An Empirical Comparison of Functional Forms for Engel Relationships

By Larry Salathe*

The functional form used to represent expenditures or consumption as a function of income and household size (Engel relationship) dramatically affects estimates of elasticities of these variables. This impact also holds true when the elasticities are computed at the mean of the sample used. When per capita expenditures were expressed as a function of per capita income, the double-and semi-log functional forms provided the best statistical fit. When expenditures were expressed as a function of household size and income, the quadratic functional form provided the best statistical fit.

Keywords:

Engel curves
functional form
goodness of fit
income
household size

INTRODUCTION

A variety of functional forms have been suggested to represent Engel relationships.[1] The most widely used include the linear, quadratic, double logarithmic, semi-logarithmic, inverse, and logarithmic-inverse. Because each functional form possesses some desirable characteristics, no single form has found general acceptance among economists (*2*, *8*, *5*, *9*, *6*).[2]

Few researchers have examined the discrepancies in results obtained by assuming different functional forms for Engel relationships or the ability of these different functional forms to "fit" the same data. Previous research indicates that the choice of functional form can substantially influence the (estimated) income elasticity. Income (expenditure) elasticity for a particular product can vary by 50 percent or more at the means because of differences in the functional form (*9*).

Prais and Houthakker compared the fit of the linear, double logarithmic, semi-logarithmic, inverse, and logarithmic-inverse functional forms using grouped data. They measured goodness of fit by the correlation between actual and predicted values of the dependent variable.

The interpretation of this measure of goodness of fit varies, depending on whether the dependent variable is transformed before estimation. For example, for the double logarithmic functional form, the computed correlation coefficient measures the correlation between the natural logarithm of observed expenditure (quantity) and the predicted value for the natural logarithm of expenditure (quantity).

However, for the linear functional form, the correlation coefficient measures the correlation between observed expenditure (quantity) and predicted expenditure (quantity). Thus, a more consistent measure of goodness of fit would be to transform the predicted values for the double logarithmic functional form to natural numbers before computing the correlation coefficient.

The quadratic functional form has attracted only limited attention from economists (*7*). This disinterest is somewhat puzzling because the form allows the marginal propensity to consume (spend) and the income elasticity to vary with the level of income. Such flexibility is particularly useful for analyzing expenditures or consumption of commodities considered to be necessities.

OBJECTIVES

The objectives of this article are (1) to examine differences in estimated household size and income elasticities generated by different functional forms including quadratic and (2) to compare the ability of different functional forms to fit ungrouped data. Results should provide a better understanding of the relationship between functional form and estimated income and household size elasticities. In addition, since one criterion for selecting functional forms is goodness of fit, the study should indicate which forms are most appropriate for estimating Engel relationships.

RESULTS

The data used in the analysis consist of 7,143 households in the spring portion of the 1965 USDA Household Food Consumption Survey. The data on food expenditures were grouped into seven expenditure groups: dairy products (excluding butter), fats and oils, flour and cereals, beef and pork, vegetables, fruits, and total food consumed at home.

---

*Larry Salathe is an agricultural economist with the National Economic Analysis Division, ESCS.

[1] An Engel relationship can be defined as describing how expenditures or consumption of a particular commodity varies with household income and size.

[2] Italicized numbers in parentheses refer to items in References at the end of this article.

The functional form used to represent expenditures as a function of income and household size (Engel relationship) dramatically affects estimates of elasticities of these variables. This impact also holds true when the elasticities are computed at the mean of the sample used. When per capita expenditures were expressed as a function of per capita income, the double- and semi-log functional forms provided the best statistical fit. When they were expressed as a function of household size and income, the quadratic functional form provided the best statistical fit.

Per Capita Specification

The first set of results was generated by specifying six different functional relationships between per capita and income. Table 1 contains the mathematical form of the six functional forms and summarizes the properties of each. According to economic theory, the functional form used in estimating Engel relationships should satisfy the adding up constraint. This property implies that predicted expenditures for each good add up to total expenditures. This is the only property which economic theory gives us. Economists have also suggested that the demand for certain goods, in particular, food, may reach a satiety level as income increases.

One disadvantage of the double-logarithmic and logarithmic-inverse functional forms is that observations having zero expenditure cannot be used in the analysis. Eliminating these observations will result in an inflated estimate for the income (expenditure) elasticity. One could assign a small number to the dependent variable when its recorded value equals zero. Here, a value of one cent was assigned as the level of expenditure when the household recorded no expenditure for a particular food group. Thus, the parameters in all the functional forms were estimated from the same data set.

Even though all expenditure-income elasticities were computed at the sample means, substantial differences still exist in the elasticities (table 2). The inverse and log-inverse functional forms generated expenditure-income elasticities considerably lower than those from the other four functional forms. Of those four forms, the double logarithmic produced the highest expenditure-income elasticity for dairy products, beef and pork, vegetables, fruits, and total food; and the lowest income elasticity for flour and cereals, and the fats and oils food groups. Compared with the linear functional form, the quadratic form provided expenditure-income elasticities having a higher absolute value for all food groups except vegetables, which was the only expenditure category in which both per capita income and per capita income squared were positive and significant.

To compare the ability of each functional form to fit the data, correlation coefficients and mean squared error

Table 1—Properties of alternative functional forms for Engel relationships, expenditures, and income expressed in per capita terms

	Functional form	Marginal propensity to spend	Expenditure-income elasticity	Adding-up constraint	Saturation level	Zero observations
Linear	$E=a+bY$	b	bY/E	holds	no	can be used
Quadratic	$E=a+bY+cY^2$	$b+2Y$	$\frac{(b+2cY)Y}{E}$	holds	no	can be used
Double-logarithmic	$1nE=a+b1nY$	bE/Y	b	does not hold	no	cannot be used
Semi-logarithmic	$E=a+b1nY$	b/Y	b/E	does not hold	no	can be used
Logarithmic-inverse	$1nE=a+b/Y$	$-bE/Y^2$	$-b/Y$	does not hold	yes	cannot be used
Inverse	$E=a+b/Y$	$-b/Y^2$	$-b/EY$	holds	yes	can be used

E is per capita expenditures, and Y is per capita income.

Source: (*6*, p. 50).

Table 2—Estimated expenditure-income elasticities from alternative specifications of Engel relationship *

Expenditure item	Functional form					
	Linear	Quadratic	Double log	Semi-log	Inverse	Log-inverse
Dairy products	.128	.150	.217	.153	.049	.083
Fats and oils	.168	.177	.151	.163	.042	.045
Flour and cereals	-.095	-.112	-.225	-.111	-.032	-.054
Beef and pork	.299	.319	.361	.300	.078	.118
Vegetables	.283	.269	.322	.250	.059	.096
Fruits	.293	.325	.519	.295	.078	.178
Total, food[1]	.212	.229	.236	.217	.059	.075

*Calculated at sample means.

[1] Includes only food consumed at home.

Table 3—Mean squared error statistics for various functional forms

Expenditure item	Linear	Quadratic	Double log	Semi-log	Inverse	Log-inverse
			Dollars/week			
Dairy products	5.416	5.257	4.958L	4.997	5.738H	5.121
Fats and oils	.545	.542	.529L	.536	.588H	.540
Flour and cereal	.634	.633	.738	.632L	.652	1.043H
Beef and pork	26.321	25.817	25.314L	25.341	33.404H	27.290
Vegetables	3.673	3.722	3.447L	3.693	4.554H	3.694
Fruits	2.844	2.773	2.907	2.754L	3.408H	2.997
Total, food	167.973	161.922	139.517L	153.044	219.228H	167.123

H—highest value for each expenditure group.
L—lowest value for each expenditure group.

statistics were computed. In every case, the correlation coefficients measure the correlation between observed and predicted expenditures in natural numbers. To provide greater detail on each functional form's ability to fit the data, mean squared error statistics were also computed by converting observed and predicted expenditure values to natural numbers.

Only the mean error statistics appear (table 3) because the two sets of statistics gave the same results. Generally, the double- and semi-log functional forms have the lowest mean squared error while the inverse functional form had the highest mean squared error. However, for the flour and cereals group, the linear, quadratic, semi-log, and inverse functional forms fit the data better than the double log. Since the estimated expenditure-income elasticity for the flour and cereals subgroup was negative, the double logarithmic functional form appears to be a poor choice when estimating Engel relationships for commodities with negative income elasticities.

The double-logarithmic and semi-logarithmic functional forms may be appropriate when per capita expenditures are expressed as a function of per capita income. However, they may not be the most appropriate when income and household size are treated as separate independent regressors.

Household Size and Income as Separate Regressors

In some recent studies, researchers have specified household expenditures as a function of income and household size rather than expressing expenditures and income in per capita terms (*6*, *3*). The double logarithmic and semi-logarithmic functional forms may be appropriate when per capita expenditures are expressed as a function of per capita income. However, they may not be the most appropriate when income and household size are treated as separate independent regressors.

When expenditures and income are expressed in per capita terms, multiplying income and household size by the same constant does not alter per capita expenditures. This implies that when per capita expenditures are expressed as a function of per capita income, the estimated income and household size elasticities are restricted to sum to one. This restriction is relaxed when income and household size are used as separate regressors.

The estimated household size and income elasticities for 15 alternative functional forms appear in table 4. The relations expressing expenditures as a function of the inverse of income provided the lowest expenditure-income elasticities. The relations expressing the natural logarithm of expenditures as a function of the natural logarithm of income usually produced the highest expenditure-income elasticities. The household size elasticities also exhibited the same patterns, but their relative differences are considerably smaller. After excluding the functional forms expressing expenditures as a function of the inverse of income, the estimated expenditure-income elasticities continued to vary, by as

Table 4—Estimated expenditure-income and household size elasticities for functional forms, income and household size as separate regressors*

Functional form	Expenditure-income elasticities							Household size elasticities						
	Dairy products	Fats and oils	Flour and cereals	Beef and pork	Vegetables	Fruits	All food[1]	Dairy products	Fats and oils	Flour and cereals	Beef and pork	Vegetables	Fruits	All food[1]
(1) $E=a+bY+cY^2+dS+fS^2$	.139	.106	−.142	.290	.199	.292	.195	.592	.593	1.086	.452	.425	.378	.535
(2) $E=a+bY+cY^2+d/S$	.125	.075	−.210	.270	.179	.274	.173	.504	.488	.802	.401	.381	.319	.449
(3) $E=a+bY+cY^2+d\ln S$	.134	.084	−.213	.281	.190	.280	.181	.599	.525	1.021	.462	.436	.378	.534
(4) $\ln E=a+b\ln Y+c\ln S$	.210	.135	−.178	.367	.292	.489	.196	.762	.799	1.370	.657	.585	.389	.643
(5) $\ln E=a+b\ln Y+cS+dS^2$	.212	.150	−.150	.376	.298	.488	.206	.701	.731	1.307	.591	.520	.347	.598
(6) $\ln E=a+b\ln Y+c/S$	.176	.093	−.228	.330	.257	.468	.168	.710	.766	1.237	.640	.577	.374	.597
(7) $E=a+b\ln Y+cS+dS^2$	.109	.088	−.113	.232	.152	.226	.155	.591	.575	1.084	.447	.421	.377	.535
(8) $E=a+b\ln Y+c\ln S$	.102	.065	−.182	.222	.142	.214	.139	.598	.570	1.022	.455	.429	.376	.531
(9) $E=a+b\ln Y+c/S$	.091	.053	−.186	.211	.131	.207	.130	.504	.483	.806	.391	.372	.315	.445
(10) $\ln E=a+b/Y+c\ln S$	.097	.058	−.063	.164	.128	.202	.082	.794	.822	1.330	.717	.634	.480	.678
(11) $\ln E=a+b/Y+cS+dS^2$	.098	.066	−.048	.167	.131	.200	.087	.726	.751	1.278	.639	.560	.418	.627
(12) $\ln E=a+b/Y+c/S$	.077	.035	−.089	.142	.108	.187	.066	.741	.787	1.190	.698	.624	.468	.632
(13) $E=a+b/Y+c/S$	.026	.013	−.076	.067	.036	.065	.038	.530	.500	.770	.447	.412	.370	.481
(14) $E=a+b/Y+c\ln S$	.034	.021	−.072	.076	.044	.072	.046	.622	.587	.986	.507	.466	.427	.565
(15) $E=a+b/Y+cS+dS^2$	.038	.032	−.038	.080	.049	.077	.053	.610	.590	1.064	.488	.450	.417	.560

*Calculated at sample means. E is expenditure, Y is income, and S is household size.

[1] Includes only food consumed at home.

much as 100 percent. For all functional forms, the estimated household size elasticities varied by about 50 percent.

Comparing the estimated expenditure-income elasticities for the per capita models with the functional forms having income and household size as separate regressors (tables 2 and 4) provided additional insights. The expenditure-income elasticities obtained for the quadratic, double-log, semi-log, inverse, and log-inverse forms when expenditures and income were expressed in per capita terms usually exceeded the elasticities for these same functional forms when household size and income were treated as separate regressors. In addition, the sum of the expenditure-income and household size elasticities usually fell far below one.

Tables 5 and 6 present the rankings by functional form for the mean square error (lowest to highest) and correlation coefficients (highest to lowest), respectively. By both criteria, functional form (1)—expenditures as a function of income, income squared, household size, and household size squared—performed the best (produced the lowest mean squared error and highest

Table 5—Rankings of mean squared error statistics, 15 Engel functional forms

Expenditure item	Functional form*														
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Dairy products	1	7	3	10	8	13	2	4	9	14	12	15	11	6	5
Fats and oils	1	7	2	10	13	12	3	4	8	11	15	14	9	6	5
Flour and cereals	1	11	4	7	9	14	2	5	12	8	10	15	13	6	3
Beef and pork	2	5	1	10	12	9	4	3	6	14	15	13	11	7	8
Vegetables	2	5	1	11	12	10	4	3	6	14	15	13	9	7	8
Fruits	1	5	2	10	11	12	3	4	6	13	14	15	9	7	8
Total, food	1	11	2	5	6	7	3	4	14	10	12	13	15	9	8

*Numbers correspond to the equations in table 4.

Table 6—Rankings of correlation coefficients, 15 Engel functional forms

Expenditure item	Functional form*														
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Dairy products	2	13	5	6	1	4	3	7	14	12	9	10	15	11	8
Fats and oils	1	13	2	5	11	8	3	4	14	9	12	10	15	7	6
Flour and cereals	1	13	8	4	6	10	2	9	14	5	7	12	15	11	3
Beef and pork	3	8	2	4	7	1	6	5	9	12	13	10	15	11	14
Vegetables	2	7	1	6	8	3	5	4	9	13	14	10	15	11	12
Fruits	1	8	2	3	4	7	5	6	9	10	12	11	15	13	14
Total, food	2	13	3	1	5	6	4	7	14	11	12	9	15	10	8

*Numbers correspond to the equations in table 4.

Engel relationships which express expenditures and income in per capita terms may be too restrictive, as they force the sum of the income and household size elasticities to equal one.

correlation coefficient). Functional form (3)—expenditures as a function of income, income squared, and the natural logarithm of household size—also provided an above average fit to the data. Both of these functional forms produced only moderately different expenditure-income and household size elasticities at the sample means. The mean squared error statistics were above average for the linear logarithmic functional form for all food groups except total food, and flour and cereals (functional form 4, table 5).

CONCLUSIONS

The choice of the functional form dramatically affects estimated income and household size elasticities. Income elasticities derived from the inverse and log-inverse functional forms should be interpreted with caution, as, in this study, these forms provided very low income elasticities and poor statistical fits to the data. The double log usually provided the best statistical fit and also the highest income elasticity for models expressing per capita expenditure as a function of per capita income. The double-log fit poorly the flour and cereals expenditure data, which suggests that it is a poor choice when estimating Engel relationships for inferior commodities. The semi-log and quadratic functional forms provided better statistical fits to the data than the linear, inverse, or log-inverse functional forms.

When expenditures were expressed as a function of household size and income, the quadratic form having income, income squared, household size, and household size squared as explanatory variables provided the best statistical fit. For the 15 functional forms analyzed, the linear logarithmic functional form's fit to the data was about average. Thus, the double-logarithmic functional form seems appropriate when per capita expenditures are expressed as a function of per capita income. The linear logarithmic seems, however, to be a poor choice when income and household size are used as separate regressors in the Engel function. The estimated income and household size elasticities generated from the different models in which income and household size are treated as separate regressors suggest that the sum of these elasticities is usually different from one. Thus, Engel relationships which express expenditures and income in per capita terms may be too restrictive, as they force the sum of the income and household size elasticities to equal one.

REFERENCES

(*1*) Benus, J., J. Kmenta, and H. Shapiro, "The Dynamics of Household Budget Allocation to Food Expenditures." *Rev. Econ. and Stat.*, 58(1976):129-138.

(*2*) Brown, A., and A. Deaton, "Surveys in Applied Economics, Models of Consumer Behavior." *Econ. J.* 82(1972): 1,145-1,236.

(*3*) Buse, Rueben C., and Larry E. Salathe, "Adult Equivalent Scales: An Alternative Approach." *Am. J. Agr. Econ.* 60(1978):460-468.

(*4*) Chang, Hui-shyong, "Functional Forms and the Demand for Meat in the United States." *Rev. Econ. and Stat.* 59(1977): 355-360.

(*5*) Goreux, L. M., "Income and Food Consumption." *Monthly Bul. Agr. Econ. and Stat.* 9(1960):1-13.

(*6*) Hassan, Zuhair A., and S. R. Johnson, "Urban Food Consumption Patterns in Canada." *Agriculture Canada*. Pub. No. 77/1, Jan. 1977.

(*7*) Howe, Howard, "Cross-Section Application of Linear Expenditure Systems Responses to Sociodemographic Effects." *Am. J. Agr. Econ.* 59(1977):141-48.

(*8*) Leser, C.E.V., "Forms of Engel Functions." *Econometrica* 31-(1963):694-703 .

(*9*) Prais, S. J., and H. S. Houthakker,*The Analysis of Family Budgets.* Cambridge Univ. Press, Cambridge, 1955.

(*10*) Tomek, William G. "Empirical Analysis of the Demand for Food: A Review." Cornell Agr. Econ. Staff, Paper No. 77-8, Cornell Univ., Apr. 1977.

(*11*) Zarembka, Paul. "Transformation of Variables in Econometrics." *Frontiers in Econometrics*, Paul Zarembka, ed., Academic Press, New York, Apr. 1977, pp. 81-104.