Analyzing Fresh Vegetable Consumption From Household Survey Data

Anderson Reynolds

Abstract

To analyze fresh vegetable consumption using household survey data, the tobit model and a more flexible parameterization to the tobit model—the "double hurdle" model—were considered. Based on the likelihood ratio test, the tobit model was rejected against the "double hurdle" specification. Moreover, the results suggest that the tobit model underestimated the impact of the explanatory variables on fresh vegetable expenditures. Other results indicate that total food expenditures (proxy for income), age, household composition, sex, race, marital status, urbanization, region, and seasonality are all important determinants of fresh vegetable expenditures.

Key words: tobit, double hurdle, log likelihood, socioeconomic, demographic, consumption

Data generated from household expenditure surveys are usually characterized by a large proportion of the households reporting zero expenditures, thus rendering standard regression methods inappropriate for conducting the consumption analysis. Recognizing this, researchers have commonly employed the tobit model to perform the analysis. However, the tobit model is just one among several censored regression models that can be used to model consumption behavior. The specification of an appropriate model depends on the phenomenon that is assumed to give rise to the zeros.

The tobit model assumes that zero expenditures are observed when desired consumption is nonpositive; thus the dependent variable is truncated at zero. A second explanation for the occurrence of zero expenditures is provided by Cragg, who recognized that although the household may desire a positive amount of the good, impediments to acquisition (such as availability of the good and transaction cost) may effectively prohibit purchases. In conformity to this explanation, Cragg proposed a more flexible parameterization to the tobit model, which he termed the "double hurdle" model. A third explanation is embodied in the "purchase infrequency" model. The model is based on the proposition that, in the case of infrequently purchased goods, zero expenditures may have been recorded because the household purchased a stock of the good outside of the survey period.

In analyzing adult women's consumption of ten food groups, Haines *et al.* compared the tobit model with Cragg's "double hurdle" model. The hypothesis that the tobit model was correctly specified against the alternative, Cragg's "double hurdle" model, was rejected for nine of those food groups. Similarly, in the case of the United Kingdom's clothing consumption, Blundell and Meghir rejected the tobit specification against the "purchase infrequency" model.

Despite the demonstrated possibility that the tobit model may be a misrepresentation of households' underlying consumption behavior, studies that have employed the tobit model to analyze fresh vegetable consumption (Huang *et al.*; Capps and Love; Smallwood and Blaylock; Blaylock and Smallwood) have failed to consider alternative specifications. Because fresh vegetables may be considered a frequently purchased item, it is conceivable that the purchase infrequency model does not apply. However, the "double hurdle" model appears to be a viable alternative to the tobit model.

This study provides an analysis of the impact of socioeconomic and demographic variables on U.S. consumption of fresh vegetables. Statistical testing is performed to determine which specification—the tobit or the "double hurdle" model—is most consistent with the underlying consumption behavior.

Information on the impact or relative impact of various socioeconomic factors on the consumption of fresh vegetables can benefit both producers and consumers and may facilitate the decision making of policy makers. For example, such information can enable the industry to focus its limited advertising dollars on the subsector of the population most likely to respond favorably. In addition, the informa-

Anderson Reynolds is a Postdoctoral Fellow in the Department of Agricultural Economics at the University of Guelph; Guelph, Ontario. The author would like to express thanks to J.S. Shonkwiler for his assistance in the development of the paper.

tion can be used to forecast or to project consumer expenditures as the explanatory variables change over time, thus enabling the industry to adjust in an appropriate manner. Moreover, knowledge of how socioeconomic variables affect food consumption patterns can allow policy makers to anticipate the dietary effects of assistance programs such as food stamps.

MODEL SPECIFICATION

The tobit model as developed by Tobin may be specified as follows:

$$
(1)\quad \begin{aligned} y_i^* &= x_i\beta + e_i \\ y_i &= y_i^* \quad \text{if } y_i^* > 0 \\ &= 0 \quad \text{otherwise} \end{aligned}
$$

where, with regard to the analysis at hand, y_i is the ith individual household's observed expenditures on fresh vegetables, y_i* is the desired or optimal consumption (in monetary terms) level of that household and can be construed as the solution to a utility maximization problem, x_i represents a vector of socioeconomic and demographic variables that characterizes the household's preferences and/or influences the household's purchasing behavior, and the error term, e_i, is assumed to be independently normally distributed with zero mean and constant variance, σ^2. According to this specification, observed expenditures are equal to the desired expenditure level if desired expenditures are greater than zero; otherwise zero expenditures are observed. Desired expenditures, y_i^*, can take on negative values; however, values of y_i^* less than zero are unobserved, hence, y_i is censored at zero.

The log likelihood for equation (1) has the form

$$
(2)\quad \begin{aligned} \text{Log L} &= \sum_0 \log(1 - \Phi_i(w)) - \frac{1}{2}\sum_1 \log 2\pi \\ &- \frac{1}{2}\sum_1 \log \sigma^2 - \sum_1 \frac{(y_i - x_i\beta)^2}{\sigma^2} \end{aligned}
$$

where Φ_i denotes the standard normal distribution function evaluated at $w_i = \frac{x_i\beta}{\sigma}$ and the sumation indexes refer to the limit and nonlimit observations. The first term on the right-hand side of equation (2) is the contribution of the limit observations to the log likelihood function, while the remaining terms represent the contribution of the nonlimit observations. The first and second derivatives with respect to β and σ^2 may be found in Amemiya.

Cragg's "double hurdle" model generalizes the tobit model in that it recognizes that, although the household may desire a positive amount of the good, impediments to purchase may prohibit consumption. This recognition led to the modeling of consumption behavior in two stages: first, based on impediments to acquisition, the household decides whether or not to purchase the good, and second, according to the intensity of the desire for the good, the household decides on how much to purchase. The "double hurdle" model is represented as

$$
(3)\quad \begin{aligned} y_i^* &= x_i\beta + e_i \\ D_i &= z_i\theta + v_i \\ y_i &= y_i^* \quad \text{if } D_i > 0 \\ &= 0 \quad \text{otherwise} \end{aligned}
$$

where y_i and y_i* are previously defined, and D_i characterizes the decision of whether to purchase. It is assumed that only the sign of D_i is observed and that y_i* is observed only when D_i is positive. The vectors of independent variables need not be different, and the error terms (e_i,v_i) are assumed to be independently normally distributed with zero means and constant variances (σ^2,1). This specification pinpoints the essential difference between the tobit model and the "double hurdle" model. In the tobit model the same variables (x_i) and parameters (β_i) explain the decision of whether to purchase and of how much to purchase. In contrast, the "double hurdle" model allows different sets of variables (x_i,z_i) and parameters (β, θ) to characterize the two decisions.

The log likelihood function for the "double hurdle" model follows as

$$
(4)\quad \begin{aligned} \text{Log L} &= \sum_0 \log(1 - \Phi_i(m)) + \sum_1 (\log \Phi_i(m) \\ &- \log \Phi_i(w)) - \frac{1}{2}\sum_1 \log 2\pi - \frac{1}{2}\sum_1 \log \sigma^2 \\ &- \sum_1 \frac{(y_i - x_i\beta)^2}{\sigma^2} \end{aligned}
$$

where $m = z_i\theta$. The first derivatives with respect to the parameters (β, θ, and σ^2) can be found in Reynolds (p. 39).

Given equations (2) and (4) and their corresponding derivatives, maximum likelihood estimates of the tobit and "double hurdle" models can be obtained via the method of scoring or the modified method of scoring (Judge *et al.*).

ESTIMATION AND DISCUSSION

The data for the study were obtained from the 1984 Consumer Expenditure Survey sponsored by the

Bureau of Labor Statistics. The data relate to fresh vegetable (excluding potatoes) expenditures by U.S. households during the months of March, April, May, June, November, and December. The commodity—fresh vegetables—and the specific six months chosen for the analysis represent part of an ongoing research project to study the demand for Florida's fresh vegetables during its major production months. The resulting sample consisted of 3368 households (observations). Of these, 1088 reported no fresh vegetable expenditures. This significant portion of observations on fresh vegetable expenditures (the dependent variable) taking a zero value provides justification for considering censored regression models as an appropriate framework for conducting the present investigation.

Other than household income, traditional economic theory generally does not give specific indications of the variables (variables that comprise the vector x_i) to include in the specification of an Engel curve. Consequently, logic, results of past studies, and, to a limited extent, economic theory are used to guide the selection of explanatory variables. To begin with, household production theory would suggest that variables characterizing labor market participation (hours of work, for example) should influence fresh vegetable consumption. This is expected because labor market participation, in part, reduces the amount of time available to the household for the transformation of fresh vegetables to meal items, ultimately constraining the household production function and hence the household's fresh vegetable expenditures. Household size is another variable that can be expected to influence consumption. Apart from the fact that larger households will generally need more food than smaller households, household size introduces economies of scale into consumption. The family life cycle hypothesis provides justification for including household age composition. According to the life cycle concept, biological and psychological changes associated with aging give rise to changing nutritional needs. Thus, the age of household members can be expected to influence food consumption patterns. For similar reasons the sex of household members can be expected to affect food intake. The educational level of the household head can also be anticipated to influence consumption, provided that the level of education affects the dietary choice of the meal planner. Due to differences in tradition, environment, and opportunities (availability of certain goods) associated with rural or urban location or regional differences, the location of the household is likely to have an impact on its consumption pattern. Varying traditions and consumption habits among races can also influence current and future consumption patterns. The results of past studies (Salathe; Huang *et al.*, Capps and Love; Matsumoto; Smallwood and Blaylock; Blaylock and Smallwood) suggest that most of these variables do have an impact on fresh vegetable consumption.

Table 1 provides a description of the variables included in the analysis. The variables described by "if" statements were one-zero variables. Average weekly household fresh vegetable (excluding potatoes) expenditures were used as the dependent variable. The independent variables include total household food expenditures, household size and household size squared, urbanization, the age, sex, race, education, and marital status of the household head, age distribution of the household, the region in which the household is located, and the months during which the household was surveyed. Obtaining reliable income data on individual households can be quite elusive. For example, some households in the sample did not provide complete information on their incomes. To circumvent this problem, total food expenditure was used in lieu of household income.

Apart from the included explanatory variables, variables such as the number of earners in the household and hours per week the household head worked, designed to characterize the household's labor force participation, were considered but found to be statistically insignificant. In addition, low order polynomials involving food expenditures, family size, and age were considered, but the insignificant coefficients associated with these variables implied that the interactive effects among these different variables were minimal.

The estimation results for the tobit and "double hurdle" models are presented in Table 2. Gauss (Edlefsen and Jones), a micro computer software programing language, was used to conduct the estimation. Since both the first and second analytical derivatives of the log likelihood function of the tobit model are easily obtained, maximum likelihood estimates of the tobit model were obtained via the method of scoring that uses the first and second derivations of the log likelihood function. Second derivatives of the log likelihood of the "double hurdle" model are not as easily derived; therefore, the modified method of scoring that uses only the first derivatives was utilized. Least squares estimates were used as starting values for β, while estimates generated from a probit among observations above and below the limit provided starting values for θ. Recall that in the tobit model both the decision of whether to purchase and how much to purchase are captured in the β parameters, while in

Table 1. Variable Definitions and Mean Values

Variable	Mean	Definition
Dependent Variable	1.5132	Weekly fresh winter vegetable (excluding potatoes) expenditures ($)
(Food Expenditure)$^{1/2}$	6.2798	Square root of total food at home expenditure ($)
Household Size	2.6113	Number of household occupants
(Household Size)^2	6.8189	Household size squared
Age	46.6093	Age of reference person
Sex	0.6698	= 1 if reference person is male
Race		
White	0.8548	Omitted base group
Black	0.1146	= 1 if reference person is black
Nonwhite/Nonblack	0.0306	= 1 if reference person is nowhite/nonblack
Education	0.7289	= 1 if reference person completed H.S.
Marital Status	0.5751	= 1 if reference person is married
Urban	0.8925	= 1 if household resides in urban area
Region		
Northeast	0.3124	Omitted base region
Midwest	0.2360	= 1 if household resides in the MW
South	0.2369	= 1 if household resides in the South
West	0.2147	= 1 if household resides in the West
Season	0.4486	= 1 if household was surveyed during the winter months of November and December
Household Composition		
Children < 5	0.0221	Proportion of household 0-4 yrs old
Children 5 to 13	0.0835	Proportion of household 5-13 yrs old
Persons 14 to 24	0.1866	Proportion of household 14-24 yrs old
Persons 25 to 44	0.3021	Omitted base group
Persons 45 to 64	0.2346	Proportion of household 45-64 yrs old
Persons > 65	0.1712	Proportion of household over 65 yrs old

the "double hurdle" model the decision of whether to purchase is embodied in θ, and β embodies the second decision of how much to purchase.

Because the tobit model is nested with respect to the "double hurdle" model, the likelihood ratio test can be used to test the tobit specification against the "double hurdle" model. Specifically, the "double hurdle" model is reduced to the tobit model when $\theta = \frac{\beta}{\sigma}$. Thus the nested test involving the two models is a test of the null hypothesis that $\theta = \frac{\beta}{\sigma}$. To test this hypothesis, the likelihood ratio test statistic, which is distributed asymptotically as a chi-square with 20 degrees of freedom, was calculated and found to be 529. Comparing this computed value with a critical chi-square statistic value at conventional probability levels leads to a rejection of the null hypothesis that the restrictions embodied in the tobit model are valid. This conclusion suggests that the "double hurdle" model is a better representation of the data generating process. Consequently, the remaining analysis focuses on the results of the "double hurdle" model.

RESULTS

With the exception of the sex and the household composition variables associated with the proportion of persons in the household between 45 and 65 years of age, the signs of the β coefficients were uniform across the tobit and "double hurdle" models. However, except for the marital status variable, the coefficients of all the variables whose signs were uniform across the two models were consistently larger in absolute value in the "double hurdle" model than in the tobit model. For example, the β coefficient associated with persons over 65 years in the "double hurdle" model was 31 times the size of its counterpart (in absolute value) in the tobit model. Similarly, the coefficients associated with children

Table 2. Estimated Tobit and "Double Hurdle" Models of Fresh Vegetable Expenditures

Independent Variable	Tobit Model	Double Hurdle Model	
	β_i	β_i	θ_i
Constant	-3.4623 (0.2761)[a]	-12.5874 (1.0694)	-1.4435 (0.2026)
$(\text{Food Exp.})^{1/2}$	0.7401 (0.0192)	1.4037 (0.0529)	0.3874 (0.0136)
Household Size	-0.2676 (0.1204)	-0.5657 (0.3602)	0.0227 (0.0935)
$(\text{Household Size})^2$	0.0238 (0.0128)	0.0466 (0.0342)	-0.0106 (0.0114)
Age	0.0089 (0.0057)	0.0405 (0.0179)	-0.0020 (0.0041)
Sex	-0.3489 (0.1044)	0.2044 (0.3478)	-0.3673 (0.0731)
Black	0.1784 (0.1331)	0.8692 (0.5035)	-0.0177 (0.0910)
Nonwhite/Nonblack	1.4836 (0.2177)	3.1078 (0.4557)	0.3453 (0.1792)
Education	0.1426 (0.0929)	0.4393 (0.2928)	0.0522 (0.0681)
Marital Status	0.3153 (0.1317)	0.1223 (0.4279)	0.3049 (0.0933)
Urban	0.4474 (0.1451)	1.9717 (0.4662)	0.0212 (0.1066)
Midwest	-0.2501 (0.1160)	-0.9518 (0.3744)	-0.0454 (0.0810)
South	-0.0958 (0.1164)	-0.8377 (0.3687)	0.0642 (0.0824)
West	0.1859 (0.1190)	0.2956 (0.3590)	0.1266 (0.0861)
Season	-0.3724 (0.0770)	-1.1025 (0.2517)	-0.1628 (0.0540)
Children < 5	-1.0018 (0.5374)	-4.2207 (1.9484)	-0.1862 (0.3923)
Children 5 to 13	-0.7699 (0.3219)	-1.5978 (0.9292)	-0.4352 (0.2295)
Persons 14 to 24	-0.3635 (0.1752)	-1.3906 (0.6022)	-0.1703 (0.1134)
Persons 45 to 65	0.0895 (0.2016)	-0.6074 (0.6406)	0.2823 (0.1438)
Persons > 65	-0.0689 (0.2860)	-2.1951 (0.9251)	0.4740 (0.2043)
Variance	4.1119 (0.1245)	9.6261 (0.5141)	
Log Likelihood	-5437.9	-5173.3	

[a] Calculated standard errors appear in parentheses.

less than 5 years old, the southern region, urbanization, age and black households were all at least 4 times the size of the corresponding coefficients in the tobit model. This implies that the tobit specification underestimated the impact of the explanatory variables on the household's decision on how much to spend on fresh vegetables.

Total food expenditures, the proxy for household income, appeared to be an important factor in both the decision on whether to purchase (θ) and on how much to purchase β. The associate θ and β coefficients were both positive, and they were more than twice the size of their standard errors.

The results indicated that household size had a positive impact on the decision of whether to purchase, but a negative impact on the decision of how much to purchase. However, the associated θ coefficient was smaller than the corresponding standard error, and the β coefficient was less than twice the size of its standard error. In contrast, the θ coefficient associated with the square of household size was negative while the β coefficient had a positive value. Once again, the coefficients were less than twice the size of their corresponding standard errors. *A priori* and in conformity with most other studies, household size was expected to have a positive impact on vegetable expenditures (indicating that large households tend to consume more than their smaller counterparts), while a negative sign was expected for the coefficient of the household size square variable (indicating economies of scale in consumption). However, only the θ coefficients, which embodied the decision on how much to purchase, conformed with such expectations.

The age of the household head appeared to have a positive and significant impact on the household's decision on how much fresh vegetables to purchase. The associated β coefficient was positive and was more than twice the size of its standard error. In contrast, based on the θ coefficient, the age of the household head seemed to have a negative but insignificant impact on the household's decision regarding whether to purchase fresh vegetables. Household age composition also appeared to affect fresh vegetable expenditures. All the β coefficients associated with household composition had negative signs, implying that households whose compositions were skewed toward members of ages 25 to 44 years were more predisposed to spending on fresh vegetables than households whose compositions were skewed toward other age groups. The β coefficients associated with children less than 5 years, persons 14 to 24 years, and persons more than 65 years were all at least twice the size of their standard errors; that associated with persons 5 to 13 years was over 1.5 times the size of the corresponding standard error; and the coefficient of persons 45 to 65 years was slightly less than its standard error. According to the θ coefficients, households with children less than 5, children 5 to 13 years, and persons 14 to 24 years were less likely to purchase fresh vegetables than households with persons 25 to 44 years of age. However, the coefficients of all these variables were less than twice the size of their corresponding standard errors. In contrast, the signs of the θ coefficients associated with persons 45 to 65 years and persons more than 65 were positive and larger than twice the size of their corresponding standard errors, implying that the presence of these two groups in a household predisposed the decision to purchase fresh vegetables more than the presence of persons 25 to 44 years old. The negative signs of the β coefficients indicated that, once the household had decided to purchase fresh vegetables, it tended to purchase a greater amount if the household was predominantly composed of persons 25 to 44 years old as opposed to persons greater than 44 years of age.

The results indicated that, while a household headed by a female was more likely to decide to purchase fresh vegetables than a male-headed household, a male-headed household was predisposed to a greater expenditure outlay than a female-headed household. However, since the associated θ coefficient was over twice the size of the corresponding standard error, while that of the β coefficient was smaller than its standard error, the first effect seemed more important.

With regard to race, blacks were less likely than whites to decide to purchase fresh vegetables, but once the purchasing decision was made, blacks tended to spend more on fresh vegetables than whites. Note, however, that both the β and θ coefficients were less than twice the size of their corresponding standard errors. In comparison, the results clearly suggest that nonwhites/nonblacks as a group were more likely than whites to decide to purchase fresh vegetables. Also, they had a tendency to spend a great amount on fresh vegetables. The associated β and θ coefficients were both more than twice the size of their corresponding standard errors.

The signs of the β and θ coefficients associated with the education variable were positive, indicating that high school graduates were more likely to purchase fresh vegetables and at greater expenditure levels than nongraduates. However, that difference between high school graduates and nongraduates did not appear to be significant.

The marital status of the household appeared to have a significant positive impact on the decision of whether to purchase fresh vegetables. The sign of θ was positive and more than three times the size of the corresponding standard error. The impact on expenditure levels, though positive, was insignificant.

The location variables appeared to influence expenditure levels significantly, but not the decision on whether to purchase (the coefficients of the associated θ coefficients were all less than twice the size of their corresponding standard errors). According to the β coefficient associated with the urban variable, urban households tended to spend significantly more on fresh vegetables than their rural counterparts. The coefficient was over four times the size of the corresponding standard error. The greater incidence of home gardens in rural areas may partly account for that result. The signs of the β coefficients associated with the Midwest and the South were negative and twice the size of their corresponding standard errors, while that of the West variable was positive and less than twice the size of its standard error. This implies that households located in the Midwest and the South tended to spend significantly less on fresh vegetables than households located in the Northeast, while households located in the West spent an insignificantly greater amount on fresh vegetables than their northeastern counterparts.

Apparently, seasonality influenced fresh vegetable consumption. Both the β and θ coefficients associated with the season variable were negative and were at least twice the size of their corresponding standard errors. This implies that households surveyed in November or December, as opposed to the other four months, were less likely to decide to purchase fresh vegetables, and when they decided to purchase fresh vegetables they were likely to purchase a smaller amount.

Promotional programs and advertising campaigns are commonly used as means to expand food consumption. In view of limited funds, the success of such promotional efforts may depend on targeting the population with the greatest potential or tendency to consume the food item in question. In that respect, the above results may have important implications for the fresh vegetable industry. The study suggested that fresh vegetable consumption was positively related to education, income, and age. Urban dwellers tended to spend more on fresh vegetables than their rural counterparts. Households located in the West had a greater tendency to spend on fresh vegetables than those located in other parts of the country. Households whose occupants were nonwhite/nonblack tended to spend more on fresh vegetables than households whose occupants were either black or white.

SUMMARY

This study provides further empirical evidence that the tobit model may, in some cases, be an inappropriate representation of consumers' underlying consumption behavior. In the case of fresh vegetable consumption, the tobit specification was rejected in favor of the more flexible parameterization represented by Cragg's "double hurdle" model. Unlike the tobit model, the "double hurdle" model allows different sets of variables and parameters to embody the two-step decision of whether to purchase and how much to purchase. The results of the "double hurdle" model indicated that, indeed, the same variable may impact these two decisions differently in terms of direction, magnitude, and significance level. The results also indicate that income, age, household composition, sex, race, marital status, urbanization, regional location, and seasonality had significant impacts on fresh vegetable consumption.

REFERENCES

Amemiya, T. "Tobit Models: A Survey." *Econometrics.* 24(1984):3-61.

Berndt, E., B. Hall, and J. Hausman. "Estimation and Inference in Nonlinear Structural Models." *Ann. Econ. Social Meas.* 3(1974):653-665.

Blaylock, J.R., and D.M. Smallwood. "U.S. Demand for Food: Household Expenditures, Demographics, and Projections." Washington, DC: USDA ERS Technical Bulletin No. 1713, February, 1986.

Blundell, R., and C. Meghir. "Bivariate Alternatives to the Tobit Model." *J. Econometrics.* 34(1987):1979-2000.

Capps, Jr., O., and J.M. Love. "Determinants of Household Expenditure on Fresh Vegetables." *So. J. Agr. Econ.* 15(1983):127-132.

Cragg, J.G. "Some Statistical Models for Limited Dependent Variables with Applications to the Demand for Durable Goods." *Econometrica.* 39(1971):829-844.

Edlefsen L.E., and S.D. Jones. *Gauss Version 1.28* Washington: Applied Technical Systems, Inc., 1985.

Haines, P.S., D.K. Guilkey, and B.M. Popkin. "Modeling Food Consumption Decisions as a Two Step Process." *Am. J. Agr. Econ.* 70(1988):543-552.

Huang, C.L., S.M. Fletcher, and R. Raunikar. "Modeling the Effects of the Food Stamp Program on Participating Households' Purchases: An Empirical Application." *So. J. Agr. Econ. 13(1981):27-33.*

Judge, G., W. Griffiths, R. Hill, H. Lutkepohl, and T. Lee. *The Theory and Practice of Econometrics.* New York: Wiley, 1985.

Matsumoto, M. "Interregional Variation in Food Expenditure Patterns of Low-Income Households." Washington, DC: USDA Economic Research Service National Economics Division, June, 1984.

Reynolds, A. "An Econometric Analysis of Fresh Winter Vegetable Consumption: Extensions of the Tobit Model." Unpublished Ph.D. Dissertation, University of Florida, 1989.

Salathe, L.E. "Household Expenditure Patterns in the United States." USDA Econ. Stat. Coop. Service Technical Bulletin No. 1603, April, 1979.

Smallwood, D.M., and J.R. Blaylock. "Household Expenditures for Fruits, Vegetables and Potatoes." Washington, DC: USDA ERS Technical Bulletin No. 1650, May, 1984.

Tobin J. "Estimation of Relationships for Limited Dependent Variables." *Econometrica. 26(1958):24-36.*

U.S. Department of Labor, Bureau of Labor Statistics. *Consumer Expenditure Survey: (Interview Survey/Dairy Survey),* Washington DC: U.S. Government Printing Office, 1980-85.