

Household Demand for Fresh Potatoes: A Disaggregated Cross-Sectional Analysis

Thomas L. Cox, Rod F. Ziemer and Jean-Paul Chavas

A model of household fresh potato consumption incorporating prices, income, family size and other socioeconomic effects is estimated by maximum likelihood Tobit procedures. The effects of truncation bias due to non-purchasing households are evaluated and decompositions of the Tobit elasticities are performed for various sub-groups of the data. The market development implications of this type of disaggregated commodity analysis are explored.

The use of socioeconomic variables to augment the more traditional money income specification of household food expenditure functions from cross-section data has been increasingly accepted. Price *et al.* and others have focused upon expenditure function analysis for broad food aggregates incorporating socioeconomic and demographic factors. Adrian and Daniel, Allen and Gadson and others have focused upon the impact of household socioeconomic characteristics on selected food nutrients. In more commodity specific frameworks, Price *et al.* (fruits and vegetables), Huang *et al.* (whole and low-fat fresh milk), and Capps and Love (fresh vegetables) have demonstrated the explanatory usefulness of socioeconomic variables for disaggregated cross-sectional analysis.

Most food consumption analysis of cross-section data derives from traditional con-

sumer demand theory, where demand is a function of own price, the prices of substitutes and complements, income, and household size. This traditional specification is then commonly augmented with socioeconomic variables as proxies for household taste characteristics. Unfortunately, much cross-section data, particularly from Bureau of Labor Statistics (BLS) sources, contain only household expenditures and socioeconomic information. Given the lack of price information, prices are generally not included in cross-sectional analysis (e.g., West and Price; Buse and Salathe).

Recent cross-section data sources such as the USDA's 1977-78 Nationwide Food Consumption Survey (NFCS), contain detailed information on household socioeconomic characteristics as well as food expenditures and their corresponding quantities consumed for the survey week. The inclusion of both quantities consumed and expenditures can then be used to derive commodity price information. If one considers disaggregated commodity prices from the 1977-78 NFCS it is observed that prices are generally anything but constant in this cross-section demand data. Assuming that the structure of commodity demand is constant over the survey period and that regional and quarterly differences in cross-sectional prices

The authors are respectively Research Associate and Assistant Professor, Department of Agricultural Economics, Texas A&M University, College Station, Texas, and Associate Professor of Agricultural Economics, University of Wisconsin, Madison, Wisconsin.

The authors acknowledge the helpful comments of John P. Nichols, Michael K. Wohlgenant, the editors and anonymous *Journal* reviewers.

Texas Agricultural Experiment Station, Texas A&M University, Technical Article No. 19264.

reflect commodity supply forces, then these prices can be hypothesized to identify a commodity demand curve as in time-series data. Therefore, it appears that the NFCS can provide a basis for investigating the role of prices and substitution relationships in disaggregated, cross-sectional demand analysis.

The primary objective of this paper is to investigate price effects in disaggregated cross-sectional analysis of food consumption. This objective is accomplished by developing a household consumption model for the commodity fresh potatoes. The analysis is based on households surveyed over a week period in the 1977-78 NFCS and located in the Western region of the United States. By including the prices of commodities that are close substitutes for fresh potatoes (canned or frozen, and dehydrated potatoes), the model provides useful information concerning issues such as substitutions between product form (e.g., fresh versus processed vegetables). A secondary objective of the paper is to estimate the influence of a number of socio-demographic variables on household fresh potato consumption.

A problem with increased disaggregation is that a number of households become less likely to consume particular commodities during the one week survey period, possibly leading to a large number of zero valued observations on the dependent consumption variable. Tobin has shown that traditional least squares regression analysis based on a sample characterized by a truncated dependent variable can lead to biased and inconsistent parameter estimates. In this paper, asymptotically efficient estimates of model parameters and associated elasticities are obtained through maximum likelihood (ML) procedures. To provide some evidence on the importance of truncation bias, these estimates are compared and contrasted to the traditional ordinary least squares (OLS) estimates obtained from the

full sample (i.e., all households) as well as from the truncated sample (i.e., including only purchasing households). The ML estimates are then discussed and interpreted in the context of a disaggregated demand analysis. Implications of the results for market development strategies are explored.

The plan of the paper is as follows. Section 2 discusses theoretical considerations from previous research and the empirical model. Section 3 then describes the data and estimation procedures. Section 4 presents the results from the alternative estimation procedures and their implied elasticities. Conclusions are offered in Section 5.

Theoretical Considerations

Traditional consumer theory assumes that consumption units (households) attempt to maximize utility from the services of goods purchased in the marketplace subject to a money income constraint. This motivates the inclusion of prices and income in demand specifications. However, numerous non-market socioeconomic factors such as family size, age/sex composition, education, occupational and life-cycle variables have been shown to influence consumption decisions (e.g., West and Price; Buse and Salathe). Life-cycle concepts (Ferber) and the "new household economic theory" (Becker; Lancaster) have extended the applicability of traditional consumer theory and have motivated the incorporation of household socioeconomic characteristics via the household production framework. In particular, household production theory can help explain the allocation of household resources to competing non-market factors and improve the predictive ability of consumption models (Davis).

As an approximation to the underlying behavior suggested by traditional and

household production theory¹, a reduced form household demand specification is hypothesized as:

$$Q_{ij} = f(P_{ij}, PS_{ij}, I_i, FS_i, SE_i) \quad i = 1, \dots, N, \quad (1)$$

where Q_{ij} and P_{ij} represent the quantity and price, respectively of the j^{th} commodity consumed by the i^{th} household; PS_{ij} represents the prices of commodities hypothesized to be substitutes for the j^{th} commodity; I_i is total household income; FS_i represents the age/sex composition of the i^{th} household; and SE_i represents other relevant socioeconomic characteristics of the household. Previous research suggests that relevant socioeconomic variables include occupation of the household head (Price *et al.*), degree of urbanization (Adrian and Daniel), region (Burk), number of meals consumed (Allen and Gadson), season (Be-loian), life-cycle proxies such as age of household head (Ferber; Allen and Gadson), sex and education of the meal planner (Allen and Gadson). In addition, equivalence scale research suggests that the age/sex composition of the family is relevant to household consumption decisions (Price; Buse and Salathe).

Since a number of households did not consume a given disaggregated commodity during the survey period, the demand relation (1) is specified for estimation purposes as a truncated dependent variable or Tobit model:

$$\begin{aligned} Q_{ij} &= X_{ij}\beta_j + u_{ij} && \text{if } X_{ij}\beta_j + u_{ij} > 0 \\ &= 0 && \text{if } X_{ij}\beta_j + u_{ij} \leq 0 \end{aligned} \quad (2)$$

$i = 1, 2, \dots, N,$

where X_{ij} is a $(1 \times K)$ vector of relevant exogenous variable values, β_j is a $(K \times 1)$

parameter vector, and u_{ij} is a random variable assumed to be distributed as normal with mean zero and variance σ^2 .

A number of approaches are available for estimating the parameters in (2). We will distinguish two OLS estimators: (1) truncated OLS, resulting from the use of just the non-limit observations (i.e., those households where $Q_{ij} > 0$); and (2) full sample OLS, resulting from the use of all available observations (i.e., including those households where $Q_{ij} = 0$). Both OLS estimators have been shown to yield biased and inconsistent estimates of β_j and σ^2 for the Tobit model in (2) (Tobin; Greene).

Procedures to correct for truncation bias include: a consistent method proposed by Heckman involving a probit instrumental variable, two-step procedure; Amemiya's maximum likelihood procedure derived from the truncated normal likelihood function and the method of Newton with a consistent initial estimator; an iterative maximum likelihood procedure proposed by Fair; and, a relatively simple, method of moments approximation proposed by Greene. Since the truncated normal likelihood function has a unique maximum (Olsen), Fair's iterative procedure is utilized as it is computationally easier than Amemiya's maximum likelihood procedure.²

The interpretation of the coefficients and the elasticities which result from Tobit analysis differ from those of the traditional normal linear model due to the correction for possible truncation bias. The expected value of consumption from (2) has been shown by Tobin to be:

$$E(Q) = X\beta\Phi(Z) + \sigma\phi(Z), \quad (3)$$

¹ The lack of a computationally tractable simultaneous equation Tobit estimator makes it difficult to test Tobit demand systems for the adding up and symmetry restrictions implied by demand theory. Therefore the estimated equation(s) should be considered approximations to the underlying behavior suggested by demand theory.

² The maximum likelihood procedures of both Amemiya and Fair were evaluated. Since the results are equivalent, there is no loss in using the computationally easier Fair procedure. The covariance matrix for the Fair estimates, however, is calculated by the asymptotic covariance matrix proposed by Amemiya (p. 1006).

where $E(Q)$ is the expected value of consumption, $Z = X\beta/\sigma$, $\phi(Z)$ is the standard normal density function, $\Phi(Z)$ is the standard normal distribution function, and σ is the standard deviation of the normal error term from (2). Amemiya has shown that the expected value of conditional consumption (i.e., conditional upon being non-zero), is simply $X\beta$ plus the expected value of the truncated normal, conditional error term:

$$\begin{aligned} E(Q^*) &= E(Q|Q > 0) \\ &= E(Q|u > -X\beta) \\ &= X\beta + \sigma\phi(Z)/\Phi(Z). \end{aligned} \tag{4}$$

Therefore, expected consumption is directly related to the expected conditional consumption via $\Phi(Z)$, the probability of non-zero consumption as follows:

$$E(Q) = \phi(Z)E(Q^*). \tag{5}$$

Notice that these predicted values, and hence their derivatives with respect to the variables in the design X , can be considerably different than those from the traditional linear model in which $E(Q) = X\beta$.

McDonald and Moffitt suggest the useful decomposition of the marginal effects on (5) due to a change in the k^{th} variable of X :

$$\begin{aligned} \partial E(Q)/\partial X_k &= \Phi(Z)(\partial E(Q^*)/\partial X_k) \\ &+ E(Q^*)(\partial \Phi(Z)/\partial X_k). \end{aligned} \tag{6}$$

This decomposition of the Tobit total predicted response indicates two effects: the change in quantity consumed of the purchasing households weighted by the probability of being a purchasing household (the first component on the right hand side of (6)); and, the change in the probability of being a purchasing household weighted by the expected value of consumption for such a household (the second term on the right hand side of (6)). Thus, expression (6) decomposes the total effect of a change in X_k on expected consumption $E(Q)$ into two additive terms: the conditional effect

(given $Q > 0$) plus the probability or participation effect. The conditional and probability effect components of (6) are useful for interpretation of Tobit estimates.

Note that (6) can be alternatively expressed in elasticity form by multiplying by $X_k/E(Q)$ and using (5). This gives the following elasticity decomposition (see Huang *et al.*)

$$\begin{aligned} \eta_k &= (\partial E(Q^*)/\partial X_k)(X_k/E(Q^*)) \\ &+ (\partial \Phi(Z)/\partial X_k)(X_k/\Phi(Z)) \end{aligned} \tag{7}$$

where $\eta_{jk} = (\partial E(Q_j)/\partial X_k)(X_k/E(Q_j))$. As in (6) above, expression (7) decomposes the total effect elasticity, η_{jk} , into the conditional elasticity associated with non-zero consumption (the first term on the right hand side of (7)) plus the probability effect or participation elasticity as the percentage change in the probability of becoming a consuming household associated with a percentage change in X_j (the second term on the right hand side of (7)). This formulation will be useful in the discussion of price and income elasticities from our model.³

Model Specification and Data

The Western region defined in the NFCS is utilized for this analysis. Total household income from wage and non-wage sources is aggregated over individual household members. After deleting those household observations where income was not reported (11.2 percent),

³ It is important to note that while the Huang, Carley, and Raunikar formulation enhances the interpretation of the elasticity decomposition's components, it suggests the unweighted or marginal effects of a change in X_k on $E(Q^*)$ and $\Phi(Z)$ as the appropriate conditional and probability slopes, respectively. Interest in the *quantity* effects of the decomposition, however, suggests that the components of (6) are more appropriate and so will be used as the slope components of the Tobit elasticities for the purpose of this paper.

some 2,221 households were retained for the analysis. The quantities consumed (pounds per household per week) of fresh potatoes (including home produced potatoes) are used as dependent variables. From quantity and expenditure data, the prices for fresh potatoes (including sweet potatoes and yams) and their close substitutes, commercially canned or frozen and dehydrated potatoes, were obtained. These prices (PFRESH, PFROZEN, and PDEHYDRA for fresh, commercially frozen or canned, and dehydrated potato prices (\$/pound), respectively), were derived by dividing household expenditures by the quantities consumed for each commodity.⁴

The degree of truncation (i.e., the percentage of zero-valued quantity consumed observations) in the selected sample is 29.8, 87.6, and 94.8 percent for the fresh, canned/frozen, and dehydrated categories respectively. Because of this truncation, only 6 households consumed all three commodities during the survey week, resulting in missing price values. In order to obtain price information for each household, the missing price values were estimated using a mean price "grid" procedure. Each missing price of a household was estimated as the average price for households from the same geographic sub-region (Mountain or Pacific) and for the same quarter (spring, summer, fall, or winter). Finer grid procedures were evaluated but subsample cell size was deemed unacceptable.⁵

⁴ Quality dimensions to these "implicit" prices were evaluated under the Prais and Houthakker hypothesis that quality effects would be manifest as price (hence, quality) increases with income. Price equations as a function of region, urbanization, quarter and income were estimated to evaluate this hypothesis. The lack of significance of the estimated income coefficients of these equations supports the conclusion that quality effects in the prices used for this analysis are not significant.

⁵ In estimating missing prices there is a tradeoff between the significance of the variables in explaining the actual prices and the size of the resulting cells

Various specifications with respect to functional form in family size and income were considered. Given the nonlinearity of the Tobit likelihood function, linear-in-parameters functional forms are generally specified to reduce analytical and computational difficulties. Prais and Houthakker suggest that semi-logarithmic income specifications are particularly appropriate for food items as they allow commodities to appear as luxuries at low income levels and as necessities at higher levels of income (Phlips, p. 111). Given this theoretical consideration and its empirical performance, the natural logarithm of income (LNINCOME) was retained in the final model specification.⁶

Demographic changes in family composition and the equivalence scale literature motivates the decomposition of family size into age/sex categories. The non-significance of many of the traditional age/sex categories (particularly for children) and an interest in parsimony for computational reasons, led to the following specification of household composition: the number of children ages 5 or less (CHILD < 5); the number of children ages 6-15 (CHILD > 5); the number of adult males (ADULT_M); and, the number of adult females (ADULT_F). A quadratic term in family size (SQFAMSZ) was retained in the final model to complement the age/sex categories (which are a linear decomposition of family size) and to evaluate potential economies of scale in household consumption. The 21MEALSZ variable has the standard definition, that is,

(that is, the number of observations from which the average grid price is computed). Using geographic sub-region and quarter represented what we felt was the best tradeoff between cell size and explanatory power. The USDA uses a similar average grid price for households consuming a commodity but lacking price information when constructing this data.

⁶ Food stamp recipients comprised less than 5 percent of sample households. Therefore, no variables relating to food stamp expenditures or participation were considered in the model.

the total number of meals eaten from household food supplies in the past week (including refreshment meal equivalents), divided by 21. The 21-meal-equivalent family size captures effects due to family size and the proportion of meals consumed at home.

A number of specifications with respect to SE of (1) were considered. Occupation of the household head, season (quarter), and age of the household head were not found to have significant impacts upon fresh potato consumption. Therefore these variables were dropped from the final model in the interest of parsimony. Discrete zero-one variables for the sex and education of the meal planner (MALEPLAN for male meal planners; ELEMED for meal planners with elementary school education; and COLLED for college educated meal planners), urbanization (SUBURBAN; NONMETRO) and geographic subregion (PACIFIC) were retained in the final model. Estimated parameters on the own, cross-price and income parameters were found to be highly stable with respect to the alternative specifications of SE and FS in (1) which were evaluated.⁷

Results

Parameter Estimates

The parameter estimates from the full sample OLS, the truncated sample OLS, and the ML Tobit results are presented in Table 1 along with the sample means for the full sample (i.e., all households), the limit sample (i.e., non-purchasing households), and the non-limit sample (i.e., pur-

chasing households). Note that prices, income, urbanization and geographic subregion are fairly comparable for all three sample means, while considerably more variation is evident in the family size/composition and sex/education of the meal planner variables.

The maximum likelihood (ML) Tobit estimates closely parallel the full sample OLS results with respect to sign. Notable exceptions are that ADULT_F and the INTERCEPT coefficients are of opposite signs (but statistically insignificant at the $\alpha = .10$ level). The magnitudes of the coefficients are generally different. In 12 out of 18 coefficients, the ML Tobit results are greater in relative value (14 out of 18 in absolute value) than the full sample OLS results. These results reflect the known downward asymptotic bias of OLS in a limited dependent variable model (Greene).

Comparison of the truncated OLS with the ML Tobit results indicates more variation with respect to signs of the coefficients while the general downward bias of the OLS results is again evident. Coefficient signs are opposite in the two models for the INTERCEPT, SQFAMSIZ, MALEPLAN, SUBURBAN and ADULT_F. With respect to downward bias, in 11 of 18 coefficients the ML Tobit results are larger in relative value (13 of 18 in absolute value) than the truncated OLS results.

The ML Tobit results generally appear reasonable with respect to signs and magnitudes. The own and cross price, as well as income coefficients conform to *a priori* expectations.⁸ Commercially canned/frozen and dehydrated potatoes are found to be substitutes for fresh potatoes, which is

⁷ While simultaneity bias, errors in variables, multicollinearity, and other standard econometric problems can be argued to exist (as in most econometric models), we feel that these concerns do not seriously detract from the objectives or results of this paper.

⁸ Comparison of the final model with and without prices yielded a log-likelihood ratio statistic of 20.59. Given a chi-square critical value of 7.81 for $\alpha = .05$ and 3 degrees of freedom, the model with prices is preferred.

TABLE 1. Comparison of the Full Sample and Truncated OLS Results with ML Tobit for Household Fresh Potato Consumption in the Western Region of the 1977-78 NFCS.

Variable	Sample Means			Estimated Coefficients ^a		
	Full Data	Limit Data	Non-Limit Data	Full Sample OLS	Truncated OLS	ML Tobit
INTERCEPT	—	—	—	1.4245 (1.0090)	2.0888 (1.2668)	-.4885 (1.3715)
PFRESH	.145	.146	.144	-6.1755 (.9088)***	-5.9617 (.9784)***	-4.9268 (1.3039)***
PFROZEN	.480	.484	.478	1.8265 (.7505)**	2.7370 (.9297)***	1.8279 (.9954)**
PDEHYDRA	.995	1.008	.990	.4410 (.3675)	.4071 (.4759)	.7679 (.5007)
SQFAMSIZ	10.604	7.592	11.883	-.0038 (.0210)	.0316 (.0265)	-.0766 (.0279)***
21MEALSZ	2.537	1.889	2.813	1.0407 (.1177)***	.9630 (.1554)***	1.5317 (.1576)***
LNINCOME	9.316	9.303	9.321	-.1382 (.0965)	-.1711 (.1213)	-.2546 (.1310)**
MALEPLAN	.111	.210	.069	-.2947 (.2998)	.4696 (.4230)	-1.1880 (.4212)***
ELEMED	.095	.057	.112	.6728 (.2634)**	.2892 (.3164)	1.1360 (.3496)***
COLLED	.405	.506	.362	-.7007 (.1586)***	-.7145 (.2047)***	-.9658 (.2173)***
SUBURBAN	.417	.429	.412	.0170 (.1687)	-.0348 (.2196)	.0588 (.2320)
NONMETRO	.222	.147	.255	.6831 (.2054)***	.5453 (.2549)**	.9172 (.2761)***
PACIFIC	.727	.799	.697	-.3332 (.2022)*	-.1444 (.2544)	-.5795 (.2737)***
CHILD < 5	.285	.215	.315	-.6174 (.2237)***	-.7577 (.2891)***	-.4294 (.2998)
CHILD > 5	.498	.331	.569	.1379 (.2159)	.0894 (.2764)	.4966 (.2880)*
ADULT—M	.967	.808	1.035	.3995 (.2094)*	.1694 (.2799)	1.0104 (.2839)***
ADULT—F	1.090	.955	1.148	-.2826 (.2483)	-.2327 (.3178)	.0533 (.3333)
VARIANCE				11.6024	13.1023	18.7369

^a Asymptotic standard errors in parenthesis.

Asterisks indicate: * significant at the $\alpha = .10$ level; and ** significant at the $\alpha = .05$ level; and *** significant at the $\alpha = .01$ level.

indicated as an inferior good (negative income effect). The family composition effects appear quite reasonable; households

with younger children and female adults are indicated to consume fewer fresh potatoes on average than households with

adult males and older children.⁹ Also, the sign of SQFAMSIZ indicates that, within some range of the data, fresh potato consumption increases at a decreasing rate with respect to family size.¹⁰

Households with a male meal planner form a larger proportion of the limit (non-purchasing) sample. Furthermore, they are almost entirely households without a female household head present, tend to have higher educations and incomes, live in urban or suburban areas, and tend to have smaller family sizes (fewer children and adult females) than the full and non-limit samples. In general, households with male meal planners tend to reflect the characteristics of the non-purchasing sample of households. These notions are supported by the negative sign of the statistically significant coefficient for MALEPLAN.

Predicted Consumption

Given the differences between the alternative parameter estimates, the general downward bias of the OLS results, and the

statistical significance of many of the discrete zero-one variables, one would expect similar differences with respect to predicted values and elasticity measures. These differences should be manifest both between models and within model results evaluated at various subsample means. In general, one would expect the OLS predicted values to be lower than those from ML Tobit due to the downward bias of OLS. With respect to the elasticities the issue is less certain due to the generally lower OLS slope values being offset by larger weights used in the computation of the elasticities (i.e., the lower predicted consumption $E(Q)$). The variation of the elasticity results due to the difference of subsample mean evaluation points (that is, the specific values of X at which the elasticity is evaluated) suggests similar expectations. Based on the ML Tobit results, urban households in the Pacific region with higher incomes, more educated meal planners, fewer older children and fewer adult males would generally be expected to consume fewer fresh potatoes.

Table 2 compares the predicted fresh potato consumption (pounds per household per week) for the ML Tobit, the full sample and truncated OLS results evaluated at various subsample means (evaluation points). Components used in the McDonald and Moffitt decompositions are also included. As expected, there is frequently a large difference between the results at different evaluation points. Some of the evaluation points considered include the Pacific sub-region, urban or non-metro, female only or both male and female headed households, and high school or college educated meal planners. The subsample data headings indicate: Region (P = Pacific); Urbanization (U = urban, N = non-metro); Household Head Status (F = only female head, B = both male and female head); and Education of the Meal Planner (H = high school, C = college). Those subsamples designated MEAN are evaluated at their respective subsample

⁹ The following alternative decomposition of family size was considered: children less than 1, 1-5, 6-10, and 11-15 years of age with ADULT_M and ADULT_F. This specification yielded the ML parameter estimates -4.9575 (PFRESH), 1.8454 (PFROZEN), 0.8153 (PDEHYDRA), -0.0770 (SQFAMSZ), 1.5311 (21MEALSZ), and -0.2608 (LNINCOME) in contrast to the results presented in Table 1. Given the robustness of these coefficients as well as those of the dummy variables (not presented) to alternative decompositions of family size, the 6-10 and 11-15 year old children categories were aggregated in the interests of parsimony in the final model.

¹⁰ Utilizing the sum of the household composition coefficients (CHILD < 5, CHILD > 5, ADULT_M, ADULT_F) as the linear term of the quadratic family size relationship, fresh potato consumption increases with family size up to about 7.4 household members. Given a full sample mean household size of 2.85 members, with a standard deviation of 1.59, the positive linear component of the quadratic family size relationship dominates the negative quadratic component for most of the full sample households.

TABLE 2. Comparison of ML Tobit Predicted Consumption Values with the OLS Results and Components of the McDonald and Moffitt Decompositions.

Sub-Sample ^a	ML Tobit Expected Conditional Consumption: E(Q*)	ML Tobit Expected Consumption: E(Q)	Truncated OLS Expected Value: X _β — Truncated	Full Sample OLS Expected Value: X _β — Full Sample	Probability Non-Zero Consumption: Φ(Z)	Fraction of Mean Total Response Due To Conditional Response
FULLMEAN	4.1822	2.7555	3.6585	2.7914	0.6589	0.4620
PUFCMEAN	3.0146	1.1452	1.6490	0.6130	0.3799	0.3021
PUFHMEAN	3.6332	1.9762	2.9097	1.8618	0.5439	0.3881
PUBCMEAN	3.8366	2.2632	3.0037	2.1213	0.5899	0.4159
PUBHMEAN	4.3933	3.0561	3.9141	3.0268	0.6956	0.4894
PNBCMEAN	4.6214	3.3792	4.2038	3.4198	0.7312	0.5183
PNBHMEAN	5.2881	4.3043	5.0040	4.2546	0.8140	0.5980
PUBCH10	3.6290	1.9704	2.4331	1.6327	0.5430	0.3876
PUBCH01	4.0115	2.5120	3.2802	2.3879	0.6262	0.4394
PUBCL10	3.7692	2.1677	2.6703	1.8243	0.5751	0.4067
PUBCL01	4.2302	2.8239	3.6786	2.7228	0.6676	0.4683

Note: Predicted consumption values are measured in pounds per household per week.

^a P = Pacific region; U = urban, N = non-metro; F = female head only, B = both male and female head; H = high school education, C = college educated; MEAN = evaluated at sample means for other explanatory variables; income level: H = \$20,000, L = 5,000; 10 = younger child, 01 = older child.

means (derived from the full data) for all other independent variables. Thus, for example, PUBCMEAN indicates an average household in the Pacific region, in an urban area, both male and female household heads present, and a college educated meal planner. In addition, those subsamples not designated MEAN, indicate the influence of income level (H = \$20,000, L = \$5,000) and contrast a younger child (. . . 10) versus an older child (. . . 01). Thus, for example, PUBCH10 represents a PUBC household as above with the exceptions that a high income level (\$20,000), one younger and no older children have replaced the sample means for these variables in the evaluation point.

Since the full sample OLS model utilizes the full sample of data, its predicted value should reflect the expected unconditional consumption of all observations. As expected, the full sample OLS predicted values are generally less than those from the ML Tobit E(Q) due to the downward bias of the OLS coefficients. The exceptions are the FULLMEAN and PUFCMEAN evaluation points. Similarly,

since the truncated OLS model utilizes only the non-limit observations (i.e., those conditional on non-zero consumption), its predicted value should reflect expected conditional consumption comparable to the ML Tobit E(Q*) results. The ML Tobit E(Q*) are again larger than the truncated OLS predicted values, as expected.

Note that expected conditional consumption is always larger than the unconditional due to the probability aspects of (5). Furthermore, as expected, urban households with only female household heads, and college educated meal planners are all predicted to consume less fresh potatoes than their counterparts, i.e., non-urban households with both male and female household heads and high school educated meal planners, *ceteris paribus*. The last four subsamples reflect the influence of income level and age of children. As expected, higher income households and those with a young child are predicted to consume fewer fresh potatoes.

It is also interesting to note that the differences between the full sample OLS predicted consumption and the ML Tobit

$E(Q)$ decrease with increased expected consumption. The same relationship is true for the differences between the truncated OLS predicted consumption and the ML Tobit $E(Q^*)$. Thus, for example, the full sample and truncated OLS predicted consumption levels are 54.7 and 53.5 percent of their ML Tobit equivalents ($E(Q)$ and $E(Q^*)$, respectively) for the lowest expected consumption subsample, PUFCEAN. In contrast, for the highest expected consumption subsample, PNBHMEAN, the full sample and truncated OLS predicted consumption levels are 98.8 and 94.7 percent of the ML Tobit $E(Q)$ and $E(Q^*)$, respectively. In general, the full sample OLS predicted consumption levels are closer to the appropriate ML Tobit predicted consumption levels than the truncated OLS results. This is not surprising given that truncated OLS discards the information contained in the limit (non-purchasing) households.

Tobit Decompositions

Table 2 also presents two components used in the McDonald and Moffitt decompositions. The probability of non-zero consumption as evaluated by the cumulative distribution function $\Phi(Z)$, reflects the ordering of the ML Tobit estimates for $E(Q)$ and $E(Q^*)$ (see (5)). Thus, higher expected consumption reflects a higher probability of non-zero consumption. The fraction of mean total response due to conditional response (i.e., due to the response of actual consuming households), follows a similar pattern.¹¹ At the two extreme subsample evaluation points, PUFCEAN households have a predicted probability of non-zero consumption of 38 versus 81 percent for PNBHMEAN households. The percentage of the aver-

age total response due to the response of non-limit households is 30 percent for the PUFCEAN households. In contrast, 60 percent of the average total response for PNBHMEAN households is due to the response above the limit. This type of result indicates that the Tobit model can provide useful information on the factors influencing the probability of consumption during a particular period.

Table 3 compares the Tobit decompositions as elasticity and slope (quantity) responses for selected independent variables. These comparisons are evaluated at the full sample means (FULLMEAN) and the subsample extremes with respect to expected consumption, PUFCEAN and PNBHMEAN. Approximate asymptotic standard errors computed using the general procedures in Silvey, allow inferences concerning the significance of these evaluation points and their differences. As would be expected, the tests of elasticity or slope differences from zero generally follow the significance of the associated ML Tobit parameter estimates. An evaluation of significant differences between the PUFCEAN and PNBHMEAN Tobit decompositions is summarized in Table 4.¹²

Several characteristics of the results in Tables 3 and 4 are worthy of note. First, the subsample evaluation points generally differ from the FULLMEAN and from each other. Table 4 indicates that these differences are statistically significant for all responses except the PDEHYDRA estimates. The significantly different sub-

¹¹ This component, part of the $\partial E(Q^*)/\partial X_k$ term of the conditional quantity response in (6), is derivable as $[1 - Z\phi(Z)/\Phi(Z) - \phi(Z)^2/\Phi(Z)^2]$ (McDonald and Moffitt).

¹² Treating the PUFCEAN and PNBHMEAN Tobit decompositions as asymptotically normal random variables, their difference is thus a normal random variable with variance equal to the sum of their respective variances minus twice their covariance (Mood *et al.*, p.178-79). Thus, the test statistic $(\text{PUFCMEAN} - \text{PNBHMEAN})/(\text{var}(\text{PUFCMEAN}) + \text{var}(\text{PNBHMEAN}) - 2*\text{cov}(\text{PUFCMEAN}, \text{PNBHMEAN}))^{1/2}$ will be asymptotically distributed as $N(0, 1)$ under the null hypothesis that the two subsample results are equal.

sample responses could provide useful information for market segmentation analysis and the targeting of promotion strategies. As well, these results indicate the inferential usefulness of standard errors for the Tobit decomposition components. This additional information is frequently lacking in applied Tobit analysis. Second, the elasticity measures can be misleading where larger predicted *percentage* changes actually refer to significantly smaller expected consumption, and hence, imply smaller predicted absolute *quantity* responses. For example, contrast of the total effect elasticities and their slope components generally reflects an opposite movement with respect to increasing expected consumption. Third, there is a considerably more detailed indication of consumption behavior through the use of Tobit analysis as reflected by the decompositions and their probability components than with OLS in the traditional, normal linear model. Lastly, given the general downward bias of OLS, one would expect OLS derived results to frequently overestimate consumption elasticities and underestimate expected consumption, particularly as the degree of truncation (and, hence the potential bias of OLS) increases.

The variation in elasticity and slope components due to the point of evaluation in Table 3 is most noticeable in the total effect slopes and elasticities, the conditional effect slopes, and the probability effect elasticities. Table 4 indicates that these components are statistically different between the PUFCEAN and PNBHMEAN households for all estimated responses except the PDEHYDRA measures. The point estimates of the statistically different own price (PFRESH) total elasticities of Table 3 indicate that PUFCEAN is predicted to be twice as responsive as PNBHMEAN in percentage terms. In contrast, however, the corresponding total effect PFRESH slopes (which are statistically different from each other as well)

indicate that the implied quantity change in pounds of fresh potatoes due to unit changes in PFRESH are exactly opposite to the relationship of the elasticity results. Thus, for a \$.10 change in the price of fresh potatoes, PNBHMEAN households are predicted to change consumption by 0.4 pounds (a 9.3 percent change) whereas PUFCEAN households are predicted to change fresh potato consumption by 0.2 pounds (a 17.5 percent change).

A second example of the usefulness of the standard errors and slope components to augment the Tobit elasticities concerns the cross price effects of the two substitute goods, frozen and dehydrated potatoes. Note that the total, conditional and probability effect elasticities for PFROZEN and PDEHYDRA are fairly close in magnitude. The slope components of these elasticities, however, again indicate that the quantities implied by approximately equal predicted percentage changes in consumption, are frequently quite different. The predicted substitution response of fresh potato consumption to PFROZEN considerably dominates the PDEHYDRA response in quantity terms. For example, although the magnitude of their total elasticities are nearly identical, a \$.10 change in PFROZEN is predicted to induce a 0.07 and 0.15 pound total change in fresh potato consumption for PUFCEAN and PNBHMEAN households, respectively. In contrast, a \$.10 change in PDEHYDRA is predicted to induce only a 0.03 and 0.06 pound total change for PUFCEAN and PNBCMEAN households, respectively. The standard errors of Table 3, however, indicate that the effects of PDEHYDRA on fresh potato consumption are not statistically different from zero at the $\alpha = 0.10$ level of significance. Hence, comparisons of substitution effects due to PFROZEN and PDEHYDRA estimated here, should be treated cautiously. Table 4 likewise suggests that the estimated substitution effects due to PDEHYDRA do not vary significantly among the subsam-

TABLE 3. Comparison of the Total, Conditional and Probability Tobit Effects as Elasticity and Slope (Quantity) Measures for Household Fresh Potato Consumption in the Western Region of the 1977-78 NFCS.^a

Total Effects:											
Elasticities						Slopes					
Subsample	PFRESH	PFROZEN	PDE-HYDRA	LN-INCOME	21MEALSZ	PFRESH	PFROZEN	PDE-HYDRA	LN-INCOME	21MEALSZ	
FULLMEAN	-.1708*** (.0453)	.2098* (.1143)	.1827 (.1191)	-.5672* (.2918)	.9292*** (.0965)	-3.2460*** (.8584)	1.2043* (.6558)	.5059 (.3299)	-.1678* (.0863)	1.0092*** (.1041)	
PUFCMEAN	-.2468*** (.0652)	.2935* (.1602)	.2667 (.1744)	-.7615* (.3911)	.6682*** (.0720)	-1.8717*** (.5151)	.6944* (.3794)	.2917 (.1901)	-.0967* (.0504)	.5819*** (.0679)	
PNBHMEAN	-.1276*** (.0340)	.1694* (.0921)	.1535 (.1003)	-.4516* (.2330)	.9118*** (.0968)	-4.0102*** (1.0631)	1.4878* (.8122)	.6250 (.4072)	-.2073* (.1065)	1.2468*** (.1311)	
Conditional Effects:											
Elasticities						Slopes					
Subsample	PFRESH	PFROZEN	PDE-HYDRA	LN-INCOME	21MEALSZ	PFRESH	PFROZEN	PDE-HYDRA	LN-INCOME	21MEALSZ	
FULLMEAN	-.0789*** (.0209)	.0969* (.0528)	.0844 (.0550)	-.2620* (.1348)	.4293*** (.0446)	-1.4996*** (.3976)	.5564* (.3032)	.2337 (.1525)	-.0775* (.0399)	.4662*** (.0493)	
PUFCMEAN	-.0745*** (.0198)	.0886* (.0483)	.0806 (.0525)	-.2300* (.1185)	.2018*** (.0210)	-.5654*** (.1627)	.2098* (.1155)	.0881 (.0577)	-.0292* (.0154)	.1758*** (.0243)	
PNBHMEAN	-.0763*** (.0202)	.1013* (.0552)	.0918 (.0598)	-.2700* (.1389)	.5452*** (.0565)	-2.3980*** (.6448)	.8897* (.4888)	.3737 (.2437)	-.1239* (.0639)	.7455*** (.0865)	
Probability Effects:											
Elasticities						Slopes					
Subsample	PFRESH	PFROZEN	PDE-HYDRA	LN-INCOME	21MEALSZ	PFRESH	PFROZEN	PDE-HYDRA	LN-INCOME	21MEALSZ	
FULLMEAN	-.0919*** (.0245)	.1129* (.0615)	.0983 (.0641)	-.3052* (.1571)	.4999*** (.0529)	-1.7464*** (.4621)	.6480* (.3528)	.2722 (.1775)	-.0903* (.0464)	.5430*** (.0558)	
PUFCMEAN	-.1722*** (.0457)	.2048* (.1121)	.1861 (.1220)	-.5315* (.2731)	.4663*** (.0528)	-1.3063*** (.3542)	.4847* (.2642)	.2036 (.1326)	-.0675* (.0350)	.4061*** (.0447)	
PNBHMEAN	-.0513*** (.0141)	.0681* (.0372)	.0617 (.0407)	-.1815* (.0947)	.3666*** (.0460)	-1.6122*** (.4295)	.5982* (.3254)	.2513 (.1642)	-.0833* (.0430)	.5012*** (.0531)	

^a Asymptotic standard errors in parentheses. Significance levels of $\alpha = .10, .05, \text{ and } .01$ are noted by *, **, and ***, respectively.

TABLE 4. Estimated Differences between the PUFCEAN and PNBHMEAN Subsamples with Respect to the Tobit Elasticities and Slopes.^a

	Elasticities					Slopes				
	PFRESH	PFROZEN	PDE-HYDRA	LN-INCOME	21MEALSZ	PFRESH	PFROZEN	PDE-HYDRA	LN-INCOME	21MEALSZ
Total Effects:										
Difference	-.1192*** (.0320)	.1241* (.0687)	.1132 (.0745)	-.3099* (.1599)	.2436*** (.0404)	2.1385*** (.5749)	-.7934* (.4377)	-.3333 (.2188)	.1106* (.0569)	-.6649*** (.0830)
Conditional Effects:										
Difference	.0018** (.0009)	-.0127* (.0070)	-.0112 (.0074)	.0400* (.0206)	-.3434*** (.0357)	1.8326*** (.4986)	-.6799* (.3764)	-.2856 (.1872)	.0947* (.0489)	-.5697*** (.0736)
Probability Effects:										
Difference	-.1209*** (.0326)	.1367* (.0757)	.1244 (.0819)	-.3500* (.1805)	.0997*** (.0371)	.3059*** (.1063)	-.1135* (.0673)	-.0477 (.0337)	.0158* (.0089)	-.0951*** (.0248)

^a Asymptotic standard errors in parentheses. Significance levels of $\alpha = .10, .05, \text{ and } .01$ are noted by *, **, and ***, respectively.

TABLE 5. Comparison of the Full Sample and Truncated OLS Elasticities for Household Fresh Potato Consumption in the Western Region of the 1977-78 NFCS.

Subsample	Full Sample				
	PFRESH	PFROZEN	PDEHYDRA	LNINCOME	21MEALSZ
FULLMEAN	-0.3208	0.3141	0.1572	-0.4615	0.9459
PUFCMEAN	-1.5213	1.4422	0.7532	-2.0337	2.2327
PUFHMEAN	-0.4876	0.4768	0.2532	-0.6444	0.9749
PUBCMEAN	-0.4512	0.4202	0.2212	-0.6332	1.2957
PUBHMEAN	-0.3060	0.2927	0.1534	-0.4348	0.9497
PNBCMEAN	-0.2655	0.2494	0.1350	-0.3877	0.9446
PNBHMEAN	-0.1989	0.2104	0.1096	-0.3048	0.7700
PUBCH10	-0.5863	0.5459	0.2874	-0.8388	1.6835
PUBCH01	-0.4008	0.3733	0.1965	-0.5735	1.1510
PUBCL10	-0.5247	0.4886	0.2572	-0.6456	1.5066
PUBCL01	-0.3402	0.3253	0.1705	-0.4326	1.0557

Subsample	Truncated				
	PFRESH	PFROZEN	PDEHYDRA	LNINCOME	21MEALSZ
FULLMEAN	-0.2363	0.3591	0.1107	-0.4356	0.6678
PUFCMEAN	-0.5459	0.8033	0.2585	-0.9352	0.7679
PUFHMEAN	-0.3012	0.4572	0.1496	-0.5101	0.5772
PUBCMEAN	-0.3076	0.4447	0.1442	-0.5532	0.8467
PUBHMEAN	-0.2285	0.3391	0.1095	-0.4159	0.6795
PNBCMEAN	-0.2085	0.3041	0.1014	-0.3901	0.7111
PNBHMEAN	-0.1632	0.2680	0.0860	-0.3206	0.6058
PUBCH10	-0.3798	0.5490	0.1780	-0.6963	1.0453
PUBCH01	-0.2817	0.4072	0.1320	-0.5165	0.7753
PUBCL10	-0.3461	0.5002	0.1622	-0.5456	0.9524
PUBCL01	-0.2431	0.3609	0.1165	-0.3961	0.7230

ples evaluated. Thus, the standard errors for the Tobit decompositions can provide useful insight concerning inferences and interpretation of Tobit model results.

Comparison of the conditional and probability effects of the Tobit decomposition further indicates the information contained in the ML Tobit estimates for disaggregated commodity analysis. The conditional effect elasticities and slopes both increase in predicted responsiveness as expected consumption increases. In contrast, the probability effects parallel the magnitude relationships of the total effects. Thus, the elasticities decrease and the slopes increase in responsiveness as expected consumption increases. Note that the conditional effects for the highest expected consumption subsample (PNBHMEAN) dominate the corresponding

probability effects of the decomposition. This relationship is reversed for the lowest expected consumption subsample (PUFCMEAN). This result is also indicated by the fraction of mean total response due to response above the limit (see Table 2), suggesting that both percentage and quantity responses due to changes in the probability of non-zero consumption are more dominant than the conditional responses among the lower expected consumption households.

OLS Elasticities

A comparison of the ML Tobit results of Table 3 with the full sample and truncated OLS elasticities of Table 5, summarizes their differences. Aside from the elasticity magnitude differences, where,

as expected, the OLS results generally overstate responsiveness relative to the ML Tobit results, one very important difference between results is that the ML Tobit quantity effects (slopes) change with the evaluation point while the OLS slopes are constant. Given the range of magnitudes of the ML Tobit quantity effects with respect to different household characteristics, ML Tobit results appear more informative relative to OLS for market segmentation analysis.¹³

A second major difference between the results is the Tobit decomposition and its probability components. Following the argument that truncated OLS reflects conditional consumption behavior, the predicted elasticities from the truncated OLS are considerably larger in absolute value than either the estimated conditional or total ML Tobit elasticities, a result possibly due to their failure to adjust for the probability response. It should be noted that the apparent effects of truncation bias exhibited here reflect a sample that is not severely truncated, with only 30 percent non-purchasing households. In general, one would expect the discrepancies between OLS and ML Tobit to increase with greater sample truncation.

Conclusions

This paper has applied the McDonald and Moffitt decomposition to ML Tobit results, compared predicted consumption and elasticities with full sample and truncated OLS results, and indicated the potential usefulness of Tobit analysis in dis-

aggregated, cross-sectional analysis of regional fresh potato consumption. The inclusion of price variables in the disaggregated specification was found to yield reasonable parameter and elasticity estimates, thus providing useful evidence on the extent of substitution in fresh potato consumption among socioeconomic sub-groups in the Western region of the NFCS. Subsample evaluation revealed that urbanization, household head status, and education of the meal planner have a measurable influence on price and income elasticities. The estimation of the standard errors for the Tobit decomposition indicated that the differences between the extremes of the subsample evaluation points (PUFCMEAN and PNBHMEAN) were statistically significant for all estimated response measures except the PDEHYDRA measures.

Our results suggest that decomposition of Tobit elasticity and slope estimates into conditional and probability effect components can be useful for disaggregated market development purposes. Socioeconomic variables in the analysis suggested that demand responsiveness to own price and income may vary considerably among market sub-groups. Thus, the approach proposed here could be useful in the identification of target groups for market promotion strategies. In addition, given recent demographic trends such as regional population shifts, increasing numbers of single parent households, urban to rural migration, etc., the procedure utilized in this paper may contribute to a better understanding of regional food consumption patterns and their future projection.

Comparison with full sample and truncated OLS results suggests that moderate truncation may induce notable bias in OLS results. ML Tobit estimates appear preferable for two reasons. First, they have desirable asymptotic properties; in particular they are consistent and asymptotic efficient. Second, through the McDonald and Moffitt decomposition, they provide use-

¹³ Market segmentation seeks to identify market sub-groups with somewhat homogeneous characteristics and behavior that is distinct from other sub-groups. Based upon their distinctive responses and characteristics (if any), market promotion strategies can be devised to target specific sub-groups of interest. In this case, sub-group response measures would be preferred to aggregate response measures for predicting potentially distinct sub-group behavior.

ful information in the analysis of disaggregated consumption behavior. Based on our results, evaluation of price effects along with those of socioeconomic variables in other disaggregated, cross-sectional commodity models with differing degrees of truncation and aggregation, appears to be an interesting avenue for further research.

References

- Adrian, J. and R. Daniel. "Impact of Socioeconomic Factors on Consumption of Selected Food Nutrients in the United States." *American Journal of Agricultural Economics*, 58(1976): 1 31-8.
- Allen, J. E. and K. E. Gadson. "The Influence of Selected Socioeconomic Variables on At-Home Nutrient Intake." Selected paper presented at the Southern Agricultural Economics Association meeting in Orlando, Florida, February 7-10, 1982.
- Amemiya, T. "Regression Analysis When the Dependent Variable is Truncated Normal." *Econometrica*, 41(1973): 6 997-1016.
- Becker, G. S. *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press, 1976.
- Beloian, A. M. "Seasonal Variations in U.S. Diets." *Family Economics Review*, (March 1971): 1 25-9.
- Burk, M. C. *Influences of Economic and Social Factors on U.S. Food Consumption*. Minneapolis: Burgess Publishing Co., 1961.
- Buse, R. C. and L. E. Salathe. "Adult Equivalent Scales: An Alternative Approach." *American Journal of Agricultural Economics*, 60(1978): 3 460-8.
- Capps, O. and J. M. Love. "Determinants of Household Expenditure on Fresh Vegetables." *Southern Journal of Agricultural Economics*, 15(1983): 1 127-32.
- Davis, Carlton G. *Linkages Between Socioeconomic Characteristics, Food Expenditure Patterns and Nutritional Status of Low-Income Households: A Critical Review*. Florida Agricultural Experiment Station Journal Series No. 3886, August 1982.
- Fair, R. C. "Note on the Computation of the Tobit Estimator." *Econometrica*, 45(1977): 7 1723-7.
- Ferber, R. "Consumer Economics, A Survey." *Journal of Economic Literature*, 11(1973): 4 1303-42.
- Greene, W. H. "On the Asymptotic Bias of the Ordinary Least Squares Estimator of the Tobit Model." *Econometrica*, 49(1981): 2 505-13.
- Heckman, J. J. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1979): 1 153-61.
- Huang, C. L., D. H. Carley, and R. Raunikar. "Use of Elasticities: Implications of Market Segmentation for Fluid Milk." Contributed paper presented at the American Agricultural Economics Association meetings in Logan, Utah, August 1-4, 1982.
- Lancaster, K. L. "A New Approach to Consumer Theory." *Journal of Political Economy*, 74(1966): 1 132-57.
- McDonald, J. F. and R. A. Moffitt. "The Uses of Tobit Analysis." *Review of Economics and Statistics*, 62(1980): 2 318-21.
- Mood, A. M., F. A. Graybill, and D. C. Boes. *Introduction to the Theory of Statistics*. Third Edition. McGraw-Hill, Inc., New York, 1974.
- Olsen, R. "Note on the Uniqueness of the Maximum Likelihood Estimator for the Tobit Model." *Econometrica*, 46(1978): 5 1211-5.
- Phlips, L. *Applied Consumption Analysis*. North Holland Publishing Company, Amsterdam, 1974.
- Prais, S. J. and H. S. Houthakker. *The Analysis of Family Budgets*. Cambridge University Press, Cambridge, 1955.
- Price, D. N. "Unit Equivalent Scales for Specific Food Commodities." *American Journal of Agricultural Economics*, 52(1970): 2 224-33.
- Price, D. W., D. Z. Price, and D. A. West. *Traditional and Nontraditional Determinants of Household Expenditures on Selected Fruits and Vegetables*. Washington State University, College of Agriculture Research Center, Scientific Paper No. 5420, 1980.
- Silvey, S. D. *Statistical Inference*. Chapman and Hall, London, 1975.
- Tobin, J. "Estimation of Relationships for Limited Dependent Variables." *Econometrica*, 26(1958): 1 24-36.

United States Department of Agriculture. Science and Education Administration. *Nationwide Food Consumption Survey 1977-78*. Household Basic Sample Public Use Tapes.

West, D. A. and D. W. Price. "The Effects of Income, Assets, Food Programs and Household Size on Food Consumption." *American Journal of Agricultural Economics*, 58(1976): 4 725-30.