Aggregation, Rank, and Some Historical Engel Curves

Dora L. Costa[*]

April 14, 2000

ABSTRACT

I find evidence for an Engel curve rank of 3 in both historical and recent data within the context of polynomial Engel curves. The results have implications for specification and for aggregation of consumer demands.

JEL Classifications: D12

Keywords: Engel curves, aggregation, rank

Contact information:
Dora L. Costa
MIT
Department of Economics, E52-274C
50 Memorial Drive
Cambridge, MA 02142
costa@mit.edu

1 Introduction

Estimation of Engel curves has a long history in economics; but, economic theory provides little guidance in the specification of Engel curves. The most common systems such as AIDS, translog, linear expenditure, PIGL, and PIGLOG have been favored because of their exact aggregation or representative agent properties, not because of their superior fit. Lewbel (1991) found that although most household demands could be modeled as PIGLOG, a more complicated model is required when households with very low or high expenditures are included in the sample. Hausman et al. (1991) found that Engel curves containing cubic terms in functions of expenditures fit the data better than specifications containing lower order polynomials. Such a system will be exactly aggregable provided that the rank of the matrix of coefficients is at most three (Gorman 1981).

This note uses data on consumer expenditures from as far back as 1888-1890 to estimate rank within the context of polynomial Engel curves and to test whether the rank three condition is met when Engel curves contain up to a cubic term in the logarithm of total expenditures. The rank test is a useful prespecification test providing information on aggregation and separability. Failure to meet the rank three condition in specifications containing polynomials in functions of expenditures implies that it is not possible to aggregate across different income levels in estimating consumer demands. Failure might be more likely with historical data because the Bureau of Labor Statistics had not yet perfected its interview methods. Rank may be lower in historical than in recent data because the older consumer expenditure surveys did not cover the entire income distribution.

2 Data and Empirical Methodology

I use consumer expenditure surveys from 1888-90, 1917-19, 1935-36, 1972-73, and 1992. The post-war surveys are a representative sample of the population. The earlier surveys were more

specialized. One major difference is that they did not cover "slum" or "charity" families and those before 1935 did not cover higher income families. I impose several sample restrictions to make the datasets more comparable across time, to minimize income correlated demographic variation, and to exclude suspect observations or observations where demand may be unusual. I restrict all surveys to urban families, to husband and wife families, to families in the post-war period who were not receiving welfare or food stamps, and to families where the husband was age 21-64. I exclude observations in which the share of expenditures devoted to food was less than 0.05 or greater than 0.8. In the post-war period I also exclude families containing adults over age 25 other than the husband and wife. I adjust total expenditures in the 1917-19 and 1972-73 surveys for inflation.

I estimate Engel curves of the form

$$w_i = \alpha + \beta_1 \log(z) + \beta_2 \log^2(z) + \beta_3 \log^3(z) + \mathbf{x'_i}\beta + u_i \tag{1}$$

for each year of consumer expenditures where w is the share of total expenditures spent on a specific commodity group, z is total expenditures, x is a vector of demographic coefficients, and u is an error term. The expenditure shares that I examine are those devoted to food, shelter, utilities, recreation, clothing, health, education, and transportation. The demographic variables include the total number of children, the total number of children squared, the age of the household head, and the number of household members above age 18. Because of price fluctuations during the inflation accompanying World War I and because the 1917-19 survey was taken at different times, the Engel curve for 1917-19 also includes dummies (8, including the omitted variable) indicating when the survey was taken.

I test whether all higher order terms for all commodity groups in Equation 1 are equal to 0. A system that is linear in the logarithm of income has a rank of 2. Any system with higher

order terms will have a rank of at most 3 (Gorman 1981). Gorman's (1981) theorem implies that higher order polynomials in log income will have a linear relationship to the lower-order terms since the matrix of coefficients is at most of rank 3. For Equation 1 the rank restriction takes the form that the ratio of coefficients of the cubic terms to the coefficients of the quadratic terms will be constant across equations (Hausman et al. 1995). I therefore estimate the Gorman ratios to see whether the matrix of coefficients has rank 3. I also test whether a cubic or quadratic specification is preferred by testing whether all cubic terms for all commodity groups are equal to 0.

3 Results and Conclusion

I reject the hypothesis that all higher order terms in Equation 1 are equal to 0 for all commodity groups. The $\chi^2(12)$ statistic for 1888-90 is 127.63 and the $\chi^2(18)$ statistics for 1917-19, 1935-36, 1972-73, and 1992 are 201.14, 84.49, 195.82, and 41.43, respectively. I also reject the hypothesis that all cubic terms for all commodity groups are equal to 0. The $\chi^2(6)$ statistic is 25.81 in 1888-90 and the $\chi^2(9)$ statistics for 1917-19, 1935-36, 1972-73, and 1992 are 78.40, 39.24, 108.14, and 19.70, respectively.

Table 1 shows that within each year the Gorman ratios are remarkably similar across commodity groups. When the ratios do differ, the standard error on the ratio is large. A χ^2 test for equality of the ratios across commodity groups cannot reject the hypothesis that the Gorman ratios are equal at the 5 percent level of significance. Only in 1917-19 could the hypothesis that the Gorman ratios are equal be rejected at the 10 percent level of significance, largely because the ratios on food and clothing are different. However, when households who had income from gardens and therefore probably had some degree of self-sufficiency in food are excluded from the sample the $\chi^2(8)$ statistic falls to 5.18.

The results indicate that polynomial Engel curve specifications with at least a cubic

3

Table 1: Estimated Ratios of β_2/β_3 for Equation 1

	1888-90	1917-19	1935-36	1972-73	1992
Recreation	-16.833	-27.319	-22.836	-17.711	-31.262
	(1.176)	(0.129)	(1.096)	(96.639)	(0.451)
Food	-35.632	-28.057	-22.442	-34.039	-31.156
	(71.037)	(0.348)	(0.424)	(3.651)	(1.063)
Clothing	-19.790	-26.336	-22.023	-28.106	-30.992
	(2.486)	(0.429)	(1.001)	(2.283)	(0.861)
Furniture	-19.469	-27.298	-9.952	-21.402	-30.737
	(0.296)	(0.171)	(57.727)	(71.042)	(1.020)
Fuel	-20.016	-27.999	-22.590	-29.672	-31.081
	(0.272)	(0.384)	(0.284)	(0.155)	(0.289)
Health	-21.085	-27.296	-21.799	-29.733	-30.566
	(2.835)	(0.233)	(1.904)	(0.280)	(2.243)
Education		-27.269	-20.237	-20.965	-26.550
		(0.204)	(2.182)	(11.796)	(14.223)
Transportation		-18.298	-22.061	-29.168	-31.432
		(24.101)	(0.403)	(0.207)	(0.464)
Shelter		-27.090	-22.266	-26.772	-29.220
		(0.109)	(0.820)	(1.642)	(1.239)
Observations	6,407	11,849	1,650	7,602	1,677
$\chi^2(5)$, $\chi^2(8)$	8.78	14.4	4.53	9.96	3.43

The χ^2 test is for equality of the ratios across commodity groups. Standard errors are in parentheses. Only shelter expenses for renters are known in 1888-90. Population weights were used in the 1972-73 and 1992 estimation.

term are preferred to specifications with lower order terms in both recent and historical data. The findings also provide strong support for the Gorman rank condition which limits polynomial Engel curve specifications to rank 3. Better fitting Engel curves specifications than those of rank 2 can therefore be used while still preserving aggregation properties.

Data Sources

1. Sabelhaus, John. 1996. *Consumer Expenditure Survey: Family Level Extracts, 1980:1-1995:1*. http://www.nber.org.

2. United States Department of Labor. 1986. *Cost of Living of Industrial Workers in the United States and Europe, 1888-1890. ICPSR 7711*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

3. United States Department of Labor. Bureau of Labor Statistics. 1986. *Cost of Living in the United States, 1917-1919. ICPSR 8299*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

4. United States Department of Labor. Bureau of Labor Statistics. 1987. *Survey of Consumer Expenditures, 1972-1973. ICPSR 9034*. Inter-university Consortium for Political and Social Research.

5. United States Department of Labor, Bureau of Labor Statistics and United States Department of Agriculture, Bureau of Home Economics, et al. 1999. *Study of Consumer Purchases in the United States, 1935-1936. ICPSR 8908*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

References

[1] Gorman, W.M. 1981. "Some Engel Curves." In, A. Deaton, Ed. *Essays in the theory and measurement of consumer behavior in honor of Richard Stone.* Cambridge: Cambridge University Press: 7-29.

[2] Hausman, J.A., W.K. Newey, J.L. Powell. 1995. "Nonlinear errors in variables. Estimation of some Engel curves." *Journal of Econometrics*. 65: 205-233.

[3] Lewbel, Arthur. "The Rank of Demand Systems: Theory and Nonparametric Estimation." *Econometrica*. 59(3): 711-30.