Modeling Consumption with Limited Dependent Variables: Applications to Pork and Cheese

Dietary Assessment Research Series Report 3

Steven T. Yen and Helen H. Jensen

Staff Report 95-SR 76
September 1995

Modeling Consumption with Limited Dependent Variables: Applications to Pork and Cheese

Dietary Assessment Research Series Report 3

Steven T. Yen and Helen H. Jensen

Staff Report 95-SR 76
September 1995

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011

Steven T. Yen is a lecturer in the Department of Economics, University of Illinois at Urbana-Champaign, and a former CARD visiting assistant professor; and Helen H. Jensen is an associate professor of economics, Iowa State University, and head of the Food and Nutrition Policy Division, CARD.

Partial funding support was provided under Cooperative Research Agreement #58-3198-2-006 between the Agricultural Research Service, USDA, and Iowa State University. Mary Hama provided helpful comments on an earlier draft.

CONTENTS

TABLES

ABSTRACT

The double-hurdle and infrequency-of-purchase models are applied to pork and cheese consumption using the 1987-88 Nationwide Food Consumption Survey data. The models are generalized with the inverse hyperbolic sine transformation in the dependent variable, and this transformation results in a more flexible parameterization and error distribution than the untransformed models. Nonnested LR tests suggest that the IHS double-hurdle model provides better characterization of the data-generating process in household pork consumption than the IHS infrequency-of-purchase model but the elasticities derived from these models are similar. For household cheese consumption, the two models fit the data equally well. The IHS double-hurdle model and the IHS infrequency-of-purchase model generate very similar elasticities. Own-price effects on the probability and level of consumption are negative and significant for both pork and cheese, while income and cross-price effects are not significant. Household age composition, education, gender of meal planner, and race are among the demographic variables that affect consumption.

Key words: double-hurdle model, heteroskedasticity, infrequency of purchase, inverse hyperbolic sine, pork consumption, Nationwide Food Consumption Survey

MODELING CONSUMPTION WITH LIMITED DEPENDENT VARIABLES: APPLICATIONS TO PORK AND CHEESE

Modeling consumption relationships with microdata is complicated by observed zero consumption in the sample. The Tobit model (Tobin) has often been used in empirical applications and has commonly been estimated with homoskedastic and truncated normal errors. However, in limited dependent variable models, maximum-likelihood (ML) estimation produces biased and inconsistent parameter estimates when the errors are heteroskedastic (Arabmazar and Schmidt 1981) or nonnormally distributed (Arabmazar and Schmidt 1982; Robinson 1982). Nonnormal and heteroskedastic errors are features of data in food demand analysis (Yen 1993).

Another restrictive feature that may render the Tobit model unpalatable for empirical analysis is the type of parameterization used (Cragg 1971; Lee and Maddala 1985; Lin and Schmidt 1984). In particular, the Tobit model implies that the probability and level of consumption are determined by the same sets of parameters and variables. Such parameterization has been rejected in food demand analysis (Haines et al. 1988; Yen 1993, 1994).

A more flexible parameterization than the Tobit model, suggested by Cragg (1971) and Atkinson et al. (1984), is the double-hurdle model. Recent applications include Haines et al. (1988), Blaylock and Blisard (1992, 1993a, 1993b), and Blisard and Blaylock (1993). The double-hurdle model features two stochastic processes that determine the probability and conditional level of consumption, and accounts for zero observations resulting from true nonconsumption determined by economic and market determinants (i.e., corner solutions) as well as other factors such as "conscientious abstention" (Pudney 1988). However, in cross-section data, typically collected in short-duration surveys, other factors such as infrequency of purchases cannot be ruled out as the causes of zero observations. Deaton and Irish (1984) and Blundell and Meghir (1987) have proposed an alternative framework for modeling demand with zero observations that result from infrequent purchases. In either case, most empirical studies of the double-hurdle and infrequency-of-purchase models have been estimated with homoskedastic and truncated normal errors.

The specification of models to allow for a nonnormal error structure have been proposed for the Tobit model. Examples include the exponential Tobit (Maddala 1983, pp. 187-90), the Gamma Tobit (Atkinson et al. 1990), and the lognormal Tobit (Amemiya and Boskin 1981). Like the standard

Tobit model, these models are based on specific error distributions and, therefore, are subject to specification errors. Poirier (1978) proposed the Box-Cox transformation in a truncated regression, which allows skewness in the error distribution. More recently, Lankford and Wyckoff (1991) incorporated the Box-Cox transformation into the Tobit model. These models feature nonnormal errors, but the Tobit parameterization and homoskedasticity of errors are maintained.

Models that generalize both the parameterization and distributional assumptions of the Tobit model are less common. They include the Box-Cox double-hurdle model (Jones and Yen 1994; Yen 1993, 1994; Yen and Jones 1994), in which normality, homoskedasticity, and parameterization of the Tobit model are all relaxed.

As is well known, the Box-Cox transformation cannot be used when the random variable can take on nonpositive values. Yet, this is the case when a latent consumption variable is used to explain the observed zero consumption. The inverse hyperbolic sine (IHS) transformation, due to Johnson (1949), accommodates zero, negative, and positive values for the random variable and is known to better handle extreme values. Reynolds and Shonkwiler (1991) have applied the IHS transformation to the Tobit model. In this paper, the IHS transformation is incorporated in the heteroskedastic double-hurdle and infrequency-of-purchase models. The resulting specifications feature flexible parameterization and accommodate nonnormal and heteroskedastic errors. The specification of the heteroskedastic double-hurdle and infrequency-of-purchase models is proposed as a generalized model of household consumption decisions when zeros are present in the observed consumption. In this paper, the IHS-transformed double-hurdle and infrequency-of-purchase models are applied to the examples of pork and cheese.

The IHS Double-Hurdle and Infrequency-of-Purchase Models

The double-hurdle model, developed by Cragg and Atkinson et al. (1984), is characterized by a latent participation variable D_t^* specified as a linear function of the first-hurdle regressors ($\mathbf{z}_t$)

$$D_t^* = z_t \boldsymbol{\alpha} + \upsilon_t, \tag{1}$$

and a latent consumption variable y_t^* specified as a linear function of the second-hurdle regressors ($\mathbf{x}_t$)

$$y_t^* = \boldsymbol{x}_t \boldsymbol{\beta} + \varepsilon_t, \tag{2}$$

where α and β are conformable parameter vectors, and ε_t and υ_t are independent random errors such that $\varepsilon_t \sim N(0,1)$ and $\upsilon_t \sim N(0,\sigma_t)$. The observed dependent variable y_t is such that

$$y_t = y_t^* \quad \text{if} \quad D_t^* > 0 \quad \text{and} \quad y_t^* > 0 \; , \qquad \qquad (3)$$
$$= 0 \quad \text{otherwise}.$$

Thus, there are two separate stochastic processes that determine the value of y_t. In particular, for a positive consumption to occur, two hurdles have to be overcome: to participate in the market (i.e., to be a potential consumer), and to actually consume.

As noted earlier, one problem with the double-hurdle model is that ML estimates are biased and inconsistent when the errors ε_t are nonnormally distributed (Arabmazar and Schmidt 1982; Robinson 1982) or heteroskedastic (Arabmazar and Schmidt 1981). In traditional regression models, transformations of the random variables, such as the Box-Cox transformation (Box and Cox 1964) and the IHS transformation (Burbidge et al. 1988; Johnson 1949), have been used when normality or homoskedasticity of the errors is in doubt. The Box-Cox transformation has been used in limited dependent variable models by Jones and Yen (1994), Lankford and Wyckoff (1991), Poirier (1978), and Yen (1993, 1994). However, as noted earlier, the Box-Cox transformation cannot be performed on random variables that can take on zero or negative values. In this paper, we use an alternative transformation.[1]

Consider the inverse hyperbolic sine (IHS) transformation on random variable v,

$$T(v) = \log[\theta v + (\theta^2 v^2 + 1)^{1/2}]/\theta \qquad \qquad (4)$$
$$= \sinh^{-1}(\theta v)/\theta,$$

defined over all values of θ. Because the transformed variable is symmetric about 0 in θ, we can consider only $\theta \geq 0$. The transformation is linear when θ approaches zero and behaves logarithmically for large values of θ (Burbidge et al. 1988). In addition, such transformation can be performed on random variables that can take on any values.

Applying the IHS transformation to the double-hurdle model, the latent consumption equation becomes

$$T(y_t^*) = x_t\beta + \varepsilon_t . \tag{5}$$

With the IHS transformation on the latent variable y_t^*, the error term υ_t has a better chance of satisfying the normality and homoskedasticity of errors. Note that the IHS transformation is carried out on the latent variable y_t^* that, by definition, can be positive, negative, or zero. The censoring rule (3) can be modified as

$$\begin{aligned} T(y_t) &= T(y_t^*) \quad \text{if} \quad D_t^* > 0 \quad \text{and} \quad T(y_t^*) > 0 \ , \\ &= y_t = 0 \quad \text{otherwise}. \end{aligned} \tag{6}$$

With the normality assumption of ε_t, the conditional density of y_t is

$$\mathrm{f}(y_t | y_t > 0) = \left[\Phi\left(\frac{x_t\beta}{\sigma_t}\right)\right]^{-1} \frac{1}{\sigma_t} \phi\left[\frac{T(y_t) - x_t\beta}{\sigma_t}\right] \frac{1}{\left(1 + \theta^2 y_t^2\right)^{1/2}}, \tag{7}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the univariate standard normal distribution and density functions. The last term on the right hand side of (7) is the Jacobian of the transformation from $T(y_t)$ to y_t. Equation (7) indicates that the IHS transformation allows for skewness in the pre-truncated conditional density of the observed random variable y_t.

Define a dichotomous variable I_t such that $I_t = 1$ if $y_t > 0$, and 0 otherwise. Then, for an independent sample of size N, the sample likelihood function for the IHS double-hurdle model is

$$\mathrm{L} = \prod_{t=1}^{N} \left\{ 1 - \Phi(z_t\alpha)\, \Phi\left(\frac{x_t\beta}{\sigma_t}\right)\right\}^{1-I_t} \left\{ \Phi(z_t\alpha) \frac{1}{\sigma_t} \phi\left[\frac{T(y_t) - x_t\beta}{\sigma_t}\right] \frac{1}{\left(1 + \theta^2 y_t^2\right)^{1/2}} \right\}^{I_t} \tag{8}$$

The likelihood function (8) reduces to the IHS Tobit model when $\Phi(z_t\alpha) = 1$. In addition, imposing the restriction $\theta = 0$ on the IHS double-hurdle and IHS Tobit models leads to the standard (untransformed) double-hurdle and Tobit models.

For the double-hurdle models considered here, zero observations are attributed to conscientious abstention and/or economic determinants of nonconsumption (Pudney 1988). This rules out other sources of zeros, such as infrequency of purchases (which are also plausible reasons for the zero observations).

The infrequency-of-purchase model is characterized by a latent consumption equation defined in (2), and a latent purchase equation specified as a linear function of exogenous variables $\mathbf{w}_t$ and conformable parameter vector $\boldsymbol{\delta}$:

$$S_t^* = w_t \boldsymbol{\delta} + u_t, \tag{9}$$

where S_t^* is a latent variable for purchase. Assuming the random error u_t is distributed as $N(0,1)$, the probability of purchase can be expressed as $\Phi(\mathbf{w}_t\delta)$. The observed dependent variable y_t is such that (Deaton and Irish 1984, eq. 10)

$$\begin{aligned} \Phi(w_t\boldsymbol{\delta})y_t &= y_t^* \quad \text{if } y_t^* > 0 \quad \text{and} \quad S_t^* > 0 \ , \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{10}$$

Like the double-hurdle model, the infrequency-of-purchase model also contains two separate stochastic processes: a probit mechanism for the purchase decision, and a Tobit mechanism that generates zeroes for potential nonconsumers. Thus, a zero observation occurs when a household does not purchase due to shopping and inventory practices, for example, or does not consume due to economic determinants (i.e., a corner solution). The likelihood function can be written using (2), (9), and (10). The resulting model generalizes the constant P-Tobit model of Deaton and Irish (1984, eq. 11) by replacing the probability of purchase with the cumulative standard normal probability. It also amounts to the "Tobit with infrequency of purchase" model suggested by Blundell and Meghir (1987).

To incorporate the IHS transformation into the infrequency-of-purchase model, the censoring rule (10) can be modified as

$$\begin{aligned} T[\Phi(w_t\boldsymbol{\delta})y_t] &= T(y_t^*) \quad \text{if } T(y_t^*) > 0 \quad \text{and} \quad S_t^* > 0 \ , \\ &= y_t = 0 \qquad \text{otherwise.} \end{aligned} \tag{11}$$

Based on (5) and (11), the conditional density of y_t is

$$f(y_t \mid y_t > 0) = \left[\Phi\left(\frac{x_t\beta}{\sigma_t}\right)\right]^{-1} \frac{1}{\sigma_t}\,\phi\left\{\frac{T[\Phi(w_t\delta)y_t] - x_t\beta}{\sigma_t}\right\} \times \frac{\Phi(w_t\delta)}{\left\{1 + [\Phi(w_t\delta)]^2\theta^2 y_t^2\right\}^{1/2}}, \tag{12}$$

where the final expression in (12) is the Jacobian of transformation from $T[\Phi(w_t\delta)y_t]$ to y_t. Using (9), (11), and (12), the sample likelihood function for the IHS infrequency-of-purchase model is

$$L = \prod_{t=1}^{N}\left\{1 - \Phi(w_t\delta)\,\Phi\left(\frac{x_t\beta}{\sigma_t}\right)\right\}^{1-I_t} \times \left\{[\Phi(w_t\delta)]^2 \frac{1}{\sigma_t}\,\phi\left[\frac{T[\Phi(w_t\delta)y_t] - x_t\beta}{\sigma_t}\right] \frac{1}{\left(1 + [\Phi(w_t\delta)]^2\theta^2 y_t^2\right)^{1/2}}\right\}^{I_t}. \tag{13}$$

The likelihood function (13) nests the infrequency-of-purchase model (with Tobit) when $\theta = 0$ (Blundell and Meghir 1987, Table 1), and reduces to the IHS Tobit model when $\Phi(w_t\delta) = 1$.

Deaton and Irish (1984) have suggested that the double-hurdle model and the infrequency-of-purchase model are observationally equivalent when the participation probability $\Phi(z_t\alpha)$ and purchase probability $\Phi(w_t\delta)$ are constant.[2] However, when the probabilities are generalized to vary across observations, the two models are not so intimately related. Because both models contain the IHS Tobit and Tobit models as special cases and provide quite plausible accounting for the zero observations for the commodity and data considered in this study, the results from the two models yield useful comparisons.

A final issue is that of heteroskedasticity of the error terms. To incorporate heteroskedasticity into the IHS double-hurdle and IHS infrequency-of-purchase models, the standard deviation can be specified as

$$\sigma_t = \exp(h_t\gamma), \tag{14}$$

where $\mathbf{h}_t$ is a vector of exogenous variables (can be a subset of $\mathbf{x}_t$, $\mathbf{w}_t$, or $\mathbf{z}_t$) and γ is a conformable parameter vector.

Examining the Effects of Variables

In limited dependent variable models, it is useful to examine the effects of variables on the probability, conditional level, and unconditional level of consumption, as suggested by McDonald and Moffitt (1980) for the Tobit model. Such decomposition of effects is especially important for the models considered here, because the transformation of the dependent variable, the parameterization, and the heteroskedastic error specification all complicate evaluating the effects of the independent variables. The probabilities of consumption (a positive observation) are

$$P(y_t > 0) = \Phi(z_t\boldsymbol{\alpha})\,\Phi\left(\frac{\boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right) \tag{15}$$

for the IHS double-hurdle model, and

$$P(y_t > 0) = \Phi(\boldsymbol{w}_t\boldsymbol{\delta})\,\Phi\left(\frac{\boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right) \tag{16}$$

for the IHS infrequency-of-purchase model. The conditional means of y_t are

$$E(y_t \mid y_t > 0) = \left[\Phi\left(\frac{\boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right)\right]^{-1} \int_0^{\infty} y_t \frac{1}{\sigma_t}\,\phi\left[\frac{T(y_t) - \boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right] \frac{1}{\left(1 + \theta^2 y_t^2\right)^{1/2}}\,dy_t \tag{17}$$

for the IHS double-hurdle model, and

$$\begin{aligned} E(y_t \mid y_t > 0) &= \Phi(\boldsymbol{w}_t\boldsymbol{\delta})\left[\Phi\left(\frac{\boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right)\right]^{-1} \\ &\times \int_0^{\infty} y_t \frac{1}{\sigma_t}\,\phi\left\{\frac{T[\Phi(\boldsymbol{w}_t\boldsymbol{\delta})y_t] - \boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right\} \frac{1}{\left\{1 + [\Phi(\boldsymbol{w}_t\boldsymbol{\delta})]^2\,\theta^2 y_t^2\right\}^{1/2}}\,dy_t . \end{aligned} \tag{18}$$

for the IHS infrequency-of-purchase model. Then, using these probabilities and conditional means for the respective models, the unconditional mean of y_t is

$$E(y_t) = E(y_t \mid y_t > 0)\,P(y_t > 0). \tag{19}$$

The elasticities of the probability of consumption can be derived by differentiating (15) for the IHS double-hurdle model and (16) for the IHS infrequency-of-purchase models. Likewise, the elasticities of the conditional level can be derived by differentiating (17) and (18). Then, the elasticities of the unconditional level follow from the adding-up property (19).

Note that for the standard (untransformed) models ($\theta = 0$), the Jacobians of the transformation on the right hand side of (17) and (18) vanish and the integrations lead to closed forms (Amemiya 1985; Maddala 1983). For the unrestricted case, a closed form does not exist and the conditional means must be evaluated by numerical integrations. The analytic derivatives of the conditional means with respect to exogenous variables unfortunately also involve integrations. Experience suggests that evaluating the conditional means by Gausian quadratures and then numerically differentiating the conditional means works the best. Using the parameter estimates, these elasticities can be computed at the sample means of variables.

For the binary explanatory variables, evaluating the effects of these variables by elasticities is not strictly correct. The effects of each dummy variable can be calculated from the differences in probability, conditional level, and unconditional level as the value of each variable changes from zero to one.

Data and Samples

This study investigates U.S. household consumption of pork and cheese. The data came from the 1987-88 U.S. Nationwide Food Consumption Survey (NFCS), which was conducted by the Human Nutrition Information Service (HNIS, now Agricultural Research Service) of the U.S. Department of Agriculture from April 1987 to August 1988. The survey collected information on household characteristics as well as the quantity and cost of each food item used at home during the seven-day period prior to the interview.

For each product, the quantity (in pounds) and cost (in dollars) of weekly household consumption were recorded. Price was derived as the unit value obtained by dividing the cost by the quantity available for consumption.[3] For households reporting zero expenditure on a commodity, the price was derived as follows: First, households were grouped by nine geographical divisions (New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific) and three urbanization codes (central cities, metropolitan areas, and nonmetropolitan areas). These divisions provide the greatest degree of geographic specificity available in the data set. Group averages of the price were computed for the consuming households. Then, each group average was used as the price for nonconsuming households belonging to that group. To accommodate cross-price effects, prices for "other products" were derived in the same manner as for unit values and imputed prices.

The original sample contained 4,495 households, but as suggested by the HNIS, only data for 4,273 housekeeping households are suitable for analysis (U.S. Department of Agriculture 1992). Households with missing information on important variables and outliers with prices greater than five standard deviations were excluded. This resulted in 4,198 households for the pork sample and 4,217 households for the cheese sample.

For each application (of pork and cheese), quantity (in pounds) used in the household from purchases, home production, and inventory is used as the dependent variable. Economic theory suggests prices and income as the major determinants of demand. In addition, demographics, household composition, and tastes are expected to play a role in consumption of food products. Thus, the explanatory variables include: (1) own- and cross-prices described above; (2) income as weekly household income in hundreds of dollars; (3) household composition in the five age categories of 9 or under, 10-19, 20-44, 45-64, and 65 or over; (4) education of household head, recoded to 1 for primary school or under, up to 4 for college or higher; and (5) dummy variables indicating regions, ethnic groups (Hispanic or otherwise), gender of person in charge of meal planning, food stamp recipient status, urbanization, home ownership, and race (see Tables 1 and 5).

Both samples are censored, with 1,255 households (29.9 percent) reporting nonconsumption of pork, and 763 households (18.09 percent) reporting nonconsumption of cheese, during the sampling period. The mean weekly consumption of pork is 2.89 pounds for the consuming households and 2.03 pounds for the full sample. For cheese, the mean weekly consumption is 1.29 pounds for the consuming households and 1.05 pounds for the full sample. The sample statistics are reported in Tables 1 and 5.

Results

The following sections discuss the empirical findings, which include parameter estimates and the effects of explanatory variables.

Pork

The IHS double-hurdle and IHS infrequency-of-purchase models were estimated by maximizing the logarithms of the likelihood functions (8) and (13). For each model, numerical optimization was accomplished with the quadratic hill-climbing algorithm (Goldfeld et al. 1966), with the log-likelihood, analytic gradient (see Appendixes A and B), and covariance matrix of the gradient (Berndt et al. 1974) programmed in FORTRAN. The negative Hessian matrix was derived by numerically differentiating the analytic gradient and was inverted to give the variance-covariance matrix of parameters.[4]

Estimation of the IHS double-hurdle model was extremely difficult and it was concluded that some exclusion conditions on the first- and second-hurdle regressors were necessary to identify the two stochastic processes. The choice of regressors proceeded first by a probit analysis based on the purchase-nonpurchase dichotomy, using the whole list of economic and demographic variables discussed above as the regressors. Insignificant variables were gradually excluded from the probit equation, and the remaining significant regressors were used as the first-hurdle regressors in the IHS double-hurdle model.[5] The whole list of variables was included in the consumption equation.

The choice of regressors in the heteroskedasticity equation is yet another specification issue. In numerous preliminary runs of the IHS double-hurdle model, all continuous variables were experimentally included, separately and in various combinations, in the heteroskedasticity equation. Most variables were found insignificant.[6] The search concluded with five age composition variables (plus a constant) in the heteroskedasticity equation with the variables "aged 10-19" and "aged 20-44" being significant at the 0.10 level. Thus, although the IHS transformation is known to be a variance-stabilizing device, the results suggest that some residual heteroskedasticity remained and needed to be adjusted.

For comparisons, the same configuration of variables was used in estimating the IHS infrequency-of-purchase model. One of the age composition variables (10-19) was also significant in the heteroskedasticity equation so the homoskedastic specification would be inappropriate. The results suggest a lower log-likelihood than the IHS double-hurdle model. Because the two models are nonnested, some nonnested model specification test was needed to compare the two models. Following Vuong's (1989) nonnested likelihood-ratio (LR) test procedure (p. 318, eq. 5.6), the standard normal statistic was computed as $Z = 4.14$ (P-Value < 0.001), which suggests that the IHS double-hurdle model is preferable to the IHS infrequency-of-purchase model in that the former is a better characterization of the data-generating process than the latter.

The ML estimates of both models are presented in Table 2. The IHS parameter (θ) was significant (i.e., significantly different from zero) at the 0.01 level for both models; this finding rejects their untransformed (truncated normal) counterparts. All age composition variables, regional dummies for the Midwest and South, and racial dummy white were significant in both the consumption and participation/purchase equations in both models. Income and the price of other meat are not significant in the consumption equation, and the own-price effect is negative and significant.

Because the models involve different parameterization, and because the nonlinear (IHS) transformation and heteroskedasticity specification further complicate the effect of variables, the comparisons of results between the two models can better be accomplished according to the elasticities.

Based on the ML estimates, the elasticities of probability, conditional level, and unconditional level for both models were computed at the sample means of all variables. In addition, the standard errors for these elasticities were computed by mathematical approximations (Fuller 1987). The results are presented in Table 3.

Despite the LR test result, which suggests selection of the IHS double-hurdle model over the IHS infrequency-of-purchase model, the elasticities derived from these models are extremely close. In particular, in comparing elasticities between the two models, very few pairs of corresponding elasticities differ by more than one standard deviation (of either elasticity), and none differ by more than one-and-a-half standard deviations. In view of the recent debates over the suitability of the double-hurdle model and the infrequency-of-purchase model in modeling demand with zero observations (see, for example, Blisard and Blaylock 1993), the implication of the current findings is that, while one model may be chosen over the other on statistical grounds, both models offer plausible behavioral explanations for the zero observations and perform equally well for practical purposes.

According to the elasticities derived from both models, all five age composition variables have significant and positive effects on both the probability and conditional level of consumption. Moreover, the elasticities with respect to "aged 20-44" are much greater than the corresponding elasticities with respect to other age composition variables. These results indicate that households with more members between 20 and 44 are more likely to consume pork and, conditional on consumption, also consume more than others. Relative to other households, households in the

Midwest and South and those in rural areas are more likely to consume pork and, conditional on consumption, also consume more. Households with a female meal planner are more likely to consume pork than others. Relative to other groups, those who are more educated and white are less likely to consume pork and also consume less. This finding may be because these households are more health conscious and, therefore, are more reluctant to consume red meat.[7]

The own-price effects on both the probability and level of consumption are negative and significant, while income and cross-price effects are not significant. Because none of the variables has a conflicting effect on probability and conditional level of consumption, and because of the adding-up property of the elasticities, the elasticities of the unconditional level in general are consistent with the elasticities of the conditional level. The own-price elasticity of pork is about -0.56 (-0.54 according to the IHS infrequency-of-purchase model), which suggests that the demand for pork is price-inelastic. The implication of the relatively weak cross-price effects is that price effects of other meats, though present, do not lead to significant changes in pork consumption; price increases or decreases for pork, on the other hand, can induce changes in consumption.

The average effects of dummy variables are reported in Table 4. These average effects provide a way to further assess the role of these dummy variables on the components of consumption. For instance, according to results of the IHS double-hurdle model, rural households are 5 percent more likely to consume pork and, conditional on consumption, consume about 0.07 pound more per week than others. Overall, based on the effect on the unconditional level, rural households consume about 0.18 pound more per week than others. The effects of all other (significant) dummy variables can be interpreted in the same manner. Again, the average effects derived from both models are very close.

Cheese

Estimation of the IHS double-hurdle model and IHS infrequency-of-purchase model for cheese used the same procedures as for pork. The parameter estimates for both models are presented in Table 6. As a result of a specification search, income was found significant in the heteroskedasticity equation for both models. Thus, for cheese, the assumption of homoskedastic errors was also rejected and the specification of heteroskedastic error is warranted. The IHS parameter was significantly different from zero for both models, which suggests that the normal (untransformed) specifications are inappropriate. Contrary to the findings for pork, however, the log-likelihood values for both models were very close and the nonnested LR test result (Vuong's $z = 0.78$) indicated no evidence that one model is preferable over the other. For both models, a number of variables had different effects on participation and consumption. For instance, the age composition variables "aged 10-19," "aged 45-64," and "aged ≥ 65" were significant in the consumption equation but not in the participation (purchase) equation. The signs of the regional dummy "Midwest" were even opposite in the participation and consumption equations. For these variables, the effects on probability and level of consumption can better be understood by examining the elasticities.

The elasticities computed for the IHS double-hurdle model and the IHS infrequency-of-purchase model are presented in Table 7. The elasticities are extremely close between the two models. In particular, none of the elasticities in one model differ from the corresponding model by more than one standard deviation (of the elasticity in either model). Thus, although the IHS double-hurdle model and the IHS infrequency-of-purchase model provide different underlying justifications for the zero observations, both models seem to fit the data equally well and generate similar elasticities.

As for pork, the elasticities of probability, conditional level, and unconditional level with respect to the age composition variable "aged 20-44" are much greater than the corresponding elasticities for the other age composition variables. As expected, relative to the white majority, blacks and Asians are less likely to consume cheese, and conditional on consumption, they also consume less. Income has positive effects on the probability and levels of consumption. In particular, according to either model, a 1 percent increase in income increases the probability of cheese consumption by only 0.06 percent, and conditional on consumption, increases the level of consumption by only 0.06 percent. Overall, the elasticity of the unconditional level with respect to income is only 0.12 (0.11 percent according to the IHS infrequency-of-purchase model).

Finally, the effects of binary variables are presented in Table 8. According to the IHS double-hurdle model, Asian households are 0.37 percent less likely to consume cheese than with, and conditional on consumption, these households consume 0.37 pound less per week. Overall, the effect of ethnicity on the unconditional level suggests that Asian households consume about 0.61 pound less per week than white households. The interpretation of effects of the other variables is similar. The effects of all variables are very close between the two models.

Concluding Remarks

When parameterization and distribution of the Tobit model are matters of concern for empirical analysis of consumption, generalizations of the consumption model can be used. The double-hurdle model and infrequency-of-purchase model are popular extensions of the Tobit model. We have generalized these models further by using the inverse hyperbolic sine transformation to accommodate the nonnormal distribution of random errors.

For the commodities and type of data considered in this study, economic nonconsumption, conscientious abstention, and infrequency of purchases are all plausible causes of zero observations. While it is difficult to accommodate all these causes of zeros in one unified framework, the separate models we estimated account for the zero observations very well for pork and cheese. Although the IHS double-hurdle model was found to be a better characterization of the data-generating process than the IHS infrequency-of-purchase model, both models suggest very similar elasticities.

We find own-price effects on the probability and level of consumption to be negative and significant, and income and cross-price effects not to be significant. Household age composition, education, gender of meal planner, and race are among the demographic variables that affect consumption.

Table 1. Sample statistics: Household pork consumption

	Full sample		Consuming households	
Variable	Mean	Std. dev.	Mean	Std. dev.
Weekly pork consump. (lb.)	2.027	3.208	2.891	3.490
Income (weekly, hundred $)	5.241	4.470	5.313	4.551
Education of head (years)	2.372	0.824	2.341	0.806
Price of pork ($/lb.)	2.064	0.678	2.054	0.800
Price of other meat ($/lb.)	1.791	0.787	1.766	0.759
Household age composition				
≤ 9 yr.	0.486	0.848	0.534	0.878
10-19 yr.	0.418	0.800	0.474	0.847
20-44 yr.	1.026	0.924	1.069	0.928
45-64 yr.	0.528	0.767	0.557	0.782
≥ 65 yr.	0.335	0.632	0.326	0.634
Dummy variables (yes=1, no=0)				
Regions				
Northeast	0.205		0.196	
Midwest	0.264		0.278	
South	0.345		0.360	
West (reference)	0.186		0.166	
Race				
White	0.851		0.838	
Others (reference)	0.149		0.162	
Hispanic	0.036		0.037	
Female meal planner	0.788		0.804	
Food stamp recipient	0.072		0.077	
Rural	0.296		0.317	
Homeowner	0.678		0.686	
Sample size	4,198		2,943	

SOURCE: Compiled from Nationwide Food Consumption Survey (1987-88), USDA 1992.

Table 2. ML estimates: Household pork consumption

	IHS double-hurdle			IHS infrequency-of-purchase		
Variable	Particip.	Consump.	Het.	Purchase	Consump.	Het.
Constant	-0.052	0.940***	-1.022***	0.002	0.604***	-1.324***
	(0.121)	(0.097)	(0.148)	(0.117)	(0.065)	(0.157)
≤ 9 yr.	0.148***	0.064***	0.010	0.129***	0.068***	0.018
	(0.030)	(0.014)	(0.018)	(0.029)	(0.014)	(0.018)
10-19 yr.	0.179***	0.090***	0.028*	0.160***	0.090***	0.036**
	(0.030)	(0.017)	(0.017)	(0.029)	(0.017)	(0.017)
20-44 yr.	0.188***	0.130***	-0.041*	0.145***	0.121***	-0.032
	(0.035)	(0.024)	(0.022)	(0.034)	(0.022)	(0.023)
45-64 yr.	0.282***	0.157***	0.010	0.236***	0.154***	0.025
	(0.037)	(0.028)	(0.024)	(0.035)	(0.028)	(0.025)
≥ 65 yr.	0.254***	0.105***	-0.011	0.215***	0.112***	0.001
	(0.047)	(0.023)	(0.031)	(0.045)	(0.023)	(0.032)
Education of head (years)	-0.051*	-0.027***		-0.033	-0.025***	
	(0.027)	(0.010)		(0.026)	(0.009)	
Northeast	0.085	0.107***		0.078	0.095***	
	(0.067)	(0.028)		(0.064)	(0.024)	
Midwest	0.240***	0.067***		0.228***	0.088***	
	(0.064)	(0.024)		(0.062)	(0.023)	
South	0.214***	0.087***		0.201***	0.099***	
	(0.062)	(0.025)		(0.060)	(0.024)	
Female meal planner	0.102**	0.001		0.102**	0.018	
	(0.051)	(0.017)		(0.049)	(0.015)	
Rural	0.153***	0.013		0.145***	0.030**	
	(0.049)	(0.015)		(0.047)	(0.014)	
White	-0.133**	-0.127***		-0.107*	-0.112***	
	(0.064)	(0.029)		(0.063)	(0.025)	
Food stamp recipient		0.022			0.019	
		(0.028)			(0.022)	
Homeowner		0.002			0.002	
		(0.016)			(0.013)	
Hispanic		0.014			0.011	
		(0.035)			(0.027)	
Income		0.001			0.001	
		(0.002)			(0.001)	
Price of pork		-0.114***			-0.086***	
		(0.019)			(0.014)	
Price of other meat		-0.010			-0.008	
		(0.010)			(0.008)	
θ	2.316***			3.021***		
	(0.387)			(0.540)		
Log-likelihood	-8014.075			-8023.919		

Notes: Asymptotic standard errors in parentheses. Asterisks indicate levels of significance: *** = 0.01, ** = 0.05, and * = 0.10.

Table 3. Elasticities with respect to exogenous variables: Household pork consumption

	IHS double-hurdle			IHS infrequency-of-purchase		
Variable	Proba-bility	Cond. level	Uncond. level	Proba-bility	Cond. level	Uncond. level
≤ 9 yr.	0.0352***	0.0765***	0.1117***	0.0312***	0.0770***	0.1082***
	(0.0070)	(0.0121)	(0.0140)	(0.0068)	(0.0122)	(0.0137)
10-19 yr.	0.0366***	0.0960***	0.1326***	0.0332***	0.0940***	0.1272***
	(0.0059)	(0.0101)	(0.0117)	(0.0057)	(0.0103)	(0.0113)
20-44 yr.	0.0979***	0.2838***	0.3818***	0.0790***	0.2858***	0.3649***
	(0.0169)	(0.0312)	(0.0354)	(0.0162)	(0.0316)	(0.0349)
45-64 yr.	0.0737***	0.1978***	0.2715***	0.0631***	0.1994***	0.2625***
	(0.0092)	(0.0180)	(0.0202)	(0.0087)	(0.0183)	(0.0200)
≥ 65 yr.	0.0419***	0.0797***	0.1217***	0.0363***	0.0806***	0.1169***
	(0.0074)	(0.0133)	(0.0152)	(0.0072)	(0.0135)	(0.0151)
Education in years	-0.0597*	-0.1516***	-0.2113***	-0.0408	-0.1478***	-0.1886***
	(0.0306)	(0.0530)	(0.0609)	(0.0295)	(0.0533)	(0.0598)
Northeast	0.0091	0.0515***	0.0606***	0.0087	0.0517***	0.0604***
	(0.0065)	(0.0109)	(0.0126)	(0.0062)	(0.0110)	(0.0126)
Midwest	0.0308***	0.0414***	0.0722***	0.0297***	0.0425***	0.0722***
	(0.0081)	(0.0132)	(0.0153)	(0.0078)	(0.0133)	(0.0152)
South	0.0361***	0.0704***	0.1064***	0.0348***	0.0718***	0.1066***
	(0.0101)	(0.0167)	(0.0193)	(0.0098)	(0.0168)	(0.0192)
Female meal planner	0.0382**	0.0024	0.0406	0.0386**	0.0061	0.0447
	(0.0192)	(0.0313)	(0.0365)	(0.0185)	(0.0314)	(0.0364)
Rural	0.0216***	0.0089	0.0305**	0.0208***	0.0070	0.0278**
	(0.0068)	(0.0103)	(0.0123)	(0.0067)	(0.0104)	(0.0118)
White	-0.0577**	-0.2526***	-0.3103***	-0.0488*	-0.2475***	-0.2963***
	(0.0258)	(0.0398)	(0.0476)	(0.0256)	(0.0400)	(0.0456)
Food stamp recipient	0.0001	0.0038	0.0039	0.0000	0.0042	0.0042
	(0.0001)	(0.0047)	(0.0048)	(0.0000)	(0.0047)	(0.0047)
Homeowner	0.0000	0.0029	0.0029	0.0000	0.0036	0.0036
	(0.0004)	(0.0261)	(0.0266)	(0.0000)	(0.0262)	(0.0262)
Hispanic	0.0000	0.0012	0.0012	0.0000	0.0012	0.0012
	(0.0000)	(0.0030)	(0.0031)	(0.0000)	(0.0030)	(0.0030)
Income	0.0002	0.0103	0.0105	0.0000	0.0114	0.0114
	(0.0004)	(0.0220)	(0.0223)	(0.0000)	(0.0221)	(0.0221)
Price of pork	-0.0090**	-0.5520***	-0.5610***	0.0000**	-0.5405***	-0.5405***
	(0.0037)	(0.0434)	(0.0432)	(0.0000)	(0.0439)	(0.0439)
Price of other meat	-0.0007	-0.0418	-0.0425	0.0000	-0.0427	-0.0427
	(0.0007)	(0.0407)	(0.0414)	(0.0000)	(0.0408)	(0.0408)

Notes: Asymptotic standard errors in parentheses. Asterisks indicate levels of significance: *** = 0.01, ** = 0.05, and * = 0.10.

Table 4. Average effects of binary variables: Household pork consumption

	IHS double-hurdle			IHS infrequency-of-purchase		
Variable	Proba-bility	Cond. level	Uncond. level	Proba-bility	Cond. level	Uncond. level
		(pounds)	(pounds)		(pounds)	(pounds)
Northeast	0.0306	0.6752	0.5691	0.0292	0.6698	0.5610
Midwest	0.0801	0.4046	0.5024	0.0773	0.4256	0.5088
South	0.0729	0.5245	0.5667	0.0703	0.5364	0.5670
Female meal planner	0.0351	0.0076	0.0925	0.0355	0.0152	0.0974
Rural	0.0510	0.0752	0.1818	0.0492	0.0630	0.1666
White	-0.0463	-0.8213	-0.7268	-0.0388	-0.7968	-0.6855
Food stamp recipient	0.0006	0.1337	0.0967	0.0008	0.1451	0.1055
Homeowner	0.0000	0.0106	0.0077	0.0001	0.0131	0.0095
Hispanic	0.0004	0.0803	0.0581	0.0005	0.0849	0.0618

Table 5. Sample statistics: Household cheese consumption

Variable	Full sample Mean	Full sample Std. dev.	Consuming households Mean	Consuming households Std. dev.
Weekly cheese consump. (lb.)	1.055	1.180	1.288	1.183
Weekly income (hundred $)	5.260	4.519	5.554	4.661
Price of cheese ($/lb.)	2.252	0.772	2.250	0.851
Price of other dairy products ($/lb.)	0.405	0.252	0.400	0.244
Household age composition				
≤ 9 yr.	0.487	0.849	0.518	0.869
10-19 yr.	0.419	0.801	0.429	0.802
20-44 yr.	1.028	0.925	1.071	0.927
45-64 yr.	0.528	0.768	0.534	0.775
≥ 65 yr.	0.335	0.631	0.318	0.624
Dummy variables (yes=1, no=0)				
Regions				
Northeast	0.204		0.213	
Midwest	0.263		0.268	
South	0.346		0.322	
West (reference)	0.187		0.197	
Race				
Black	0.116		0.091	
Asian	0.010		0.006	
All others (reference)	0.874		0.903	
Ethnicity				
Hispanic	0.037		0.035	
All others (reference)	0.963		0.965	
Female food planner	0.788		0.796	
Food stamp recipient	0.072		0.064	
Sample size	4,217		3,454	

SOURCE: See Table 1.

Table 6. ML estimates: Household cheese consumption

	IHS double-hurdle			IHS infrequency-of-purchase		
Variable	Particip.	Consump.	Het.	Purchase	Consump.	Het.
Constant	0.733***	0.554***	-1.463***	0.789***	0.441***	-1.670***
	(0.105)	(0.035)	(0.100)	(0.097)	(0.031)	(0.109)
≤ 9 yr.	0.123***	0.047***		0.113***	0.046***	
	(0.036)	(0.008)		(0.035)	(0.007)	
10-19 yr.	0.011	0.062***		-0.006	0.051***	
	(0.033)	(0.009)		(0.031)	(0.007)	
20-44 yr.	0.117***	0.091***		0.111***	0.084***	
	(0.042)	(0.012)		(0.039)	(0.011)	
45-64 yr.	0.064	0.092***		0.047	0.081***	
	(0.045)	(0.012)		(0.041)	(0.011)	
≥ 65 yr.	0.042	0.063***		0.024	0.054***	
	(0.052)	(0.011)		(0.048)	(0.010)	
Black	-0.555***	-0.103***		-0.515***	-0.137***	
	(0.074)	(0.019)		(0.069)	(0.020)	
Asian	-1.071***	-0.125**		-0.961***	-0.213***	
	(0.199)	(0.055)		(0.178)	(0.055)	
Income	0.039***	0.003**	0.006**	0.034***	0.004***	0.008***
	(0.008)	(0.001)	(0.003)	(0.008)	(0.001)	(0.003)
Northeast	-0.105	0.012		-0.130	0.002	
	(0.086)	(0.013)		(0.081)	(0.012)	
Midwest	-0.209***	0.028**		-0.221***	0.009	
	(0.080)	(0.013)		(0.076)	(0.011)	
South	-0.295***	-0.005		-0.300***	-0.024**	
	(0.076)	(0.012)		(0.071)	(0.011)	
Hispanic	-0.334***	0.001		-0.309***	-0.023	
	(0.122)	(0.023)		(0.116)	(0.021)	
Female food planner	0.150**	-0.017		0.158***	-0.002	
	(0.059)	(0.011)		(0.055)	(0.009)	
Food stamp recipient		0.039**			0.031**	
		(0.018)			(0.015)	
Price of cheese		-0.082***			-0.067***	
		(0.009)			(0.008)	
Price of other dairy products		-0.021			-0.016	
		(0.018)			(0.015)	
θ	2.985***			3.641***		
	(0.344)			(0.450)		
Log-likelihood	-5617.463			-5619.383		

Notes: Asymptotic standard errors in parentheses. Asterisks indicate levels of significance: *** = 0.01, ** = 0.05, and * = 0.10.

Table 7. Elasticities with respect to exogenous variables: Household cheese consumption

	IHS double-hurdle			IHS infrequency-of-purchase		
Variable	Proba-bility	Cond. level	Uncond. level	Proba-bility	Cond. level	Uncond. level
≤ 9 yr.	0.0190***	0.0705***	0.0895***	0.0178***	0.0681***	0.0859***
	(0.0049)	(0.0086)	(0.0099)	(0.0047)	(0.0087)	(0.0095)
10-19 yr.	0.0037	0.0796***	0.0833***	0.0018	0.0806***	0.0824***
	(0.0038)	(0.0071)	(0.0081)	(0.0036)	(0.0071)	(0.0079)
20-44 yr.	0.0426***	0.2880***	0.3306***	0.0415***	0.2931***	0.3346***
	(0.0119)	(0.0225)	(0.0254)	(0.0111)	(0.0227)	(0.0248)
45-64 yr.	0.0141**	0.1495***	0.1636***	0.0118*	0.1532***	0.1650***
	(0.0064)	(0.0123)	(0.0138)	(0.0061)	(0.0124)	(0.0135)
≥ 65 yr.	0.0060	0.0654***	0.0714***	0.0043	0.0661***	0.0704***
	(0.0048)	(0.0096)	(0.0106)	(0.0045)	(0.0096)	(0.0106)
Black	-0.0192***	-0.0370***	-0.0562***	-0.0185***	-0.0431***	-0.0616***
	(0.0024)	(0.0056)	(0.0061)	(0.0023)	(0.0057)	(0.0064)
Asian	-0.0032***	-0.0039**	-0.0071***	-0.0030***	-0.0054***	-0.0084***
	(0.0006)	(0.0017)	(0.0018)	(0.0005)	(0.0018)	(0.0020)
Income	0.0564***	0.0648***	0.1212***	0.0507***	0.0563***	0.1070***
	(0.0116)	(0.0202)	(0.0232)	(0.0109)	(0.0213)	(0.0222)
Northeast	-0.0058	0.0075	0.0017	-0.0073	0.0091	0.0018
	(0.0049)	(0.0084)	(0.0096)	(0.0047)	(0.0086)	(0.0091)
Midwest	-0.0147**	0.0224**	0.0076	-0.0159***	0.0253**	0.0094
	(0.0059)	(0.0103)	(0.0117)	(0.0057)	(0.0105)	(0.0112)
South	-0.0288***	-0.0051	-0.0339**	-0.0300***	-0.0028	-0.0327**
	(0.0073)	(0.0132)	(0.0149)	(0.0070)	(0.0134)	(0.0144)
Hispanic	-0.0034***	0.0001	-0.0034	-0.0033***	-0.0001	-0.0033
	(0.0013)	(0.0026)	(0.0029)	(0.0012)	(0.0026)	(0.0030)
Female meal planner	0.0319**	-0.0405	-0.0086	0.0345***	-0.0410	-0.0065
	(0.0130)	(0.0259)	(0.0287)	(0.0125)	(0.0259)	(0.0288)
Food stamp recipient	0.0003*	0.0086**	0.0089**	0.0000*	0.0084**	0.0084**
	(0.0001)	(0.0040)	(0.0041)	(0.0000)	(0.0040)	(0.0041)
Price of cheese	-0.0176***	-0.5655***	-0.5831***	0.0000***	-0.5655***	-0.5655***
	(0.0048)	(0.0357)	(0.0355)	(0.0000)	(0.0358)	(0.0358)
Price of other dairy	-0.0008	-0.0268	-0.0277	0.0000	-0.0248	-0.0248
	(0.0007)	(0.0224)	(0.0231)	(0.0000)	(0.0226)	(0.0226)

Notes: Asymptotic standard errors in parentheses. Asterisks indicate levels of significance: *** = 0.01, ** = 0.05, and * = 0.10.

Table 8. Average effects of binary variables: Household cheese consumption

Variable	IHS double-hurdle			IHS infrequency-of-purchase		
	Proba-bility	Cond. level	Uncond. level	Proba-bility	Cond. level	Uncond. level
		(pounds)	(pounds)		(pounds)	(pounds)
Black	-0.1684	-0.3300	-0.4303	-0.1663	-0.3345	-0.4289
Asian	-0.3723	-0.3749	-0.6120	-0.3622	-0.3155	-0.5676
Northeast	-0.0245	0.0433	0.0071	-0.0314	0.0541	0.0085
Midwest	-0.0496	0.1010	0.0244	-0.0536	0.1179	0.0346
South	-0.0729	-0.0171	-0.0990	-0.0760	-0.0030	-0.0896
Hispanic	-0.0905	0.0024	-0.1038	-0.0851	0.0177	-0.0837
Female meal planner	0.0356	-0.0608	-0.0082	0.0385	-0.0636	-0.0080
Food stamp recipient	0.0027	0.1465	0.1264	0.0025	0.1395	0.1204

APPENDIX A. DERIVATIVES OF THE LOG-LIKELIHOOD FUNCTION: IHS DOUBLE-HURDLE MODEL

To simplify the notations, denote

$$z_1 = \boldsymbol{z}_t\boldsymbol{\alpha}, \qquad z_2 = \frac{\boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}, \qquad z_3 = \frac{T(y_t) - \boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}$$

$$\phi_1 = \phi(\boldsymbol{z}_t\boldsymbol{\alpha}), \qquad \phi_2 = \phi\left(\frac{\boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right), \qquad \phi_3 = \phi\left[\frac{T(y_t) - \boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right]$$

$$\Phi_1 = \Phi(\boldsymbol{z}_t\boldsymbol{\alpha}), \qquad \Phi_2 = \Phi\left(\frac{\boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right), \qquad \Phi_3 = \Phi\left[\frac{T(y_t) - \boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right].$$

The log-likelihood function is

$$\log \mathrm{L} = \sum_0 \left[\log(1 - \Phi_1\Phi_2)\right] + \sum_+ \left[\log\Phi_1 + \log\phi_3 - \log\sigma_t - \frac{1}{2}\log\left(1 + \theta^2 y_t^2\right)\right].$$

The derivatives of the log-likelihood function are

$$\frac{\partial \log \mathrm{L}}{\partial \boldsymbol{\alpha}} = \sum_0 \left(\frac{-\phi_1\Phi_2}{1 - \Phi_1\Phi_2}\right)\boldsymbol{z}_t + \sum_+ \left(\frac{\phi_1}{\Phi_1}\right)\boldsymbol{z}_t$$

$$\frac{\partial \log \mathrm{L}}{\partial \boldsymbol{\beta}} = \sum_0 \frac{1}{\sigma_t}\left(\frac{-\Phi_1\phi_2}{1 - \Phi_1\Phi_2}\right)\boldsymbol{x}_t + \sum_+ \left(\frac{1}{\sigma_t} z_3\right)\boldsymbol{x}_t$$

$$\frac{\partial \log \mathrm{L}}{\partial \boldsymbol{\gamma}} = \sum_0 \frac{1}{\sigma_t}\left(\frac{\Phi_1\phi_2}{1 - \Phi_1\Phi_2} z_2\right)\boldsymbol{h}_t + \sum_+ \left(\frac{1}{\sigma_t}\right)\left(z_3^2 - 1\right)\boldsymbol{h}_t$$

$$\frac{\partial \log \mathrm{L}}{\partial \theta} = \sum_+ \left\{ - z_3\left(\frac{1}{\theta\,\sigma_t}\right)\left[\frac{y_t}{\left[1 + (\theta y_t)^2\right]^{1/2}} - T(y_t)\right] - \frac{\theta y_t^2}{1 + (\theta y_t)^2}\right\}$$

APPENDIX B. DERIVATIVES OF THE LOG-LIKELIHOOD FUNCTION: IHS INFREQUENCY-OF-PURCHASE MODEL

Denote

$$z_1 = \boldsymbol{w}_t \boldsymbol{\delta}, \qquad z_2 = \frac{\boldsymbol{x}_t \boldsymbol{\beta}}{\sigma_t}, \qquad z_3 = \frac{T[\Phi_1(\boldsymbol{w}_t\boldsymbol{\delta})y_t] - \boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}$$

$$\phi_1 = \phi(\boldsymbol{w}_t\boldsymbol{\delta}), \qquad \phi_2 = \phi\left(\frac{\boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right), \qquad \phi_3 = \phi\left[\frac{T[\Phi_1(\boldsymbol{w}_t\boldsymbol{\delta})y_t] - \boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right]$$

$$\Phi_1 = \Phi(\boldsymbol{w}_t\boldsymbol{\delta}), \qquad \Phi_2 = \Phi\left(\frac{\boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right), \qquad \Phi_3 = \Phi\left[\frac{T[\Phi_1(\boldsymbol{w}_t\boldsymbol{\delta}_t)y_t] - \boldsymbol{x}_t\boldsymbol{\beta}}{\sigma_t}\right].$$

The log-likelihood function can be written as

$$\log L = \sum_0 \log(1 - \Phi_1\Phi_2) + \sum_+ \left\{2\log\Phi_1 + \log\phi_3 - \log\sigma_t - \frac{1}{2}\log\left[1 + (\Phi(\boldsymbol{w}_t\boldsymbol{\delta}))^2\theta^2 y_t^2\right]\right\}.$$

The derivatives of the log-likelihood function are

$$\frac{\partial \log L}{\partial \boldsymbol{\delta}} = \sum_0 \left(\frac{-\phi_1\Phi_2}{1-\Phi_1\Phi_2}\right)\boldsymbol{w}_t + \sum_+ \left\{2\frac{\phi_1}{\Phi_1} - \frac{1}{\sigma_t}z_3\frac{\phi_1 y_t}{\left[1+(\theta\Phi_1 y_t)^2\right]^{1/2}} - \frac{\theta^2 y_t^2\Phi_1\phi_1}{1+(\theta\Phi_1 y_t)^2}\right\}\boldsymbol{w}_t$$

$$\frac{\partial \log L}{\partial \boldsymbol{\beta}} = \sum_0 \frac{1}{\sigma_t}\left(\frac{-\Phi_1\phi_2}{1-\Phi_1\Phi_2}\right)\boldsymbol{x}_t + \sum_+ \left(\frac{1}{\sigma_t}z_3\right)\boldsymbol{x}_t$$

$$\frac{\partial \log L}{\partial \boldsymbol{\gamma}} = \sum_0 \frac{1}{\sigma_t}\left(\frac{\Phi_1\phi_2}{1-\Phi_1\Phi_2}z_2\right)\boldsymbol{h}_t + \sum_+ \left(\frac{1}{\sigma_t}\right)\left(z_3^2 - 1\right)\boldsymbol{h}_t$$

$$\frac{\partial \log L}{\partial \theta} = \sum_+ \left\{-z_3\left(\frac{1}{\theta\sigma_t}\right)\left[\frac{\Phi_1 y_t}{\left[1+(\theta\Phi_1 y_t)^2\right]^{1/2}} - T(\Phi_1 y_t)\right] - \frac{\theta\Phi_1^2 y_t^2}{1+(\theta\Phi_1 y_t)^2}\right\}$$

ENDNOTES

1. One caveat with the Box-Cox transformation is that the transformed random variable cannot strictly be normal unless the Box-Cox parameter equals zero, and that with such inherent nonnormality, the parameter estimates are inconsistent (Amemiya and Powell 1981). Comparing results with alternative transformations may be interesting. However, we were not successful in estimating the Box-Cox models.
2. Such equivalence obviously holds even with the IHS transformation considered here because the transformation occurs only in the models' Tobit component.
3. The use of unit values for broad commodity groups poses a potential problem with heterogeneity of the quality (or composition) of the commodity bundle. The heterogeneity might be accounted for by directly modeling quality effects (Deaton 1988) or by using independent price data, such as regional price information from the Bureau of Labor Statistics. Given the complexity of the econometric models considered here, direct modeling of quality effects may not be viable.
4. The asymptotic standard errors of parameters derived from the Hessian and the covariance matrix of the gradient (Berndt et al. 1974) are very similar.
5. The ML probit estimates and LR test supporting the specification of the final probit equation are available upon request.
6. For the purpose of discussion, unless indicated otherwise, a parameter or elasticity estimate is considered "significant" if it is significant at the 0.10 level or lower, that is, if p-value $\leq$ 0.10.
7. The NFCS data do not allow such inquiry. However, the effects of consumer attitudes and health knowledge on meat consumption may be worthy of further investigation.

REFERENCES

Amemiya, T. *Advanced Econometrics*. Cambridge: Harvard University Press, 1985.

Amemiya, T., and M. Boskin. "Regression Analysis When Dependent Variable is Trun-cated Lognormal, With an Application to the Determinants of the Duration of Welfare Dependency." *International Economic Review* 15 (June 1974): 485-96.

Amemiya, T., and J.L. Powell. "A Comparison of the Box-Cox Maximum Likelihood Estimator and the Non-Linear Two Stage Least Squares Estimator." *Journal of Econometrics* 17 (December 1981): 351-81.

Arabmazar, A., and P. Schmidt. "Further Evidence on the Robustness of the Tobit Estimator to Heteroskedasticity." *Journal of Econometrics* 17 (November 1981): 253-58.

Arabmazar, A., and P. Schmidt. "An Investigation of the Robustness of the Tobit Estimator to Non-normality." *Econometrica* 50 (July 1982): 1055-63.

Atkinson, A.B., J. Gomulka, and N.H. Stern. "Household Expenditure on Tobacco 1970-1980: Evidence From the Family Expenditure Survey." ESRC Programme on Taxation, Incentives, and the Distribution of Income, London School of Economics, Discussion Paper No. 60, 1984.

Atkinson, A.B., J. Gomulka, and N.H. Stern. "Spending on Alcohol: Evidence From the Family Expenditure Survey 1970-1983." *Economic Journal* 100 (September 1990): 808-27.

Berndt, E.K., B.H. Hall, R.E. Hall, and J.A. Hausman. "Estimation and Inference in Nonlinear Structural Models." *Annals of Economic and Social Measurement* 3 (October 1974): 653-65.

Blaylock, J.R., and W.N. Blisard. "U.S. Cigarette Consumption: The Case of Low-Income Women." *American Journal of Agricultural Economics* 74 (August 1992): 698-705.

__________. "Wine Consumption by US Men." *Applied Economics* 24 (May 1993a): 645-51.

__________. "Women and the Demand for Alcohol: Estimating Participation and Consumption." *Journal of Consumer Affairs* 27 (Winter 1993b): 319-34.

Blisard, W.N., and J.R. Blaylock. "Distinguishing Between Market Participation and Infrequency-of-Purchase Models of Butter Demand." *American Journal of Agricultural Economics* 75 (May 1993): 314-20.

Blundell, R.W. and C. Meghir. "Bivariate Alternatives to the Univariate Tobit Model." *Journal of Econometrics* 34 (January/February 1987): 179-200.

Box, G.E.P., and D.R. Cox. "An Analysis of Transformations." *Journal of the Royal Statistical Society Series B* 26 (2) (1964): 211-52.

Burbidge, J.B., L. Magee, and A.L. Robb. "Alternative Transformations to Handle Extreme Values of the Dependent Variable." *Journal of the American Statistical Association* 83 (March 1988): 123-27.

Cragg, J.G. "Some Statistical Models for Limited Dependent Variables with Applications to the Demand for Durable Goods." *Econometrica* 39 (September 1971): 829-44.

Deaton, A. "Quality, Quantity, and Spatial Variation of Price." *American Economic Review* 78 (June 1988): 418-30.

Deaton, A.S., and M. Irish. "A Statistical Model for Zero Expenditures in Household Budgets." *Journal of Public Economics* 23 (February/March 1984): 59-80.

Fuller, W.A. *Measurement Error Models*. New York: Wiley, 1987.

Goldfeld, S.M., R.E. Quandt, and H.F. Trotter. "Maximization by Quadratic Hill-climbing." *Econometrica* 34 (July 1966): 541-51.

Haines, P., D. Guilkey, and B. Popkin. "Modeling Food Consumption As a Two-Step Decision." *American Journal of Agricultural Economics* 70 (August 1988): 543-52.

Johnson, N.L. "Systems of Frequency Curves Generated by Methods of Transformation." *Biometrika* 36 (June 1949): 149-76.

Jones, A.M., and S.T. Yen. "A Box-Cox Double Hurdle Model." Discussion Paper No. W94/6, Institute for Fiscal Studies, London, 1994.

Lankford, R.H., and J.H. Wyckoff. "Modeling Charitable Giving Using A Box-Cox Standard Tobit Model." *Review of Economics and Statistics* 73 (August 1991): 460-70.

Lee, L-F., and G.S. Maddala. "The Common Structure of Tests for Selectivity Bias, Serial Correlation, Heteroscedasticity and Non-normality in the Tobit Model." *International Economic Review*, 26 (February 1985): 1-20.

Lin, T., and P. Schmidt. "A Test of the Tobit Specification Against an Alternative Suggested by Cragg." *Review of Economics and Statistics* 66 (February 1984): 174-77.

Maddala, G.S. *Limited-dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press, 1983.

McDonald, J.F., and R.A. Moffitt. "The Uses of Tobit Analysis." *Review of Economics and Statistics* 62 (May 1980): 318-21.

Poirier, D.J. "The Use of the Box-Cox Transformation in Limited Dependent Variable Models." *Journal of the American Statistical Association* 73 (June 1978): 284-87.

Pudney, S.E. "Estimating Engel Curves: A Generalization of the P-Tobit Model." *Finish Economic Papers* 1 (Autumn 1988): 129-47.

Robinson, P.M. "On the Asymptotic Properties of Estimators of Models Containing Limited Dependent Variables." *Econometrica* 50 (January 1982): 27-41.

Reynolds, A., and J.S. Shonkwiler. "Testing and Correcting for Distributional Misspecifications in the Tobit Model: An Application of the Information Matrix Test." *Empirical Economics* 16 (3) (1991): 313-23.

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

U.S. Department of Agriculture, Human Nutrition Information Service, Nationwide Food Consumption Survey 1987-88: Household Food Use. Documentation of Public-Use Tape, Washington, D.C., 1992.

Vuong, Q.H. "Likelihood Ratio Tests for Model Selection and Nonnested Hypotheses." *Econometrica* 57 (March 1989): 307-33.

Yen, S.T. "Working Wives and Food Away From Home: The Box-Cox Double Hurdle Model." *American Journal of Agricultural Economics* 75 (November 1993): 884-95.

__________. "Cross-Section Estimation of U.S. Demand for Alcoholic Beverage." *Applied Economics* 26 (April 1994): 381-92.

Yen, S.T., and A.M. Jones. "Individual Cigarette Consumption and Addiction: A Flexible Limited Dependent Variable Approach." Presented at the Third European Workshop on Econometrics and Health Economics, SESO, University of Antwerp, September 28-30, 1994.