	436.750	3.992	0.270	0.000
	'77	210.276	27.338	0.011	0.000
	'78	168.944	87.886	0.096	0.000
	'79	124.310	2.081	0.086	0.043
	'80	1240.903	99.376	0.716	0.000
	'81	1233.809	179.325	0.063	0.000
	'82	947.043	107.723	0.143	0.000
	'83	564.012	157.294	0.400	0.006
	'84	1622.098	158.799	0.341	0.000

Each row shows LDU rank test χ^2 statistics for the U.S. CES and each year of the U.K. FES data. In each column r denotes the rank being tested. The test is that all except the r largest pivots are zero, so each test is consistent only against alternatives that the rank is greater than r. The degrees of freedom of the statistics are $7-r$ for the CES and $8-r$ for the FES data. For every $r>4$ in every year, the test statistics were zero to five or more decimal places.

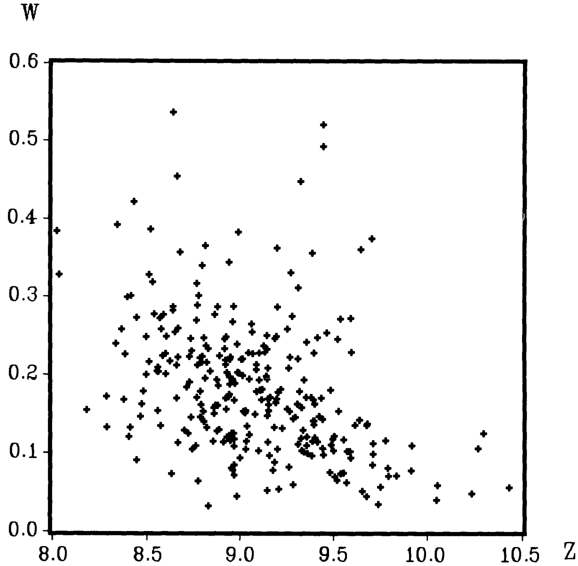

FIGURE 1.—Scatter plot of 1977 FES housing share vs. log income.

all but two pivots are zero (χ^2 statistic greater than 17 in $r = 2$ column) and a greater than 99% probability that all but three pivots are zero (χ^2 statistic less than .5 in $r = 3$ column). These results indicate, with high precision, a rank of three. The rank tests are sequential within each sample, and the test for each r is only consistent against greater values of r, but the magnitudes of the χ^2 statistics drop dramatically from $r = 2$ to $r = 3$. Therefore, any sensible procedure for adjusting the size or significance levels of the tests to account for their sequential nature would not change the conclusion that $\operatorname{rank}(Y) = 3$, and hence that M is probably three but possibly higher.

Two caveats are necessary. First, no adjustments for savings and depreciation have been made for the inclusions of durables and semidurables. Second, the results reported are only for the two selected demographic groups, and it is possible that the rank for other households may be different.

7. THE STRUCTURE OF ENGEL CURVES

This section uses both rank and kernel regressions to further analyze the structure of Engel curves. Begin with an illustrative example. Figure 1 shows a scatter plot of the budget share w_i for housing as a function of z (log total expenditures) for 1977, the middle year of the FES sample. Figure 2 shows a kernel regression of this data. The kernel is biquartic, with a bandwidth of 1.5. Given some smoothness assumptions, this kernel regression consistently estimates the true Engel curve, whatever the actual functional form is (see, e.g., Prakasa Rao (1983)). The striking feature of Figure 2 is that w_i is close to linear in z, except at the lowest and highest expenditure level households. An OLS

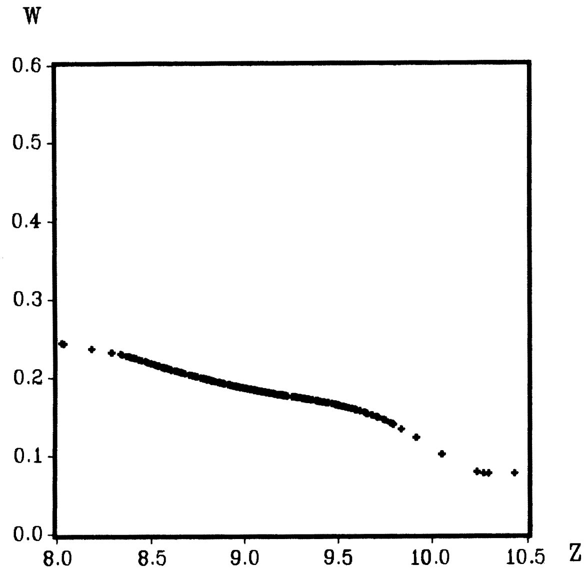

FIGURE 2.—Kernel regression of 1977 FES housing share vs. log income.

regression of w_i on powers of z gives similar results. If budget shares for the nonextreme households are actually linear in z, then for these households the derivative $\partial w_i/\partial z$ should equal the coefficient of z in an OLS regression of w_i on z. Using kernel regressions, the average derivative of w_i with respect to z can be estimated nonparametrically with an accuracy that improves with sample size at a parametric root-T rate (see Hardle and Stoker (1989), and Stoker (1990)). Table III shows that average derivative estimates are in fact quite close to OLS slopes for all the budget shares in 1977 when households in the top and

TABLE III

1977 "TRIMMED DATA" ENGEL CURVE OLS AND KERNEL AVERAGE DERIVATIVES

	Average Derivative	OLS Slope		Average Derivative	OLS Slope
Food	−.0320	−.0334	Clothing	.0199	.0162
	(.0068)	(.0061)		(.0050)	(.0045)
Housing	−.0243	−.0223	Transport	.0220	.0235
	(.0063)	(.0052)		(.0065)	(.0066)
Fuel	−.0145	−.0144	Other goods	.0057	.0044
	(.0020)	(.0021)		(.0035)	(.0032)
Durables	.0157	.0202	Services	.0076	.0061
	(.0046)	(.0048)		(.0049)	(.0042)

This table gives estimates of the derivative of budget shares with respect to z, based on 1977 FES data where the highest 5% and lowest 5% expenditure households are omitted. Average derivatives are kernel based "indirect slope" estimates (see Stoker (1990)). OLS estimates are the estimated slope coefficients in an OLS regression of the budget share on z. Standard errors are in parentheses.

TABLE IV

"TRIMMED DATA" LDU RANK TEST CHI-SQUARED STATISTICS

DATA		$r=1$	$r=2$	$r=3$
FES	'70	69.892	20.832	0.001
	'71	309.173	0.052	0.000
	'72	547.963	1.174	0.000
	'73	281.660	1.607	0.001
	'74	115.245	0.136	0.001
	'75	263.763	1.022	0.016
	'76	61.490	2.890	0.001
	'77	125.010	0.067	0.000
	'78	188.125	1.464	0.002
	'79	97.041	0.178	0.002
	'80	248.386	0.590	0.004
	'81	258.635	3.002	0.000
	'82	780.245	1.024	0.006
	'83	285.336	6.533	0.030
	'84	659.701	14.868	0.002

Each row shows LDU rank test χ^2 statistics for each year of the FES data where the highest 5% and lowest 5% expenditure households in each year are omitted. r denotes the rank tested. The test is that all except the r largest pivots are zero, so each test is consistent only against alternatives that the rank is greater than r. The degrees of freedom of the statistics are $8-r$. For every $r>3$ in every year, the test statistics were zero to over four decimal places.

bottom five percent of x values are omitted, formally confirming the near linearity apparent in Figure 2.

Demands having budget shares linear in z are called PIGLOG (see Muellbauer (1976)), and are rank two. While the tails of kernel regressions are often imprecisely estimated, the previous finding that rank equals three implies that the nonlinearities in the tails are significant. The kernel based results suggest that demands are rank two when the lowest and highest expenditure level households are dropped from the sample.

To check this, in each year the households in the top and bottom five percent of the expenditure distribution were dropped from the sample, and the rank tests were computed for the remaining ninety percent. The results, reported in Table IV, show that in almost all the years the rank drops to two in this "trimmed" sample.

These results suggest that for most households demands are reasonably modeled as rank two in general and PIGLOG in particular, but that a more complicated (i.e., rank three) model is required when households with very low or high expenditures are included in the sample.

8. EXPLAINING A REPRESENTATIVE CONSUMER PARADOX

An empirical paradox is that estimated aggregate macroeconomic demands resemble those of a utility maximizing representative consumer, yet cross sectional Engel curves are nonlinear in x. Standard aggregation theory says that Engel curves must be quasihomothetic (i.e., linear in x) for aggregate demands to match those of a utility maximizing consumer (see, e.g., Gorman (1953)). The size of the aggregation error in representative consumer models is important, because most empirical analyses of aggregate demands assume a representative consumer, and a great deal of theoretical work in economics assumes a representative consumer for tractability.

Some aggregate empirical studies (e.g., Gallant (1981), and Christensen, Jorgenson, and Lau (1975)) reject representative consumer attributes like homogeneity and Slutsky symmetry, although the rejections are not as economically significant as would be expected from aggregation theory. In contrast, the nonparametric methods employed by Varian (1983), Manser and McDonald (1988), and Diewert and Parkan (1978) do not reject the restrictions implied by a utility maximizing representative consumer.

The rank three results obtained above are consistent with many other empirical studies which find that, cross-sectionally, Engel curves are far from quasihomothetic (i.e., linear in x). Even when low and high expenditure consumers are dropped, budget shares are approximately PIGLOG (i.e., linear in z, not x).

Let x_{ht} equal household h's total consumption expenditures at time t, R_t be the vector of log prices at time t, and w_{ht} be the household h budget shares. PIGLOG demands are defined as $w_{ht} = a(R_t) + b(R_t)\ln x_{ht}$, for N vector valued functions a and b.

Lemma: *If PIGLOG demands $w = a(R) + b(R)\ln x$ satisfy the equality constraints implied by utility maximization (adding up, homogeneity, and Slutsky symmetry), then $w = a(R) + b(R)(k + \ln x)$ also satisfies these constraints for any constant k.*

Proof: $w = a(R) + b(R)\ln x$ possesses Slutsky symmetry if and only if $w = [\partial A(R)/\partial R] + [\partial B(R)/\partial R][A(R) + \ln x]$ for differentiable scalar valued functions $A(R)$ and $B(R)$ (see Muellbauer (1976)). Letting $A^*(R) = k + A(R)$ then shows that $w = a(R) + b(R)(k + \ln x)$ also satisfies symmetry. The proofs of homogeneity and adding up are analogous. *Q.E.D.*

Results very similar to Lemma 1 were exploited by Deaton and Muellbauer (1980) in their rationalization of the aggregate AIDS model, which is PIGLOG.

Let $X_t = E_h(x_{ht})$, and $W_t = E_h(w_{ht}x_{ht})/X_t$, where the operator E_h denotes averaging across households. X_t is average (per household) total expenditures in the economy and W_t is the vector of aggregate budget shares in the economy. Define X_t^* by $\ln X_t^* = E_h(x_{ht}\ln x_{ht})/X_t - k$ for some constant k. Implicitly, X_t^* is a function of k. Summing PIGLOG demands weighted by x_{ht}/X_t over households h gives

$$(8.1) \qquad W_t = a(R_t) + b(R_t)(k + \ln X_t^*).$$

The rank and Kernel results from the previous section indicate that the majority of households have approximately PIGLOG preferences. If the presence of relatively few nonPIGLOG households is swamped by the majority, aggregate demands will roughly satisfy equation (8.1).

Assuming agents maximize utility, it follows from the Lemma and equation (8.1) that macroeconomic demands will resemble those of a utility maximizing representative consumer (satisfying adding up, homogeneity, and Slutsky symmetry) if the distribution of x is such that, over time, $X_t \approx X_t^*$ for any constant k. Using the FES data that generated Table II, for each year t, X_t and $E_h(x_{ht} \ln x_{ht})$ were constructed as the sample average over all households h of x_{ht} and $x_{ht} \ln x_{ht}$, respectively. For these fifteen annual observations the regression $E_h(x_{ht} \ln x_{ht})/X_t = \hat{k} + \ln X_t + e_t$ has $R^2 = .9983$, implying that the average aggregation error of using X_t in place of X_t^* is less than 0.2 percent.

This high correlation may be due to relatively slow changes in the relative distribution of per capita income over time, and the possibility that life-cycle or permanent income style consumption smoothing results in less variation in the distribution of total consumption expenditures than in income.

The empirical finding that most agents have PIGLOG demands and that the distribution of x across agents is such that $X_t \approx X_t^*$ for some k, combined with the above lemma, explains the representative consumer paradox. Taken together, these results imply that aggregate demands will resemble those of a PIGLOG representative consumer. Joint regularities in the functional form of household demands and the distribution of x overcome the standard aggregation requirement of quasihomotheticity for generating a representative consumer.

So far I have again ignored heterogeneity of preferences, but the above results can be readily extended to the case where $a(R)$ and $b(R)$ vary across agents, in which case aggregate demands will contain terms like $E_h[a_h(R_t)]$, $E_h[b_h(R_t)]$, and possibly covariances between a_h and x_{ht} and between b_h and $x_{ht} \ln x_{ht}$.

To check these results, estimates of a PIGLOG representative consumer model are compared with the estimates of a corresponding exactly aggregated model. Assume most households h have budget shares of the PIGLOG form

$$w_t = \frac{\alpha_h + \gamma R_t + \beta[\delta + \alpha_h' R_t + .5 R_t' \gamma R_t] - [\gamma\iota + \beta(1 + R_t'\gamma\iota)] \ln x_{ht}}{1 + R_t'\gamma\iota}, \tag{8.2}$$

where α_h and β are vectors, δ is a scalar, γ is a symmetric matrix, ι is a vector of ones, $\alpha_h'\iota = 1$, $\beta'\iota = 0$, and $\iota'\gamma\iota = 0$. Note that α_h permits commodity specific preference heterogeneity. These budget shares are derived from the utility function described in Lewbel (1989f). This model is PIGLOG and is the simplest possible model that has both Deaton and Muellbauer's (1980) Almost Ideal Demand System (AIDS) and Jorgenson, Lau, and Stoker's (1982) Transcendental Logarithmic (Translog) Model nested within it as special cases. (The Translog corresponds to $\gamma\iota = 0$ and the AIDS to $\beta = 0$.)

TABLE V

ESTIMATES OF EXACTLY AGGREGATED AND REPRESENTATIVE CONSUMER MODELS

	FES		U.S. NIPA	
	Model 1 Exact agg.	Model 2 Rep. con.	Model 3 Exact agg.	Model 4 Rep. con.
Number Parameters	10	10	15	12
Log Likelihood	139.91	137.43	399.17	390.18
Own Price Elasts				
Group 1	1.13	1.03	.51	.56
Group 2	.31	.53	.25	.25
Group 3	1.32	1.22	.66	.69
Income Elasts				
Group 1	.67	.80	.94	.89
Group 2	.73	.79	1.32	1.33
Group 3	.69	.99	.70	.67

Models 1 and 2 use U.K. data constructed by directly adding up the FES households analyzed in Tables I and II. For Models 1 and 2, the groups of commodities are: Group 1: Food, 2: Shelter, and 3: Other. Model 1 is exactly aggregated; Model 2 is a representative consumer.

Models 3 and 4 use aggregate U.S. data constructed from U.S. National Income and Product Accounts. For Models 3 and 4, the groups of commodities are: Group 1: Energy, 2: Nondurables, and 3: Services. Model 3 is identical to the exactly aggregated nested model in Lewbel (1989f), except for the addition of a time trend. Model 4 is the representative consumer equivalent.

Elasticities are mean uncompensated quantity demand elasticities.

Exactly aggregating equation (8.2) gives

$$W_t = \frac{\alpha_t + \gamma R_t + \beta[\delta + \alpha_t' R_t + .5 R_t' \gamma R_t] - [\gamma\iota + \beta(1 + R_t'\gamma\iota)] E(x_{ht} \ln x_{ht})/X_t}{1 + R_t'\gamma\iota} \tag{8.3}$$

where α_t is the mean in time t of α_h. Table V summarizes estimates of equation (8.3), and for comparison gives estimates of the representative consumer version of the same demands, in which the statistic $\ln X_t$ is used in place of the distribution statistic $E_h(x_{ht} \ln x_{ht})/X_t$. For the FES data, the distribution statistic $E_h(x_{ht} \ln x_{ht})/X_t$ was constructed as described earlier. For the aggregate U.S. consumption data, this statistic was approximated using income quantile data. (See Lewbel (1989f) for details.)

In models 1, 2, and 4, in Table V, α_t is modeled as $\alpha_0 + \alpha_1 t$ for constant vectors α_0 and α_1, on the assumption that most variation in α_h is demographic, and that changes in the distribution of demographic attributes across households are mostly time trends (note that the FES data is already selected to be relatively homogeneous demographically). The time trend may also crudely proxy for dynamic effects. In model 3, some demographic variation is explicitly included in α_t, as described in Lewbel (1989f).

In both the constructed FES aggregates and in the U.S. national consumption data, the exact aggregation models are found to have similar elasticity estimates and better fits (higher likelihood values) than the representative consumer models. These aggregate results are consistent with both the theory and the disaggregate data rank and kernel results of the previous sections.

When the same exercise was performed with the "trimmed" FES data set, in which the lowest and highest five percent expenditure level households were discarded before aggregating, the log-likelihoods were 141.66 and 141.72. As expected from the rank, kernel, and x distribution results, the fit of the aggregate PIGLOG model improved when the extreme households were dropped, and the difference between exact aggregation and the representative consumer is negligible in this case.

These analyses used budget shares, but quantities or expenditures can be used instead without changing the results. For example, if expenditures on each good are analyzed, then dividing each instrument by total expenditures will yield numerically identical rank results.

9. CONCLUSIONS

This paper showed that the concept of rank can be defined for any demand system, and that rank has numerous implications for separability, for functional form, and for aggregation across goods and across agents. A simple nonparametric technique using only total expenditure and budget share data was employed to estimate a matrix with the same rank as the demand system. The matrix rank test of Gill and Lewbel (1988) was applied to this estimated matrix to yield a nonparametric estimate of the rank of the demands of both U.S. and U.K. family expenditure survey data.

The nonparametric demand rank test is a useful prespecification tool, because the type of information revealed by rank (e.g., degree of separability, aggregate utility structure, and income flexibility) is typically assumed without testing in empirical demand work.

The results strongly indicate a rank of three in both U.S. and U.K. data sets. When households in the tails of the total expenditures distribution were removed from the data, the estimated rank dropped from three to two. Combining these rank results with kernel regressions and kernel average derivatives leads to the conclusion that budget shares are approximately linear in the log of total expenditures (PIGLOG) except for households in the tails, where significant nonlinearities exist. These results confirm and expand on previous parametric Engel curve analyses, which found good fits with PIGLOG models but also found improvements with rank three specifications.

There are several implications of these rank results. First, the popular AIDS and Translog models (which are PIGLOG) fit Engel curve data relatively well, but are inadequate for encompassing households in the tails. Second, at least three terms are likely to be required in semiparametric or other series expansion based models of demand (e.g., Elbadawi, Gallant, and Souza (1983) and Barnett and Yue (1988)). Third, the number of aggregable Engel curve models that possess sufficiently high rank to fit the data is relatively small (see Gorman (1981), and Lewbel (1989a, 1989d, 1990)). Finally, three should be a lower bound on the number of indirectly separable groups of goods. Most empirical studies assume at least three groups of goods, so this lower bound does not seriously constrain current empirical practice. Formally, all these results are limited to the particular demographic cells that were analyzed.

The structure of Engel curves uncovered by rank and kernel methods also contributed to the explanation of the paradox of a PIGLOG representative consumer given substantial Engel curve nonlinearities. The deviations of disaggregate demands from PIGLOG behavior are concentrated in a relatively small number of households (i.e., those having the lowest and highest consumption levels). The paradox is resolved by combining this finding with regularities in the distribution of expenditures and properties of PIGLOG demands.

There are many implications of these representative consumer results. First, representative consumer models fit well, but exact aggregation models fit better. Also, the representative consumer model that fits well is not homothetic, contradicting the typical macroeconomic representative consumer assumption. Second, demographic changes may cause larger aggregation errors in representative consumer models than nonlinearities in income effects, though the bulk of aggregation theory has focused on the latter. Third, the representative consumer's preferences differ from those of individual households. Fourth, while the magnitude of aggregation errors in correctly specified representative consumer models is small, in theory these errors interact with prices. This may help explain why representative consumer restrictions that have been statistically rejected with parametric models have not been rejected with alternative nonparametric techniques (see, e.g., Manser and McDonald (1988)). Finally, forecasts based on representative consumer models will be slightly inferior to those based on exact aggregation models, but only if the distribution regularities observed in the past are maintained in the future. In particular, policy implications drawn from representative consumer models may be badly biased if the contemplated policy affects the income distribution.

Department of Economics, Brandeis University, Waltham, MA 02254, U.S.A.

Manuscript received January, 1989; final revision received May, 1990.

APPENDIX: TESTS OF RANK USING THE LDU DECOMPOSITION

This appendix describes how to test the rank of a random matrix using the LDU (Gaussian elimination) decomposition. See Gill and Lewbel (1988) for details, and for applications including econometric identification and state space prespecification tests.

Consider an m by p matrix Y (where $m \geqslant p$) which has rank r, and let $\hat{Y}$ be a "root-T" asymptotically normal estimate of Y based on sample size T, so $\operatorname{vec}[\sqrt{T}(\hat{Y}-Y)] \stackrel{a}{\sim} N(0,V)$, for some covariance matrix V.

There exists a unique decomposition of the form $PYQ = LDU$, where P and Q are permutation matrices, L and U' are lower triangular, D is diagonal, and the elements of D decrease in magnitude along the diagonal. The corresponding decomposition for $\hat{Y}$ is $\hat{P}\hat{Y}\hat{Q} = \hat{L}\hat{D}\hat{U}$. Let

$$\hat{L}\hat{D}\hat{U} = \begin{bmatrix} \hat{L}_{11} & 0 & 0 \\ \hat{L}_{21} & \hat{L}_{22} & 0 \\ \hat{L}_{31} & \hat{L}_{32} & I_{m-p} \end{bmatrix} \begin{bmatrix} \hat{D}_1 & 0 & 0 \\ 0 & \hat{D}_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{U}_{11} & \hat{U}_{12} \\ 0 & \hat{U}_{22} \\ 0 & 0 \end{bmatrix}$$

where the row partition is by r, $p-r$, and $m-p$ rows, and the column partition of $\hat{L}$ and $\hat{D}$ is by r, $p-r$, $m-p$ columns, and for $\hat{U}$ is by r and $p-r$ columns. Let $\hat{d}_2 = \mathrm{Diag}(\hat{D}_2)$, and let Δ be the $(p-r)^2$ by $(p-r)$ matrix such that $\mathrm{vec}(\hat{D}_2) = \Delta \hat{d}_2$. Define the matrices

$$\hat{H} = \left[-\hat{L}_{22}^{-1}\hat{L}_{21}\hat{L}_{11}^{-1} : \hat{L}_{22}^{-1} : 0 \right],$$

$$\hat{K} = \begin{bmatrix} -\hat{U}_{11}^{-1}\hat{U}_{12}\hat{U}_{22}^{-1} \\ \hat{U}_{22}^{-1} \end{bmatrix},$$

$$\hat{W} = \Delta'(\hat{K}' \otimes \hat{H})(\hat{Q}' \otimes \hat{P})\hat{V}(\hat{Q} \otimes \hat{P}')(\hat{K} \otimes \hat{H}')\Delta.$$

Under the null that $\mathrm{rank}(Y) = r$, so $d_2 = 0$, construct the χ^2 test statistic

$$T\hat{d}_2'\hat{W}^{-1}\hat{d}_2 \overset{a}{\sim} \chi^2_{p-r}.$$

Since $d_2 = 0$ is equivalent to $\mathrm{rank}(Y) \leqslant r$, this test is consistent against the alternative that $\mathrm{rank}(Y) > r$. In empirical application the above statistic can be calculated for each possible rank.

REFERENCES

Anderson, T. W. (1963): "Asymptotic Theory for Principal Component Analysis," *Annals of Mathematical Statistics*, 34, 122–148.

Barnett, W. A., and P. Yue (1988): "Semiparametric Estimation of the Asymptotically Ideal Model: The AIM Demand System," Institute of Statistics and Decision Sciences Discussion Paper, Duke University.

Cass, D., and J. E. Stiglitz (1970): "The Structure of Investor Preferences and Asset Returns, and Separability in Portfolio Allocation: A Contribution to the Pure Theory of Mutual Funds," *Journal of Economic Theory*, 2, 122–160.

Christensen, L. R., D. W. Jorgenson, and L. J. Lau (1975): "Transcendental Logarithmic Utility Functions," *American Economic Review*, 65, 367–383.

Deaton, A. S., and J. Muellbauer (1980): "An Almost Ideal Demand System," *American Economic Review*, 70, 312–326.

Diewert, W. E. (1977): "Generalized Slutsky Conditions For Aggregate Demand Functions," *Journal of Economic Theory*, 15, 353–362.

Diewert, W. E., and C. Parkan (1978): "Tests For the Consistency of Consumer Data and Nonparametric Index Numbers," University of British Columbia Working Paper 78-27.

Elbadawi, I., A. R. Gallant, and G. Souza (1983): "An Elasticity Can be Estimated Consistently Without A Priori Knowledge of Functional Form," *Econometrica*, 51, 1731–1751.

Freixas, X., and A. Mas-Colell (1987): "Engel Curves Leading to the Weak Axiom in the Aggregate," *Econometrica*, 55, 515–532.

Gallant, A. R. (1981): "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form," *Journal of Econometrics*, 15, 211–245.

Gill, L., and A. Lewbel (1988): "Testing the Rank of Estimated Matrices," Brandeis University Working Paper.

Gorman, W. M. (1953): "Community Preference Fields," *Econometrica*, 21 63–80.

——— (1959): "Separable Utility and Aggregation," *Econometrica*, 27, 469–481.

——— (1981): "Some Engel Curves," in *Essays in the Theory and Measurement of Consumer Behaviour in Honor of Sir Richard Stone*, ed. by Angus Deaton. Cambridge: Cambridge University Press.

Hardle, W., and M. Jerison (1988): "Cross Section Engel Curves Over Time," Discussion Paper Number 160, Universtat Bonn.

Hardle, W., and T. M. Stoker (1989): "Investigating Smooth Multiple Regression by the Method of Average Derivatives," *Journal of the American Statistical Association*, 84, 986–995.

Hausman, J. A., W. K. Newey, and J. L. Powell (1988): "Nonlinear Errors in Variables: Estimation of Some Engel Curves," 1988 Jacob Marschak Lecture of the Econometric Society, Australian Economics Congress.

Jorgenson, D. W., L. J. Lau, and T. M. Stoker (1982): "The Transcendental Logarithmic Model of Aggregate Consumer Behavior," in *Advances in Econometrics*, ed. by R. Basman and G. Rhodes. Greenwich: JAI Press.

LAU, L. J. (1977a): "Complete Systems of Consumer Demand Functions through Duality," in *Frontiers of Quantitative Economics IIIA*, ed. by M. D. Intriligator. Amsterdam: North Holland, 59–86.

——— (1977b): "Existence Conditions for Aggregate Demand Functions," Technical Report No. 248, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, California (Revised February, 1980).

——— (1982): "A Note on the Fundamental Theorem of Exact Aggregation," *Economics Letters*, 9, 119–126.

LAWLEY, D. N., AND A. E. MAXWELL (1971): *Factor Analysis as a Statistical Method*. London: Butterworth and Co.

LESER, C. E. V. (1963): "Forms of Engel Functions," *Econometrica*, 31, 694–703.

LEWBEL, A. (1987a): "Fractional Demand Systems," *Journal of Econometrics*, 36, 311–337.

——— (1987b): "Characterizing Some Gorman Engel Curves," *Econometrica*, 55, 1451–1459.

——— (1988): "Exact Aggregation, Distribution Parameterizations, and a Nonlinear Representative Consumer," *Advances in Econometrics*, Vol. 7, ed. by G. F. Rhodes and T. B. Fomby. Greenwich: JAI Press, 253–290.

——— (1989a): "A Demand System Rank Theorem," *Econometrica*, 57, 701–705.

——— (1989b): "Portfolio Separation With General Preference Structures: A Rank and Dual Approach," Unpublished draft.

——— (1989c): "Additive Utility, Rank, the Marginal Utility Cost Function, and Portfolio Separation," Unpublished draft.

——— (1989d): "Exact Aggregation and a Representative Consumer," *Quarterly Journal of Economics*, 104, 622–633.

——— (1989e): "Household Equivalence Scales and Welfare Comparisons," *Journal of Public Economics*, 39, 377–391.

——— (1989f): "Nesting the AIDS and Translog Demand Systems," *International Economic Review*, 30, 349–356.

——— (1990): "Full Rank Demand Systems," *International Economic Review*, 31, 289–300.

——— (1991): "Cost of Characteristics Indices and Household Equivalence Scales," *European Economic Review*, forthcoming.

MAGNUS, J. R. (1984): "On Differentiating Eigenvalues and Eigenvectors," ICERD Discussion Paper, L.S.E.

MANSER, M. E., AND R. J. MCDONALD (1988): "An Analysis of Substitution Bias in Measuring Inflation, 1959–1985," *Econometrica*, 56, 909–930.

MUELLBAUER, J. (1975): "Aggregation, Income Distribution, and Consumer Demand," *Review of Economic Studies*, 62, 269–283.

——— (1976): "Community Preferences and The Representative Consumer," *Econometrica*, 44, 525–543.

PRAKASA RAO, B. L. S. (1983): Nonparametric Functional Estimation. New York: Academic Press.

STOKER, T. M. (1984): "Exact Aggregation and Generalized Slutsky Conditions," *Journal of Economic Theory*, 33, 368–377.

——— (1990): "Equivalence of Direct, Indirect, and Slope Estimators of Average Derivatives," forthcoming in *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, ed. by W. A. Barnett, J. Powell, and G. Tauchen. Cambridge: Cambridge University Press.

VARIAN, H. R. (1983): "Non-parametric Tests of Consumer Behavior," *Review of Economic Studies*, 50, 99–110.

References

Asymptotic Theory for Principal Component Analysis
T. W. Anderson
The Annals of Mathematical Statistics, Vol. 34, No. 1. (Mar., 1963), pp. 122-148.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196303%2934%3A1%3C122%3AATFPCA%3E2.0.CO%3B2-D

Transcendental Logarithmic Utility Functions
Laurits R. Christensen; Dale W. Jorgenson; Lawrence J. Lau
The American Economic Review, Vol. 65, No. 3. (Jun., 1975), pp. 367-383.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28197506%2965%3A3%3C367%3ATLUF%3E2.0.CO%3B2-F

An Almost Ideal Demand System
Angus Deaton; John Muellbauer
The American Economic Review, Vol. 70, No. 3. (Jun., 1980), pp. 312-326.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28198006%2970%3A3%3C312%3AAAIDS%3E2.0.CO%3B2-Q

An Elasticity can be Estimated Consistently without a Priori Knowledge of Functional Form
Ibrahim Elbadawi; A. Ronald Gallant; Geraldo Souza
Econometrica, Vol. 51, No. 6. (Nov., 1983), pp. 1731-1751.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198311%2951%3A6%3C1731%3AAECBEC%3E2.0.CO%3B2-4

LINKED CITATIONS

Engel Curves Leading to the Weak Axiom in the Aggregate
Xavier Freixas; Andreu Mas-Colell
Econometrica, Vol. 55, No. 3. (May, 1987), pp. 515-531.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198705%2955%3A3%3C515%3AECLTTW%3E2.0.CO%3B2-G

Community Preference Fields
W. M. Gorman
Econometrica, Vol. 21, No. 1. (Jan., 1953), pp. 63-80.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28195301%2921%3A1%3C63%3ACPF%3E2.0.CO%3B2-9

Separable Utility and Aggregation
W. M. Gorman
Econometrica, Vol. 27, No. 3. (Jul., 1959), pp. 469-481.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28195907%2927%3A3%3C469%3ASUAA%3E2.0.CO%3B2-A

Investigating Smooth Multiple Regression by the Method of Average Derivatives
Wolfgang Hardle; Thomas M. Stoker
Journal of the American Statistical Association, Vol. 84, No. 408. (Dec., 1989), pp. 986-995.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28198912%2984%3A408%3C986%3AISMRBT%3E2.0.CO%3B2-C

Forms of Engel Functions
C. E. V. Leser
Econometrica, Vol. 31, No. 4. (Oct., 1963), pp. 694-703.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28196310%2931%3A4%3C694%3AFOEF%3E2.0.CO%3B2-I

Characterizing Some Gorman Engel Curves
Arthur Lewbel
Econometrica, Vol. 55, No. 6. (Nov., 1987), pp. 1451-1459.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198711%2955%3A6%3C1451%3ACSGEC%3E2.0.CO%3B2-O

LINKED CITATIONS

An Analysis of Substitution Bias in Measuring Inflation, 1959-85
Marilyn E. Manser; Richard J. McDonald
Econometrica, Vol. 56, No. 4. (Jul., 1988), pp. 909-930.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198807%2956%3A4%3C909%3AAAOSBI%3E2.0.CO%3B2-X

Aggregation, Income Distribution and Consumer Demand
John Muellbauer
The Review of Economic Studies, Vol. 42, No. 4. (Oct., 1975), pp. 525-543.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28197510%2942%3A4%3C525%3AAIDACD%3E2.0.CO%3B2-V

Community Preferences and the Representative Consumer
John Muellbauer
Econometrica, Vol. 44, No. 5. (Sep., 1976), pp. 979-999.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197609%2944%3A5%3C979%3ACPATRC%3E2.0.CO%3B2-K

Non-Parametric Tests of Consumer Behaviour
Hal R. Varian
The Review of Economic Studies, Vol. 50, No. 1. (Jan., 1983), pp. 99-110.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28198301%2950%3A1%3C99%3ANTOCB%3E2.0.CO%3B2-V