A Rational Rank Four Demand System

Arthur Lewbel Boston College

revised October 2001, corrected March 2003

Abstract

Past parametric tests of demand system rank employed polynomial Engel curve systems. However, by Gorman's (1981) theorem, the maximum possible rank of a utility derived polynomial demand system is three. The present paper proposes a class of demand systems that are utility derived, are close to polynomial, and have rank four. These systems nest rational polynomial demands, and so can be used to test ranks up to four. These systems are suitable for applications where high rank is likely, such as demand systems involving a large number of goods. A test of rank using this new class of systems is applied to UK consumer demand data.

Keywords: Demand Systems, Rank, Utility Functions, Engel Curves.

This research was supported in part by the National Science Foundation through grants SBR-9514977 and SES-9905010. I'd like to thank John Rust, Krishna Pendakur, and an anonymous referee for helpful comments.

Arthur Lewbel, Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA, 02467, USA. (617)-552-3678, lewbel@bc.edu, http://www2.bc.edu/~lewbel/

1 Introduction

The rank of a demand system can be defined as the dimension of the space spanned by its Engel curves. A now famous result by Gorman (1981) is commonly misinterpreted to mean that utility maximization requires demands to have rank three or less. In fact, Gorman only defines rank for the class of exactly aggregable demand systems, and his results only apply to that class. Lewbel (1991) shows that all demand systems (not just exactly aggregable ones) have a rank. Demands that are not exactly aggregable can have rank higher than three, without violating utility maximization.

The present paper proposes a large set of rank four demand systems that are consistent with utility maximization and nest commonly used exactly aggregable demand systems of lower rank as special cases. The Engel curves of these demand systems are then used to test for rank.

Demand rank has many implications for aggregation, separability, empirical modeling, and welfare analyses. The following are some examples (see Lewbel 1991 and 1997 for details). If utility is homothetically separable into L groups of goods, then the rank of the demand system is a lower bound on L. If a household's utility function is a social welfare function over M homothetic subutility functions, corresponding, e.g., to the utility functions of each member of the household, then rank is also a lower bound on M.

A demand system has rank one if and only if it is homothetic, meaning that all income elasticities equal one. Most conditions required for aggregate demands to resemble those of a representative consumer require either rank one or two. Utility derived demands that are exactly aggregable, meaning demands in the form of equation (2) below, must have rank less than or equal to three. Utility derived deflated income demands, meaning demands in the form of equation (3) below, must have less than or equal to four.

In demand functions over financial assets, the rank of the demand system equals the degree of portfolio separation, for example, in the standard capital asset pricing model, utility is maximized by holding shares in two funds (the market and the riskless asset). Therefore, the types of utility functions that give rise to the CAPM, i.e., utility functions that depend only on the mean and variance of asset returns, have rank two. See Lewbel and Perraudin (1995).

The models described here are assumed to apply to a single consuming individual or household. In empirical applications, some or all of the functions defined below could be specified to vary by demographic or other consumer characteristics.

Let w be the vector of budget shares of the J goods that a consumer buys, let p be the J vector of prices of the goods, and x be total expenditures (income for short). Any demand system can be written in the form

(1)
$$w = \sum_{r=1}^{R} A_r(p) f_r(p, x)$$

for some $R \leq J$, where for each r = 1, ..., R, f_r is a scalar valued function of income and prices and $A_r(p)$ is a J vector valued function of prices. Let A(p) be the J by R matrix having columns $A_r(p)$ for r = 1, ..., R. Define a demand system to be rational if it is derived from the maximization of a utility function.

Generalizing Gorman (1981), Lewbel (1991) defines the rank M of any demand system to be the maximum rank, over all possible price vectors p, of the matrix A(p). Equivalently, the rank M of any demand system can be defined as the maximum dimension of the function space spanned by the demand system's Engel curves. Lewbel (1991) shows that, for demands derived from utility maximization, the rank M of a demand system will also equal the minimum number of functions $\theta_m(p)$ such that the indirect utility function can be written in the form $u = V[x, \theta_1(p), ..., \theta_M(p)]$. This result gives rise to many of the above listed implications of rank, and illustrates how rank provides valuable information for specifying appropriate functional forms for utility functions and their associated demand systems.

Gorman (1981) analyzes a special class of demand systems known as exactly aggregable demands. These are demands that can be written in the form

(2)
$$w = \sum_{r=1}^{R} B_r(p) g_r(x)$$

where R is a positive integer, and for each r = 1, ..., R, $B_r(p)$ is a J vector valued function of prices and g_r is a scalar valued function of just income (not prices). Let B(P) be the J by R matrix of columns $B_r(p)$. Exactly aggregable demands are useful because, as their name implies, they can be summed across consumers to yield closed form (though not necessarily representative consumer) expressions for aggregate demands. See, e.g., Jorgenson, Lau, and Stoker (1982). The fact that utility derived demand systems must be homogeneous of degree zero in x and p greatly limits the types of Engel curves that (2) can possess, in particular, the functions g_r must either be polynomials (in levels or logs) or trigonometric functions. Gorman's (1981) famous rank theorem is that the maximum possible

rank of the matrix B(p) in any rational, exactly aggregable demand system is three.

Gorman did not define the concept of rank for demand systems other than those in the exactly aggregable class of equation (2). However, Gorman's theorem implies that, in terms of Lewbel's general definition of rank, utility derived exactly aggregable demands have a maximum rank of three. It is therefore of particular interest for empirical analysis to know if demands have rank greater than three, because if they do then they cannot be exactly aggregated.

Demand systems that are not exactly aggregable, and hence have the form of equation (1) but not (2), can be rational with a rank greater than three. For example, consider deflated income demand systems, which are defined by

(3)
$$w = \sum_{r=1}^{R} C_r(p) h_r[x/b(p)]$$

where b(p) is a scalar valued function and for each $r = 1, ..., R, C_r(p)$ is a J vector valued function of prices and h_r is a scalar valued function of deflated income x/b(p). Lewbel (1989) proposed this class of models to relax the constraints on Engel curves that are imposed by homogeneity in exactly aggregable models. Lewbel (1989) gives a proof that the maximum possible rank of a rational deflated income demand system is not three, but four.

Parametric empirical tests of demand system rank include Hausman, Newey, and Powell (1995), Banks, Blundell, and Lewbel (1997), Lyssioto and Pashardes (1997), and Nicol (2000). Ignoring the role of demographic characteristics, these parametric rank tests all use models of the log polynomial form

(4)
$$w = d(p, x) = \sum_{r=1}^{R} B_r(p) (\ln x)^{r-1}$$

This functional form is used because it nests popular models like the Translog, Almost Ideal, and Quadratic Almost Ideal (QUAIDS) demand systems, and because numerous engel curve and demand system studies confirm that budget shares are well approximated by low order polynomials. See, e.g., Howe, Pollak, and Wales (1979), Deaton and Muellbauer (1980), Jorgenson, Lau, and Stoker (1982), Bierens and Pott-Buter (1987), Härdle and Jerison (1988), Atkinson, Gomulka, and Stern (1990), Lewbel (1991), Blundell, Pashardes, and Weber (1993), Hildenbrand (1994), and Hausman, Newey, and Powell (1995).

Since (4) is a special case of (2), by Gorman's theorem rationality requires the rank of the demand system (4) to be three or less. Therefore, when studies that

use polynomials like (4) find that rank four is empirically rejected (against the alternative that the rank is three or less), we do not know if that rejection is caused by rank being really three or less, or if it is caused by the fact that polynomials having rank four must violate rationality.

Nonparametric rank tests are proposed and implemented by Lewbel (1991), Banks, Blundell, and Lewbel (1997), and Donald (1997). These tests also find a rank of three, but in general nonparametric tests have lower power than appropriately specified parametric tests.

The goal of this paper is to construct a demand system that is similar to (4), and indeed nests (4) for $R \le 3$, but can also have rank four without violating rationality. This system is then used to parametrically test the null hypothesis that the rank is three (or less) versus rank four, without imposing irrationality under the rank four alternative.

The proposed class of rank four models are deflated income demand systems. In addition to testing, these systems may prove useful in applications where high rank is known or suspected, e.g., in models involving a large number of diverse goods. For example, Nicol (2000) finds that conditional demands for some demographic groups are rank three, which implies that unconditional demands for those groups are likely to be rank four.

2 A Nearly Log Polynomial Rational Rank Four Demand System

Consider the class of indirect utility functions given by

(5)
$$u^{-1} = \left(\frac{\ln[x - a(p)] - b(p)}{c(p)}\right)^{-1} + d(p)$$

where a, b, c, and d are functions of prices. Homogeneity requires that c and d be homogeneous of degree zero and that $\exp(b)$ and a be homogeneous of degree one in p. Application of Roy's identity to this utility function yields demands of the form

(6)
$$w = \frac{A_4(p)}{x} + \left(1 - \frac{a(p)}{x}\right) \sum_{r=1}^3 A_r(p) (\ln[x - a(p)])^{r-1}$$

where

(7)
$$A_1(p) = \frac{\partial b(p)}{\partial \ln p} - \frac{\partial c(p)}{\partial \ln p} \frac{b(p)}{c(p)} + \frac{\partial d(p)}{\partial \ln p} \frac{b(p)^2}{c(p)}$$

(8)
$$A_{2}(p) = \frac{\partial c(p)}{\partial \ln p} \frac{1}{c(p)} - 2 \frac{\partial d(p)}{\partial \ln p} \frac{b(p)}{c(p)}$$
(9)
$$A_{3}(p) = \frac{\partial d(p)}{\partial \ln p} \frac{1}{c(p)}$$
(10)
$$A_{4}(p) = \frac{\partial a(p)}{\partial \ln p}$$

This is a deflated income demand system and its rank is four, provided that no one of the functions a, b, c, or d can be written as a function of the other three. This can be directly checked by examination of equations (6) to (10), or by applying Theorem 1 of Lewbel (1991) to equation (5).

If a(p) = 0 for all p, then $A_4(p)$ also equals zero, and equation (6) reduces to the quadratic logarithmic model, which is equation (4) with R = 3. Moreover, applying results in Banks, Blundell, and Lewbel (1997), all rank three quadratic logarithmic demands have utility functions given by equation (5) with a(p) = 0, so (5) nests all possible rank three quadratic logarithmic demand systems.

The presence of 1/x terms in budget share models is also supported by some empirical work. For example, The model of Howe, Pollak, and Wales (1979) has budget shares that are linear in x and 1/x.

3 A Rank Four Almost Ideal Demand System

To provide an explicit example of a utility derived rank four system, consider the following definitions, where 1 denotes a J vector of ones, α_0 is a scalar parameter, α , β , δ , and λ are J vectors of parameters, and Γ is a symmetric, J by J matrix of parameters.

(11)
$$\ln a(p) = \delta' \ln p$$
(12)
$$b(p) = \alpha_0 + \alpha' \ln p + \ln p' \Gamma \ln p/2$$
(13)
$$\ln c(p) = \beta' \ln p$$
(14)
$$d(p) = \lambda' \ln p$$

where $\alpha' \mathbf{1} = 1$, $\beta' \mathbf{1} = 0$, $\Gamma \mathbf{1} = 0\mathbf{1}$, $\delta' \mathbf{1} = 1$, and $\lambda' \mathbf{1} = 0$. This model has budget shares given by equation (6) with

(15)
$$A_1(p) = \alpha + \Gamma \ln p - \frac{b(p)}{c(p)} \beta + \frac{b(p)^2}{c(p)} \lambda$$

(16)
$$A_2(p) = \frac{1}{c(p)}\beta - \frac{2b(p)}{c(p)}\lambda$$

$$(17) \qquad A_3(p) = \frac{1}{c(p)}\lambda$$

$$(18) \qquad A_4(p) = \delta a(p)$$

or more simply

19)
$$w = \frac{a(p)}{x} \delta + \left(1 - \frac{a(p)}{x}\right) \left(\alpha + \Gamma \ln p + \frac{\ln[x - a(p)] - b(p)}{c(p)} \beta + \frac{\left[\ln[x - a(p)] - b(p)\right]^2}{c(p)} \lambda\right)$$

In empirical applications with heterogeneous households, one might let the α or δ parameters vary by demographic characteristics of the household. The rank three QUAIDS (Quadratic Almost Ideal) model of Banks, Blundell, and Lewbel (1997) equals the special case of this model in which a(p) = 0, and the rank two Almost Ideal Demand System of Deaton and Muellbauer (1980) equals the special case of a(p) = 0 and d(p) = 0.

4 Engel Curves

To avoid having to specify functional forms for the price functions in equation (6), rank will be tested directly in a half a dozen different time periods by using Engel curves. Assuming all households in each time period t in a sample face the same prices p_t , for each good j = 1, ..., J, equation (6) can be rewritten as

(20)
$$w_{jit} = \frac{A_{4jt}}{x_{it}} + \left(1 - \frac{a_t}{x_{it}}\right) \sum_{r=1}^{3} (\ln[x_{it} - a_t])^{r-1} A_{rjt} + e_{jit}$$

Where *i* indexes consuming units (households), A_{rjt} (for r = 1, 2, 3, 4 and j = 1, ..., J) and a_t are scalar constants to be estimated in each time period t, and each e_{jit} is an error term. Given a sample of n_t demographically homogeneous households in each time period t, equation (20) for j = 1, ..., J is a system of J equations that can be jointly estimated by maximum likelihood.

The errors e_{jit} are assumed to be jointly normal, independent across households and time periods, but correlated across goods. This model of the errors is what Pollak and Wales (1992) refer to as the "standard stochastic specification" for demand systems. An advantage of normality is that the errors in all equations can

be jointly normal while still maintaining the adding up constraint, which requires that $\sum_{j=1}^{J} e_{jit} = 0$. A disadvantage of normality is that it is formally inconsistent with budget shares bounded between zero and one, though this is generally not a problem unless actual budget shares are close to zero or one. Moreover, the estimator remains consistent even without normality, because it is equivalent to a weighted nonlinear least squares. A similar objection can be raised against the polynomial specification of budget shares in equation (20), since for extreme values of income polynomials will also go outside unit simplex. The restriction that budget shares sum to one is imposed by dropping the J=5 equation from the system, and just estimating the other four equations jointly using maximum likelihood. Joint normality of the errors ensures that the resulting estimates will not depend on which equation is dropped. See Lewbel (1997) for a survey on related issues regarding estimation of demand systems.

5 Test Results

The data set is the same one used by Banks, Blundell, and Lewbel (1997) to estimate the integrable Quadratic Almost Ideal Demand System (QUAIDS), a quadratic logarithmic rank three functional form. These data are a demographically homogeneous subsample (married couples without children living in London and the southeast) of the UK Family Expenditure Survey.

Here x_{it} is total expenditures on nondurables, in hundreds of british pounds per week. Nondurable expenditures are divided into J = 5 groups: j = 1 is food, 2 is fuel, 3 is clothing, 4 is alcohol, and 5 is other. Nondurable expenditures are assumed to be separable from other components of utility (such as durables and leisure). Only the demands generated by a subutility function over these five components of nondurables are estimated, and hence only the rank of this subsystem is tested. The rank of a subutility function is a lower bound on the rank of the entire utility function.

Using both parametric and nonparametric tests, Banks, Blundell, and Lewbel (1997) empirically show that quadratic logarithmic utility appears to be the best rank three functional form for modeling demands. Estimation of equation (20) will in general yield a rank four system, that, unlike a cubic in $\ln(x)$, need not violate rationality. Testing whether $a_t = 0$ and $A_{4jt} = 0$ for j = 1, 2, 3, 4 then provides a test of the rank three quadratic logarithmic system against a potentially rational rank four alternative.

Table 1 lists χ_5^2 likelihood ratio test statistics, testing the joint hypothesis that

 $a_t = 0$ and $A_{4jt} = 0$ for data in each two year period from 1970 to 1985.

TABLE 1: LIKELIHOOD RATIO TESTS OF RANK FOUR VERSUS THREE

Years								
χ_5^2 Statistic	1.96	1.81	1.30	19.17	4.90	10.46	5.28	5.61

The 95% critical value for a χ_5^2 statistic is 11.07. One pair of years, 1976-77, rejects rank three in favor of rank four at this significance level, and 1980-81 comes close to rejecting, while all other pairs fail to reject. Since households are assumed to be independent, we can construct a joint test of the hypothesis that $a_t = 0$ and $A_{4jt} = 0$ in all years by combining the above statistics. The result, which has a χ_{40}^2 distribution under the null, equals 50.49. The 95% critical value for this statistic is 55.76, while the 90% value is 51.81. So we formally fail to reject rank three, though the p-value is not far below the 90%. level.

Banks, Blundell, and Lewbel (1997) provide some graphs of nonparametrically estimated engel curves from these data. These graphs look very close to quadratic in ln(x), consistent with rank three. To further investigate the differences between rank three and rank four estimates, results are reported below in detail for two sample time periods; 1974-75 and 1976-77. These years have, respectively, the lowest and highest chi-squared statistics in Table 1. Table 2 provides data means, Table 3 presents coefficient estimates (with standard errors in parentheses) and Table 4 gives income elasticities implied by the models with these data.

TABLE 2: DATA MEANS						
	w_1	w_2	w_3	w_4	\boldsymbol{x}	
1974-75	.323	.064	.086	.051	.410	
1976-77	.310	.068	.079	.058	.587	

TABLE 3: PARAMETER ESTIMATES

	74-75, rank 3	74-75, rank 4	76-77, rank 3	76-77, rank 4
A_{11}	.162 (.011)	.161 (.035)	.218 (.007)	.329 (.092)
A_{12}	.032 (.003)	.011 (.014)	.044 (.002)	.080 (.025)
A_{13}	.095 (.013)	.081 (.048)	.102 (.006)	026 (.078)
A_{14}	.060 (.009)	.070 (.033)	.062 (.004)	.078 (.022)
A_{21}	172 (.024)	165 (.026)	165 (.018)	261 (.091)
A_{22}	032 (.007)	027 (.010)	035 (.006)	066 (.026)
A_{23}	017 (.028)	017 (.029)	.025 (.016)	.134 (.083)
A_{24}	001 (.019)	004 (.019)	017 (.009)	031 (.019)
A_{31}	008 (.012)	.004 (.049)	020 (.015)	.097 (.062)
A_{32}	001 (.004)	025 (.019)	.003 (.005)	.041 (.020)
A_{33}	023 (.014)	040 (.058)	011 (.013)	145 (.065)
A_{34}	008 (.009)	.001 (.039)	025 (.007)	008 (.025)
A_{41}		.002 (.031)		116 (.097)
A_{42}		.020 (.015)		038 (.026)
A_{43}		.015 (.040)		.134 (.077)
A_{44}		008 (.027)		016 (.024)
a		.023 (.014)		.005 (.047)

As can be seen in Table 3, none of the A_{4jt} and a_t parameters required for rank four have significant t-statistics at the 95% level. Even in 1976 – 77, the difference in overall fit between rank three and rank four is minimal, e.g., the simple R^2 statistics for the four equations from the quadratic logarithmic model in that period are .33, .20, .04, and .02, while the corresponding R^2 statistics for the rank four specification are .34, .21, .06, and .02.

TABLE 4: INCOME ELASTICITIES						
		74-75, rank 3	74-75, rank 4	76-77, rank 3	76-77, rank 4	
food	first x decile	0.888	0.869	0.951	0.526	
	mean x	0.909	0.845	0.985	0.661	
	last x decile	0.942	0.837	1.051	0.850	
fuel	first x decile	0.977	1.055	0.973	0.843	
	mean x	0.982	1.020	0.988	0.899	
	last x decile	0.989	0.991	1.009	0.970	
clothing	first x decile	1.074	1.117	1.083	1.463	
	mean x	1.063	1.062	1.096	1.262	
	last x decile	1.064	1.019	1.126	1.008	
alcohol	first x decile	1.036	0.990	1.060	0.991	
	mean x	1.038	0.992	1.046	0.982	
	last x decile	1.045	0.997	1.045	0.985	

Although the change from rank three to four is only marginally significant statistically, in some cases the behavioral implications of the two models differ. This can be seen in Table 4, which provides estimated elasticities at the means and in the tails (the first and last x decile) of the data. In particular, the rank four model in 1976-77 has what may be implausibly large variations in income responses to food and clothing, which could provide evidence of overfitting.

6 Conclusions

This study proposes a rank four demand system that is consistent with utility maximization and nests the popular log linear and log quadratic specifications. In contrast, previous parametric studies test rank four using a cubic in log income, which is inappropriate because a rank four cubic cannot be consistent with utility maximization. Empirical estimates of this system's Engel curves generally confirm the results found in previous studies that demands appear to be rank three,

although there is some evidence for the possibility of rank four. Apart from testing, this new functional form may prove useful for estimation in contexts where demands of high rank are suspected, such as systems having a large number of diverse goods.

7 Appendix

The models were estimated assuming normal errors using a concentrated likelihood function, coded in the maxlik procedure in Gauss. In the rank four model the term $\ln(x-a)$ is undefined if a exceeds the minimum value of x. To prevent numerical problems in the optimization procedures, the variable x is scaled to lie between zero and one (by defining its units as hundreds of UK pounds per week), the function $\ln(x-a)$ was recoded as 10^{-10} when a>x, and the penalty function $[1+a-.9\min(x)]^{100}$ was subtracted from the likelihood function. In every case the parameter a converged to a value sufficiently smaller than x so as to make this penalty equal zero to four decimal places.

8 References

Atkinson, A. B., J. Gomulka, and N. H. Stern (1990), "Spending on Alcohol: Evidence From the Family Expenditure Survey 1970-1983," Economic Journal, 100, 808-827.

Banks, J., R. Blundell, and A. Lewbel (1997), "Quadratic Engel Curves and Consumer Demand," Review of Economics and Statistics, 79, 527-539.

Bierens, H. and H. A. Pott-Buter (1990), "Specification of Household Expenditure Functions and Equivalence Scales by Nonparametric Regression," Econometric Reviews, vol. 9, 123-210.

Blundell, R. W., P. Pashardes, and G. Weber (1995), "What do we learn about Consumer Demand Patterns from Micro Data?", American Economic Review, vol. 83, 570-597.

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

Donald, S. G. (1997), "Inference Concerning the Number of Factors in a Multivariate Nonparametric Relationship," Econometrica, 65, 103-132.

- Gorman, W. M. (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.
- Härdle, W. and M. Jerison (1988), "Cross Section Engel Curves Over Time," Discussion Paper number 160, Universtat Bonn.
- Hausman, J.A., W. K. Newey, and J. L. Powell (1995), "Nonlinear Errors in Variables: Estimation of Some Engel Curves," Journal of Econometrics, 65, 205-253.
 - Hildenbrand, W. (1994), Market Demand, Princeton University Press.
- Howe, H., R. A. Pollak, and T. J. Wales (1979), "Theory and Time Series Estimation of the Quadratic Expenditure System," Econometrica, vol 47, 1231-1243.
- Jorgenson, D. W., L. J. Lau, and T. M. Stoker (1982), "The Transcendental Logarithmic Model of Aggregate Consumer Behavior," in: Advances in Econometrics, R. Basman and G. Rhodes, Eds., Greenwich: JAI Press.
- Lewbel, A. (1989), "A Demand System Rank Theorem," Econometrica, vol. 57, 701-705.
- Lewbel, A. (1991), "The Rank of Demand Systems: Theory and Nonparametric Estimation," Econometrica, 59, 711-730.
- Lewbel, A. (1997), "Consumer Demand Systems and Household Equivalence Scales" in: Handbook of Applied Econometrics, Vol II: Microeconomics, M. H. Pesaran and P. Schmidt, Eds., Oxford: Blackwell publishers.
- Lewbel, A. and W. Perraudin (1995), "A Theorem on Portfolio Separation With General Preferences," Journal of Economic Theory, 65, 624-626.
- Lyssiotou, P., P. Pashardes, and T. Stengos (1998), "Preference Heterogeneity and the Rank of Demand Systems," manuscript, University of Cyprus.
- Nicol, C. (2000), "The Rank and Model Specification of Demand Systems: An Empirical Analysis Using United States Microdata." Canadian Journal of Economics, forthcoming.
- Pollak, R. A. and T. J. Wales (1992), Demand System Specification and Estimation, Oxford: Oxford University Press.