CORNELL
AGRICULTURAL ECONOMICS
STAFF PAPER

THE IMPORTANCE OF HOUSEHOLD SIZE IN EXPLAINING LOW-FAT AND WHOLE MILK EXPENDITURES IN THE U.S.

by

Karim Laraki and Christine K. Ranney

March 1986 No. 86-4

Department of Agricultural Economics
Cornell University Agricultural Experiment Station
New York State College of Agriculture and Life Sciences
A Statutory College of the State University
Cornell University, Ithaca, New York, 14853

THE IMPORTANCE OF HOUSEHOLD SIZE IN EXPLAINING LOW-FAT AND WHOLE MILK EXPENDITURES IN THE U.S.

by

Karim Laraki

and

Christine K. Ranney*

*The authors are graduate student and assistant professor, respectively, in the Department of Agricultural Economics, Cornell University.

THE IMPORTANCE OF HOUSEHOLD SIZE IN EXPLAINING LOW-FAT AND WHOLE MILK EXPENDITURES IN THE U.S.

Important changes are occurring in the sociodemographic structure of the U.S. population. Average household size is decreasing and Black households represent an increasing proportion of all households. Simultaneously, household food consumption patterns are changing. Between 1960 and 1980, quantities of whole milk and low-fat milk sold dropped by 30% and increased by 900%, respectively.[1] The purpose of this paper is to analyze the effect of the new sociodemographic structure on low-fat and whole milk expenditures.

Two alternative models are usually adopted to estimate the effects of households sociodemographic characteristics on food expenditures. The per household expenditures model used by Boehm and Babb (1975) and Blaylock and Smallwood (1983) provides estimates of the effects of selected sociodemographic variables on the total expenditures of the household. The unit equivalent scales model used by Price (1970) and Buse and Salathe (1978) provides estimates of scale values that reflect the value of the consumption of each kind of individual in the household as a fraction of the value of the consumption of a base person. The sum of the scales for all the individuals in the household is used to obtain the value of the consumption of the household per base-person equivalent.

In both models the effect of the explanatory variables is often assumed to be independent of household size. The validity of this assumption is tested in this study with respect to low-fat and whole

[1] From Milk Facts, Milk Industry Foundation, 1984 Edition.

ables. To avoid these problems Tobit analysis (see Tobin, 1958 and McDonald and Moffit, 1980) is used.

In cross-section studies, it is usually argued that prices can be eliminated from the expenditure equations because prices are constant across households. This assumption is violated when quantity discounts exist. The per unit cost of milk is lower for large quantities than for small quantities. Hence, as household size increases (and quantities purchased increase), per unit cost decreases and prices are not constant across households. This problem can be avoided by specifying one expenditure equation for each household size. It, then, becomes more reasonable to assume that prices (the cost per unit of milk) are constant.

Two variables reflecting household resources, before-tax income and bonus stamps, are included as explanatory variables. We use before-tax income, rather than after-tax income, because it was reported more frequently. Purchase requirements for food stamps were still in effect when the data were collected. Eligible households had to pay part of the total value of their food stamp allotment to receive their stamps. The difference between the face value of the allotment and the amount paid for the stamps is the value of the bonus. The logarithm of bonus stamps and of income are used because income elasticities of food expenditures have been shown (Brown and Deaton) to be less than one and to decrease with income increases. They are specified separately because they may affect expenditures differently.

Specifying each source of purchasing power, income and bonus stamps, separately in the expenditure equations can lead to simultaneous equations bias. Food expenditures, Food Stamp Program participation,

Table 1. Variables E, I, Bonushat, and those for the age and sex composition have already been discussed. The other variables, race, urbanization, and region, have been shown to affect milk consumption in previous studies.

THE EFFECT OF HOUSEHOLD SIZE ON THE OTHER VARIABLES

In (1) an explanatory variable X has N estimated effects, one for each equation and hence for each household size. Whether the effect of X on milk expenditures differs by household size can be determined by comparing these effects across the N equations.

The model can be written in the general matrix form

$$E^n = X^n b^n + u^n, \ n=1,\ldots,N, \tag{2}$$

where E^n is a T^n x 1 vector representing the dependent variable, X^n is a T^n x K^n matrix representing the independent variables $X^n_1 \ldots, X^n_k$, b^n is a vector of K^n fixed but unknown parameters, u^n is T^n x 1 vector of random terms assumed to be normally distributed with mean zero and constant variance s_n^2. T^n is the number of observations in the sample of households of size n, and K^n is the number of regressors in the equation for the households of size n.

The test of equality of the effects of a variable k for all household sizes can be formulated by the null hypothesis

$$H_0: \ b^1 = b^2 = \ldots = b^n, \tag{3}$$

with the alternative hypothesis that at least two coefficients are different.

James (1951), developed a test statistic for comparing regression coefficients that are estimated from different populations. If the residuals in (2) are normally and independently distributed and if the sample is large, the statistic $h = \sum w_i (b^i)^2 - (\sum w_i b^i)^2/w$ has a Chisquare distribution with (n-1) degrees of freedom, where n is the number of coefficients being compared, $w_i=1/var[b^i]$, and $w = \sum w_i$. Hence, if the statistic h is greater than a critical value for a chosen level of significance, H_0 is rejected and at least two coefficients are different. This implies that household size has a statistically significant effect on the response to the chosen explanatory variable.

RESULTS

The estimated parameters of the whole milk and low-fat milk expenditure equations and selected summary statistics are presented in Tables 2 and 3. Race, region, education of the shopper, and the age of the two oldest members (for whole milk only), are the most significant variables.

The age of the two oldest members (particularly that of the second oldest member) have statistically significant, but opposite effects on whole milk expenditures. The latter increase with the age of the oldest member (significantly for households of size 2 and 4) and decrease with that of the second oldest. Except for household size 7-and-larger no age variables affect low-fat milk expenditures. The sex composition (proportion of males) and the interaction between the sex composition and the age of the members do not affect significantly the expenditures on either product.

Education of the shopper has a statistically strong, but small, effect on fresh milk expenditures. The effect is positive on low-fat milk and negative on whole milk. This effect was to be expected. Some authors (see Boehm, 1976) explain this tendency by the fact that households increasingly consider health factors in their food expenditure

TABLE 3

Tobit Estimates of the Low-fat Milk Expenditure Equations
for Family Sizes 1 to 7-and-Larger

Non-Normalized Tobit Coefficients
(Standard Errors in Parentheses)

Variable	1***	2***	3***	4***	5***	6*	≥7
Constant	-7.583** (4.069)	-6.313 (5.471)	-10.449* (7.924)	5.152 (13.279)	-53.178** (27.000)	-92.768** (53.088)	-81.275 (81.623)
Before-tax Income (Log form)	.473 (.301)	.215 (.280)	1.129** (.474)	-.015 (.649)	1.207 (1.030)	2.015 (2.308)	-1.501 (1.000)
Bonushat[a] (Log form)	.045 (.043)	.017 (.026)	.033 (.026)	-.016 (.033)	.037 (.041)	-.004 (.072)	-.016 (.078)
Logage1[b]	.252 (.415)	.183 (.689)	.121 (1.153)	-1.628 (2.209)	2.782 (4.049)	10.724 (8.623)	30.199*** (13.603)
Logage2		.073 (.427)	-.208 (.628)	1.102 (1.472)	1.202 (2.856)	-2.328 (5.527)	-18.825** (9.245)
Logage3			-.456 (.378)	-1.221 (1.026)	3.426* (2.376)	-2.962 (6.215)	-12.448 (10.248)
Logage4				.197 (.612)	-1.756 (2.035)	.433 (5.532)	53.921** (25.052)
Logage5					.036 (.958)	3.720 (3.751)	-32.164** (16.876)
Logage6						.700 (2.031)	-24.797 (10.626)
Loage7							8.444** (4.537)
Interaction[c]	-1.020* (.647)	.027 (.705)	-.005 (.506)	.873 (.696)	-1.024 (1.093)	-1.922 (2.281)	-1.007 (3.182)
Prop. of Males	6.277* (4.000)	-.926 (8.512)	-1.178 (8.347)	-16.714 (14.859)	27.863 (28.236)	60.972 (73.003)	-1.173 18.068)
Education	.037* (.026)	.075** (.024)	.079** (.045)	.104** (.062)	.243** (.114)	.373** (.183)	1.938*** (.490)
RACE: Black	-1.982** (.828)	-1.870*** (.597)	-1.698** (.900)	-5.428** (2.321)	-2.137 (2.162)	-24.230 d.	-9.005* (6.955)
Nonblack/ Nonwhite	-1.460* (1.097)	-1.827* (1.359)	-19.166 d.	-4.248** (2.578)	-.576 (2.703)	-.392 (5.127)	-39.168 d.
URBANIZATION: Suburban	-.300 (.359)	-.067 (.287)	.949** (.506)	-.608 (.669)	2.045** (1.226)	2.516 (1.986)	7.757** (3.724)
Central City	-.356 (.317)	-1.58 (.317)	.944* (.596)	-2.158** (.853)	-.861 (1.690)	.627 (2.519)	-5.583 (4.742)
REGION: North East	-.725** (.390)	-1.438*** (.383)	-3.490*** (.794)	-7.365*** (.922)	-2.303 (1.843)	9.417** (2.530)	(5.667)
North Central	-.383 (.369)	.466** (.345)	-.411 (.599)	.093 (.791)	-.906 (1.379)	2.068 (2.408)	18.299*** (5.268)
South	-.868** (.387)	-.761** (.365)	-2.165*** (.637)	-3.822*** (.872)	-7.730*** (1.632)	-7.078** (3.196)	-30.685** (18.376)

SUMMARY STATISTICS

Variable	1***	2***	3***	4***	5***	6*	≥7
Standard Error	1.862	2.426	3.203	4.377	5.418	5.549	6.027
% of zero observations	84.1	76.4	73.5	72.5	74.1	75.9	79.3
Total observations	473	762	430	411	232	112	67

The likelihood ratio test for the full equation as compared to the equation restricted to the effect of the constant is significant at the +++=.5% level, ++=10% level, +=20% level.

*** Significantly different from 0 at the .5% level.
** Significantly different from 0 at the 5% level.
* Significantly different from 0 at the 10% level.
a Predicted value of the bonus.
b The age variables are in months, ordered from the oldest member (age_1) to the youngest (age_7).
c Interaction = Propmale.$\sum Log(Age_i)$
d The standard errors are unusually high (over 500). Refer to text (p.10) for discussion of the problem.

for household size 7 and larger. The last equation, therefore, is not considered in the following test statistics.

THE EFFECT OF HOUSEHOLD SIZE ON THE RESPONSE TO SELECTED EXPLANATORY VARIABLES

The estimated parameters seem to vary across equations. The first important difference to notice is how the standard errors of the equations increase with household size. Using Bartlett's test (see Intriligator 1978, p. 157) the hypothesis of equal variances across equations is strongly rejected for the low-fat and the whole milk expenditure equations.

The test for the effect of household size is performed on the variables that significantly affect fresh milk expenditures (age of the two oldest members, education of the shopper, effect of black households as opposed to that of white households, and effects of households in the Northeast and the South as compared to the West). For whole milk expenditures the hypothesis of equal effects across household sizes was rejected at the 1% level for education and the two regional dummy variables (Northeast and South). For low-fat milk expenditures the hypothesis is rejected at the 1% level for the two regional dummy variables and for the effect of Black households. The absolute effect of the variables increases with household size, significantly for education (for whole milk), race (for low-fat milk), and the regional dummy variables (for both products), and insignificantly for the age variables.

CONCLUSIONS

The response to age, education of the shopper, race, and region, are the most significant variables for explaining fresh milk expenditures. The response of whole milk expenditures to the education and

REFERENCES

Blaylock, J.R. and Smallwood, D.M. (1983), "Effects of Household Socio-economic Features on Dairy Purchases", USDA, ERS, Technical Bulletin, No. 1686, August.

Boehm, W.T. (1976), "The Household Demand for Fluid Milk Consumption", Research Division Bulletin, Virginia Polytechnic Institute and State University, No. 120.

Boehm, W.T. and Babb, E.M. (1975), "Household Consumption of Beverage Milk Products", Department of Agricultural Economics, Agricultural Experiment Station, Purdue University, Station Bulletin 75, March.

Brown, A. and Deaton, A. (1972), "Surveys in Applied Economics: Models of Consumer Behavior", The Economic Journal, December, pp. 1186-1236.

Bryant, W.K. (1985), "Durable and Wives' Employment Yet Again", Mimeograph, New York State College of Human Ecology, Cornell University.

Buse, R.C. and Salathe, L.E. (1978), "Adult Equivalent Scales: An Alternative Approach" American Journal of Agricultural Economics, August, pp. 460-468.

Intriligator, M.D. (1978), Econometric Models, Techniques, and Applications, Prentice Hall: Englewood Cliffs, New Jersey.

James, G.S. (1951), "The Comparison of Several Groups of Observations When the Ratios of the Population Variances are Unknown", Biometrika, Vol. 38, pp. 324-329.

McDonald, J.F. and Moffitt, R.A. (1980), "The Use of Tobit Analysis" The Review of Economics and Statistics, Vol. 2, May, pp. 318-321.

Price, D.W. (1970), "Unit Equivalent Scales for Specific Food Commodities" American Journal of Agricultural Economics, May, pp. 224-223.

Rosenzweig, M. and Wolpin, K. (1980), "Life-Cycle Labor Supply and Fertility: Causal Inference from Household Models", Journal of Political Economy, Vol. 88, No. 2, April, pp. 328-348.

Tobin, J. (1958), "Estimation of Relationships for Limited Dependent Variables" Econometrica, Vol. 26, No. 1, January, pp. 24-36.