Determinants of Household Expenditures on Alcohol

Steven T. Yen and Helen H. Jensen

Working Paper 95-WP 144
December 1995

Center for Agricultural and Rural Development
Iowa State University
578 Heady Hall
Ames, IA 50011

Steven T. Yen is a lecturer in the Department of Economics, University of Illinois at Urbana-Champaign; Helen H. Jensen is associate professor of economics and head of the Food and Nutritional Policy division, Center for Agricultural and Rural Development, Iowa State University.

The authors gratefully acknowledge the helpful comments and suggestions by two anonymous reviewers. Appreciation is also extended to Chun-Fu Chen and Ann Simpson for excellent programming assistance in data preparation. Journal Paper No. J-15962 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 2851.

ABSTRACT

This paper examines the determinants of household alcohol expenditures by using a nonnormal and heteroscedastic double-hurdle model to accommodate zero observations in the sample. The model is a generalization of the double-hurdle model estimated in previous studies of alcohol consumption. We also examine the effects of explanatory variables by calculating and decomposing the elasticities. Findings support the use of a more generalized error distribution. Income, region, education, and household demographics are among the significant determinants of alcohol expenditures.

DETERMINANTS OF HOUSEHOLD EXPENDITURES ON ALCOHOL

There is long-standing interest in the economic aspects of alcoholic beverages, although formal economic models of alcohol consumption are a more recent phenomenon (Hilton 1993). Findings on the price responses of alcohol demand have obvious policy implications ! in promoting tax revenues and controlling consumption through taxes (Ornstein 1980). The identification of non-price determinants of alcohol consumption is also important because such information often plays a crucial role in the prevention of alcohol abuse and alcohol-related problems (Hilton 1993).

There have been numerous studies on the demand for alcoholic beverages. Ornstein (1980) and Ornstein and Levy (1983) provide surveys of the earlier studies, and Leung and Phelps (1993) review the studies that have appeared during the last decade. In addition, Selvanathan (1991) reviews the recent literature on alcohol demand systems. Most of the earlier studies have been based on aggregate time series. However, a notable trend among the more recent studies on alcohol consumption and expenditures is the growing use of survey data. The examples include Atkinson, Gomulka, and Stern (1990), Blaylock and Blisard (1993), Heien and Pompelli (1989), Pompelli and Heien (1991), and Yen (1994).

The use of microdata allows examination of the effects of detailed household characteristics and provides the degrees of freedom to estimate a large number of parameters. However, microdata also present a unique problem: the sample often contains a significant proportion of observations with reported zero expenditures. Standard econometric techniques not accounting for this data feature lead to biased and inconsistent parameter estimates (Maddala 1983). Researchers have often used the Tobit model (Tobin 1958) to estimate demand relationships with limited dependent variables. However, parameterization of the Tobit model is known to be very restrictive, which makes it unpalatable for empirical analysis.

Recent analysts have used the double-hurdle model in modeling alcohol demand (Blaylock and Blisard 1993; Yen 1994). The double-hurdle model features two separate stochastic processes for participation and consumption; it allows for examining the determinants of both participation and consumption decisions and provides more useful insights on consumer behavior than the Tobit model. Applications of the double-hurdle model in other areas of food demand include Haines, Guilkey, and Popkin (1988), Reynolds (1990), and Wang and Jensen (1994).

The model specification and estimation procedure of the double-hurdle model require specification of the distribution for the error terms of the participation and consumption equations. At issue are a number of aspects related to the error distribution: (a) dependence between the error terms, (b) heteroscedasticity, and (c) shape of the distribution. In previous studies of alcohol consumption, the error terms were found to be independent (Blaylock and Blisard 1993), while evidence of nonnormality and heteroscedasticity of errors has been reported (Yen 1994).

The purpose of this paper is to evaluate determinants of alcohol expenditures, and to explain important and recent empirical developments and considerations in modeling consumer expenditure behavior. We investigate household expenditures on alcohol, using a nonnormal and heteroscedastic double-hurdle model, and we accommodate skewness in the error terms by incorporating the inverse hyperbolic sine (IHS) transformation (Burbidge, Magee, and Robb 1988) in the dependent variable. The resulting specification allows for a more flexible error distribution and nests a range of specifications considered in the empirical literature, including the standard (Gaussian) double-hurdle, the IHS Tobit (Reynolds and Shonkwiler 1991), and the Tobit models. The more flexible specification of the IHS Tobit is a methodological improvement over earlier approaches to estimating alcohol demand that allows the researcher to relax some of the restrictions of the standard Tobit model. Using household expenditure data from the 1989 and 1990 Bureau of Labor Statistics (BLS) Consumer Expenditure Diary Surveys, we present and compare results of the IHS double-hurdle and the IHS Tobit models and examine the consequences of the Tobit parameterization with a more flexible error specification.

Theoretical Model and Empirical Specification

Following the theory of consumer demand, a household makes choices among consumer goods by maximizing utility subject to a budget constraint. That is, it solves the following constrained utility maximization problem:

$$\operatorname*{Max}_{y}\ [u(\boldsymbol{y}, \boldsymbol{d}) \mid \boldsymbol{p}\,\boldsymbol{y} = m], \tag{1}$$

where $\boldsymbol{y}$ is a vector of consumer goods, $\boldsymbol{p}$ is a vector of corresponding prices, $\boldsymbol{d}$ is a vector of household characteristics, and m is the household budget. Assuming that the utility function $u(\bullet)$ is continuous, increasing, and quasi-concave, then the *notional demand* for a commodity, say alcohol, can be expressed as a demand function $f(\boldsymbol{p},\boldsymbol{d},m)$. Since price information was not collected in the Diary Surveys, we treat the surveys as cross-sections and assume all households face the same relative prices. The expenditure equation is denoted $g(\boldsymbol{d},m)$.[i]

The notional demand and expenditure are the results of utility maximization with only the budget constraint, given household characteristics. In practice the quantity and expenditure are also subject to a nonnegativity constraint. Therefore, the optimal expenditures (E) can be either an interior solution or a corner solution. That is,

$$E = \max\,[0, g(\boldsymbol{d}, m)]\,. \tag{2}$$

To operationalize the model assume a linear functional form for $g(\cdot)$ and assume the optimal outcome is observed with errors. Then, the expenditure model can be characterized as the Tobit model (Tobin 1958)

$$\begin{aligned} E_t &= x_t\beta + u_t \quad \text{if } x_t\beta + u_t > 0 \\ &= 0 \qquad \text{otherwise}, \end{aligned} \tag{3}$$

where, for household t, $\boldsymbol{x}_t$ is a vector of explanatory variables, $\boldsymbol{\beta}$ is a conformable parameter vector, and the error term u_t is distributed as $N(0,\sigma)$.

Within the Tobit framework, zero observations represent corner solutions in consumer choices. However, for commodities like alcohol, it is unlikely that all zeros in a sample represent corner solutions. On statistical grounds, the Tobit parameterization also imposes an unnecessary restriction on the data-generating process because the same set of variables ($\boldsymbol{x}_t$) and parameters ($\boldsymbol{\beta}$) determine both the discrete probability of a non-zero outcome and the level of positive expenditures. Consequently, the probability of a non-zero outcome is tied closely to the conditional density of the positive observations and this is an undesirable property. The Tobit model has been rejected for alcohol consumption (Blaylock and Blisard 1993; Yen 1994) and for other applications in food demand (Haines, Guilkey, and Popkin 1988; Reynolds 1990; Wang and Jensen 1994).

Recent analysts of alcohol demand have used the double-hurdle model, which allows separate parameterization of the participation and consumption decisions. The double-hurdle model features a participation equation $\boldsymbol{z}_t\boldsymbol{\alpha} + \varepsilon_t$ and an expenditure equation $\boldsymbol{x}_t\boldsymbol{\beta} + u_t$ such that

$$\begin{aligned} E_t &= x_t\beta + u_t \quad \text{if } z_t\alpha + \varepsilon_t > 0 \text{ and } x_t\beta + u_t > 0 \\ &= 0 \qquad \text{otherwise}. \end{aligned} \tag{4}$$

where $\boldsymbol{z}_t$ and $\boldsymbol{x}_t$ are vectors of explanatory variables, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are vectors of parameters, and ε_t and u_t are the error terms (Blundell and Meghir 1987; Cragg 1971). The structure (4) suggests that two mechanisms determine consumption. For positive expenditures to occur, two "hurdles" have to be overcome: to participate in the market (i.e., to be a potential consuming household), and to actually consume.

Estimation of the double-hurdle model requires specification of the error structure. One commonly made assumption is that the errors ε_t and u_t are independently and normally distributed (Haines, Guilkey, and Popkin 1988; Reynolds 1990). Blaylock and Blisard (1993) estimated the double-hurdle model for alcohol consumption by U.S. women by specifying a bivariate normal distribution for the error terms, allowing for dependence between the participation and consumption decisions; the results however suggested independence between the two decisions.

Besides normality, the error term u_t has been assumed to be homoscedastic in most of these applications. In limited dependent variable models, however, maximum-likelihood (ML) estimates based on the normality and homoscedasticity assumptions are inconsistent when either assumption is violated (Arabmarzar and Schmidt 1981, 1982).

Heteroscedasticity of errors can be accommodated by allowing the standard deviation σ to vary across observations and making it a function of observed variables. Nonnormality can be accomplished by considering nonnormal error distributions. However, like the standard Tobit model, the assumption of a specific nonnormal error distribution yields models that are subject to specification errors. Another approach to nonnormality is transformation of the dependent variable. By specifying a truncated normal distribution for the transformed dependent variable, the resulting model specification allows for skewness in the original (untransformed) dependent variable. In an analysis of household consumption of alcohol in the United States, Yen (1994) incorporated the Box-Cox transformation in Cragg's double-hurdle model (Cragg 1971) and allowed a heteroscedastic specification for the error terms. The findings in Yen (1994) justified the concerns for nonnormality and heteroscedasticity in modeling alcohol consumption.

One problem with the Box-Cox transformation is that it cannot be performed on random variables that can take on zero or negative values. In addition, it is not scale-invariant so empirical results may vary with the unit of measurement used. More seriously, the transformed random variable cannot strictly be normal unless the Box-Cox parameter equals zero, and, with such inherent nonnormality, the parameter estimates are inconsistent (Amemiya and Powell 1981). One transformation that is free from the drawbacks of the Box-Cox transformation is the IHS transformation. The IHS transformation has been proposed to accommodate nonnormal error terms in traditional regression models (Burbidge, Magee, and Robb 1988; MacKinnon and Magee 1990) and in limited dependent variable models (Horowitz and Neumann 1989). It has been applied to the Tobit model by Reynolds and Shonkwiler (1991) in modeling U.S. food demand. In this study, we incorporate the IHS transformation in the double-hurdle model and apply the model to household expenditures of alcohol in the United States. Thus, our proposed model is a generalization of the IHS Tobit model in that the Tobit parameterization is relaxed; it is a generalization of the double-hurdle model in that nonnormality and heteroscedasticity of errors are accommodated.

The IHS transformation on random variable v can be written as

$$\begin{aligned} T(v) &= \log[\,\theta v + (\theta^2 v^2 + 1)^{1/2}\,] / \theta \\ &= \sinh^{-1}(\theta v) / \theta, \end{aligned} \tag{5}$$

which is defined over all values of θ (Burbidge, Magee, and Robb 1988). Because the transformed variable is symmetric about 0 in θ, one can consider only $\theta \$ 0$. The transformation is linear when θ approaches zero and behaves logarithmically for large values of v over a wide range of values for θ (MacKinnon and Magee 1990). In addition, such transformation can be performed on random variables that can take on any (zero, positive, and negative) values.

Applying the IHS transformation to the dependent variable E, the double-hurdle model can be modified as

$$T(E_t) = x_t\beta + u_t \quad \text{if } z_t\alpha + \varepsilon_t > 0 \text{ and } x_t\beta + u_t > 0$$
$$= 0 \qquad \text{otherwise}. \tag{6}$$

With the IHS transformation on the dependent variable E, the error term u_t has a better chance of satisfying the normality and homoscedasticity assumptions. Building on the empirical evidence of independence between the participation and consumption reported in the alcohol demand literature, the error terms ε_t and u_t are assumed to be independently and normally distributed such that ε_t - $N(0,1)$ and u_t - $N(0,\sigma)$.

Consider a zero-truncated normal density for $T(E)$ (Johnson and Kotz 1970, p. 81). Then, by a transformation of variables from E to $T(E)$, the conditional density of E is

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

where $\Phi(\cdot)$ and $\varphi(\cdot)$ are the univariate standard normal distribution and density functions. In (7) both the IHS transformation $T(E)$ and the Jacobian of transformation $(1+\theta^2 E)^{-1/2}$ allow for skewness in the conditional density of E. Using (6) and (7), the sample likelihood function for the IHS double-hurdle model is

$$L = \prod_{E_t=0}\left\{1 - \Phi(z_t\alpha)\,\Phi\left(\frac{x_t\beta}{\sigma_t}\right)\right\}$$
$$\prod_{E_t>0}\left\{\Phi(z_t\alpha)\,\frac{1}{\sigma_t}\,\phi\left[\frac{T(E_t) - x_t\beta}{\sigma_t}\right]\frac{1}{\left(1+\theta^2 E_t^2\right)^{1/2}}\right\} \tag{8}$$

The likelihood function (8) nests the IHS Tobit model when $\Phi(\boldsymbol{z\alpha}) = 1$.[ii] In addition, imposing restriction $\theta = 0$ on the IHS double-hurdle and IHS Tobit models leads to the standard double-hurdle and Tobit models, respectively.

To accommodate heteroscedastic errors, the standard deviation σ_t can be specified as

$$\sigma_t = \exp(w_t\gamma), \tag{9}$$

where $\boldsymbol{w}_t$ is a vector of variables and $\boldsymbol{\gamma}$ is a conformable parameter vector.

In assessing the appropriateness of the double-hurdle model, IHS transformed or not, one might note that household survey data are often collected in a relatively short sampling period. Some households record zero expenditures during the sampling period on commodities they would nevertheless purchase if observed over a longer period of time. Such zero observations could be accommodated with the infrequency-of-purchase model (Blundell and Meghir 1987). However, for commodities like alcohol, the double-hurdle mechanism presented above may better account for the zero observations. Unfortunately, most survey data do not provide detailed enough information to distinguish the true causes of zero for each sample unit. One can, however, view the double-hurdle model as the reduced form of a structural model that augments the demand equation with separate hurdles for nonbehavioral sources of zeros, such as misreporting and infrequency of purchases (Jones and Posnett 1991). Thus, the double-hurdle model is also appropriate for modeling demand with zero observations resulting from infrequency of purchases. In fact, in the special case when the probability of participation is estimated as a constant (i.e., when z contains only a unity), the double-hurdle model is "observationally equivalent" to a model of infrequency of purchase with corner solutions (Deaton and Irish 1984). When the probability is specified as $\Phi(z\alpha)$ and is allowed to vary across observations, the double-hurdle model and infrequency-of-purchase model are only closely related, in that both specifications nest the Tobit model. While we focus on the IHS double-hurdle model in the current paper, we also report briefly the results of a test which rejects the IHS infrequency-of-purchase model.

Marginal Effects of Variables

In limited dependent variable models, the effects of explanatory variables must be evaluated at the mean of the dependent variables. For the standard (homoscedastic and truncated normal) Tobit model, McDonald and Moffitt (1980) suggest decomposition of the unconditional mean of the dependent variable into the probability (of a positive observation) and the conditional mean. The effects of explanatory variables on these components can then be assessed. For the IHS double-hurdle model considered here, the unconditional mean can be decomposed in like manner, although the double-hurdle parameterization, IHS transformation, and heteroscedasticity specification all complicate the expressions for these components of consumption and the marginal effects of variables on these components. The marginal effects of continuous variables can be obtained by differentiating the probability, conditional mean, and unconditional mean of consumption. Based on these marginal responses, the elasticities are straightforward and can be computed using the parameter estimates. For discrete variables the calculation of elasticities is not strictly appropriate. The effects of these variables can be computed as the finite changes in probability, conditional level, and unconditional level resulting from a change in value of these variables form zero to one, one at a time, *ceteris paribus*.

Data

The sample for the present study was drawn from the BLS 1989 and 1990 Consumer Expenditure Diary Surveys (U.S. Department of Commerce 1989, 1990). Each year the Diary Survey was completed by each sample consumer unit during two consecutive one-week periods. We included only households with complete two-week information. Households with missing information for important variables were excluded, as were households with negative after-tax income. This resulted in a final sample of 9552 households, of which 4643 were drawn from the 1989 Diary Survey, and 4909 from the 1990 Diary Survey. A total of 4411 households (or 46.18%) reported expenditures on alcohol during the two-week period.

Household expenditures on alcohol is used as the dependent variable.[iii] The explanatory variables selected include household age composition, income, age of the household head, and dummy variables indicating regions of residence, education, marital status, gender, and race of the household head, home ownership, and seasonality. Although price information was not collected in the Diary survey, the regional and seasonal dummies are expected to account for much of the regional and seasonal price variations and avoid potential misspecification.[iv]

Detailed definitions of all variables are presented in Table 1. The mean two-week expenditures on alcohol is \$11.66 for the full sample and \$25.25 for the consuming households. Sample statistics for all the variables for the full, nonconsumer, and consumer samples are presented in Table 2.

Empirical Results

We estimated the IHS double-hurdle and the IHS Tobit models and compare the results. However, as the infrequency-of-purchase model may also provide a plausible account for the zero observations, a heteroscedastic IHS infrequency-of-purchase model was also estimated and tested against the IHS double-hurdle model using a nonnested specification test procedure (Vuong 1989) and was rejected ($z = -2.05$, where z is distributed as $N(0,1)$).[v] While in the interest of space we focus on the preferred IHS double-hurdle model and the nested IHS Tobit model in the current paper, the development and empirical results of the IHS infrequency-of-purchase model are available upon request from the authors.

Parameter Estimates

The IHS double-hurdle model was estimated by maximizing the logarithm of the likelihood functions (8). The IHS Tobit model was estimated by maximizing a separate log-likelihood function, with the univariate normal probability $\Phi(\boldsymbol{z\alpha})$ removed from (8). The results for both models are presented in Table 3.

After an extensive search, three household composition variables were included in the heteroscedasticity equation. For both the participation and consumption equations, we started with an extensive list of variables, which included income, regional dummies, seasonal dummies, and other household demographic variables, and followed an empirical approach to some "exclusion conditions." All seasonal dummies were insignificant in the participation equation and therefore were excluded; the exclusion of these seasonal dummies from the participation equation was supported by the likelihood-ratio (LR) test (χ^2 = 2.18, d.f.= 3, p-value = 0.53).

Estimation results of the IHS double-hurdle model suggest that the IHS parameter (θ) is significantly different from zero. In addition, all three household composition variables are significant in the heteroscedasticity equation. These results suggest that the normal and homoscedastic double-hurdle model, commonly estimated in previous studies of alcohol consumption and expenditures, is misspecified. Also, among the other notable results of the parameter estimates are the conflicting effects of variables. In particular, variables Members 19-64, Northeast, Midwest, South, West, High school, and Married all have conflicting signs in the participation and consumption equations. In addition, although having the same sign, variables Members # 18 and College are significant in the consumption equation but not in the participation equation, whereas the variable Male is significant in the participation equation but not in the consumption equation. These conflicting and qualitatively different effects of variables are not allowed in the Tobit model.

Based on the log-likelihood values of the two models estimated, the LR test result suggests the rejection of the IHS Tobit model (χ^2 = 121.25, d.f. = 16, p-value < 0.001). The parameters for the IHS Tobit model are very well determined. That is, most variables are significant at the 0.10 level or lower. In addition, the IHS parameter (θ) is significant at the 0.10 level, as are the three household composition variables in the heteroscedasticity equation. Thus, the specification of nonnormal and heteroscedastic errors is justified.

As indicated above, the IHS transformation, heteroscedasticity specification, and the parameterization of both models all complicate the effects of variables, we turn to elasticities and average effects for a more careful examination of the results and comparison between the two models.

Elasticities and Average Effects

Using the ML estimates of both models, the elasticities of probability, conditional level, and unconditional level with respect to all continuous variables are computed at the sample means of all variables. The elasticity of probability indicates how a variable affects the likelihood (probability) to consume alcohol, while the elasticity of conditional level measures how a variable affects the expenditure level conditional on consumption (i.e., given that a decision is made to consume). Then, the elasticity of the unconditional level indicates the overall responsiveness of a household to a variable in the expenditures on alcohol. For statistical inferences, the standard errors for all elasticities are also computed by mathematical approximations (Yen 1994). The results are presented in Table 4.

Columns 2, 3, and 4 presents the elasticities computed from the IHS double-hurdle model. Most elasticities are significant at the 0.10 level or lower, which means that these elasticities are very well determined. Income has significant and positive effects on alcohol expenditures. In particular, a one percent increase in household income increases the probability of consumption by 0.21 percent, the conditional level of expenditures by 0.13 percent, and the unconditional level of expenditures by 0.34 percent. Thus, the effects of income on alcohol expenditures are small but positive. The elasticity of the unconditional level with respect to income is higher than the elasticity (0.16) reported by Blaylock and Blisard (1993) for U.S. women,[vi] and is very close, in reference to the standard error, to the elasticity (0.40) reported by Yen (1994). In terms of the decomposed effects, Yen (1994) reported a much higher elasticity of probability (0.37) and a much lower (and insignificant) elasticity of conditional level (0.04); these estimates are for alcohol consumed at home only.

The effects of numbers of household members of ages 19-64 and over 65 are not significant on the probability, but are significant on the conditional level of alcohol expenditures. Not surprisingly, the effects of the age of the household head are negative, suggesting that older households are less likely to consume alcohol and also consume less than others. Finally, the number of children (Members # 18) has a negative impact on all components (probability, conditional level, and unconditional level) of alcohol expenditures. These results indicate that household composition is an important determinant of alcohol expenditures.

Most of the elasticities computed from the IHS Tobit model are, in general, qualitatively similar (in terms of directions and statistical significance of effects) to those computed from the IHS double-hurdle model. Some elasticities are very close to those derived from the IHS double-hurdle model. For instance, the elasticities of the unconditional level of expenditures with respect to all three household composition variables are all within one standard deviation from those derived from the IHS double-hurdle model, as are the elasticities of the unconditional level with respect to Members # 18, Members $ 65, and income. The magnitudes of the income elasticities are notably close.

There are notable differences among the elasticities as well, however. For instance, unlike for the IHS double-hurdle model, the elasticity of probability of consumption with respect to Members 19-64 computed from the IHS Tobit model is significant and positive. In terms of magnitude, all elasticities with respect to Age for the IHS Tobit model are about two standard deviations from those computed for the IHS double-hurdle model. Interestingly, the effects of Members $ 65 are negative (though insignificant) on the probability but significant and positive on the conditional and unconditional levels of expenditures. Such conflicting effects of variables are not possible in a standard (normal and homoscedastic) Tobit model. The conflicting effects are possibly due to the fact that the nonlinear (IHS) transformation weakens the link between the discrete probability of a positive observation and the shape of the (now skewed) conditional distribution of the dependent variable; see also footnote 2. Thus, although the IHS Tobit model is not the preferred model in the current study, an important implication of these findings is that, in applications where zero observations truly represent corner solutions (so that the Tobit parameterization is relevant), the IHS transformation and the heteroscedasticity error specification included here may provide the flexibility necessary to model consumer behavior using the Tobit parameterization. The degree to which they will differ is an empirical question, however. In summary, comparisons of results from the IHS double-hurdle and the IHS Tobit models are mixed: some elasticities are similar, while others are notably different. The consequences of parametric misspecification are apparent.

The effects of discrete variables for both models are presented in Table 5. Except for marital status and the three seasonal dummies, all variables have positive effects on (probability, conditional level, and unconditional level of) consumption. According to the results of the IHS double-hurdle model, relative to the rural households, the reference group in the present analysis, urban households in the Midwest are about 8 percent more likely to consume alcohol, spend about \$2.93 more during the period on alcohol conditional on consumption (and about \$3.07 more overall). Households with white household heads are more likely to consume alcohol and, conditional on consumption, also consume more. This finding appears to differ somewhat from that of Blaylock and Blisard (1993), who concluded that black women have a much lower probability of being drinkers than white but, conditional on consumption, black women do not consume any differently from others. The interpretation of the effects of all other variables used in the IHS double-hurdle model are similar. The effects suggested by the IHS Tobit model are similar in direction but generally different in magnitudes from those suggested by the IHS double-hurdle model.

Concluding Remarks

This study provides an analysis of the determinants of household alcohol expenditures in the United States. The zero observations are addressed by generalizing the error distribution of the double-hurdle model. The resulting specification accommodates heteroscedastic and nonnormal errors, which have been overlooked in previous applications of the double-hurdle model in alcohol demand. We also examine the effects of explanatory variables on consumption in greater detail than has been presented in previous studies by calculating and decomposing the elasticities.

Our empirical findings suggest that the generalizations on the error distribution are warranted. We also test the Tobit model against the double-hurdle model, with more flexible error distribution on both models, and compared the elasticities calculated from both models. The standard errors calculated for these elasticities allow for a better assessment of whether the differences between these elasticities are statistically significant. The Tobit parameterization was rejected but, interestingly, based on the estimates from the two models, some elasticities are very similar both qualitatively and quantitatively, while others are notably different, between the two models. The lessons for subsequent expenditure analysis indicate that the approach to be taken is an empirical question. The traditional Tobit model may yield estimates of elasticities that are not very different in some cases. However, differences are significant for some of the estimates, and this result can only be based on the empirical analysis.

Based on the preferred IHS double-hurdle model, income, region, and household demographics such as household composition, education, home ownership, gender and age of household head, and race are among the significant determinants of household alcohol expenditures. The magnitudes of the income elasticities indicate relatively strong responses to changes in income. Changes in the quality or composition of types of alcohol purchased could explain part of these responses and are not examined here.

Higher education levels are associated with higher probabilities and levels of expenditures. Again, quality differences may be present, as well as increased acceptance of reporting alcoholic expenditures, or differences attributed to socialization. These differences, possibly associated with observed education differences, are difficult to observe in household budget surveys.

The composition of households is associated with differences in expenditures. As expected, the number of children (18 and below) reduces household expenditures on alcohol. Holding composition constant, households with younger heads have relatively higher expenditures on alcohol, a result which suggests targeting messages designed to reduce alcohol purchases to younger adults. Finally, the evidence of differences among regions holds some promise for states interested in taxation opportunities (or other policies designed to reduce alcohol consumption). Holding other factors constant, residents in the Northeast and West urban areas are more likely to purchase alcohol and spend relatively more. These results suggest greater opportunities to reduce alcohol consumption in these areas by increased taxation, for example, relative to other regions. Finally, while the current study does not address the effects of prices and institutional factors like alcohol availability (e.g., across states), collection of such information would allow the investigation of a wider range of issues and might deserve more serious attention in future surveys.

In sum, household characteristics play a relatively important role in explaining expenditures on alcohol. Additional data on prices, or at least identification of state of residence, would improve the usefulness of the data for evaluating strategies for tax effectiveness. Differences between the effects on probability of purchase and level of purchase for some factors suggest the importance of targeting consumer education selectively.

TABLE 1.
Variable Definitions

Variable	Definition
Alcohol (dependent)	Expenditures on alcohol ($ / 2 weeks)
Income	Household income ($000 / 2 weeks)
1Age	Age of household head
Household composition:	
Members # 18	Number of members aged 18 or under
Members 19-64	Number of members aged 19-64
Members $ 65	Number of members aged 65 or over
Dummy variables (yes = 1, no = 0)	
Rural	Resides in rural area (reference group)
Urban households:	
Northeast	Resides in the North and Northeast
Midwest	Resides in the Midwest
South	Resides in the South
West	Resides in the West
High school	Household head high school educated (or some high school)
College	Household head college educated
Married	Household head is married
Homeowner	Household is homeowner
Male	Household head is male
White	Household head is white
Seasons:	
Spring	Survey occurred in spring
Summer	Survey occurred in summer
Fall	Survey occurred in fall
Winter	Survey occurred in winter (reference)

TABLE 2.
Sample Statistics: Household Alcohol Consumption in the U.S.

	Full sample (n = 9552)		Consuming (n = 4411)		Non-consuming (n = 5141)	
Variable	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
Alcohol	11.662	25.819	25.254	33.172	!	!
Income	1.094	0.944	1.312	1.024	0.906	0.823
Age	46.608	17.568	42.543	15.183	50.095	18.692
Members # 18	0.729	1.128	0.727	1.075	0.731	1.172
Members 19-64	1.572	0.992	1.736	0.897	1.432	1.046
Members $ 65	0.297	0.599	0.183	0.502	0.395	0.656
Rural	0.112		0.088		0.132	
Northeast	0.185		0.190		0.180	
Midwest	0.229		0.231		0.227	
South	0.262		0.257		0.266	
West	0.213		0.234		0.195	
High school	0.426		0.386		0.460	
College	0.477		0.560		0.406	
Married	0.584		0.622		0.551	
Homeowner	0.615		0.628		0.603	
Male	0.653		0.715		0.600	
White	0.875		0.894		0.858	
Spring	0.252		0.243		0.261	
Summer	0.235		0.236		0.235	
Fall	0.276		0.270		0.280	
Winter	0.328		0.339		0.319	

SOURCE: Compiled from BLS' 1989 and 1990 Consumer Expenditure Diary Surveys.

TABLE 3.
ML Estimation of IHS double-hurdle and IHS Tobit Models

	IHS double-hurdle			IHS Tobit	
Variable	Particip-ation	Con-sumption	Heterosce-dasticity	Con-sumption	Heterosce-dasticity
Constant	0.569	!1.916	2.770*	!6.289*	2.888*
	(0.513)	(2.410)	(0.048)	(1.934)	(0.035)
Members # 18	!0.075	!1.970*	!0.027*	!2.208*	!0.031*
	(0.082)	(0.304)	(0.011)	(0.279)	(0.011)
Members 19-64	0.454*	!0.623	0.102*	0.817*	0.083*
	(0.143)	(0.540)	(0.017)	(0.461)	(0.015)
Members $ 65	0.116	!0.989	0.103*	!0.571	0.099*
	(0.267)	(1.130)	(0.029)	(0.913)	(0.027)
Income	1.146*	3.364*		4.405*	
	(0.208)	(0.411)		(0.316)	
Age	!0.012*	!0.207*		!0.290*	
	(0.006)	(0.031)		(0.023)	
Northeast	!0.432*	6.523*		5.192*	
	(0.302)	(1.084)		(0.956)	
Midwest	!0.181	4.168*		3.651*	
	(0.304)	(1.032)		(0.917)	
South	!0.409	3.834*		3.014*	
	(0.298)	(1.029)		(0.907)	
West	!0.317	5.802*		5.193*	
	(0.297)	(1.059)		(0.937)	
High school	!0.016	4.033*		4.339*	
	(0.209)	(1.307)		(1.024)	
College	0.178	5.443*		6.549*	
	(0.224)	(1.331)		(1.055)	
Married	0.401	!4.306*		!3.151*	
	(0.361)	(0.855)		(0.734)	

Homeowner	0.161	0.603	1.394*	
	(0.149)	(0.714)	(0.586)	
Male	0.566*	3.187*	5.601*	
	(0.197)	(0.738)	(0.604)	
White	!0.416	4.551*	3.436*	
	(0.262)	(0.975)	(0.771)	
Spring		!1.960*	!1.853*	
		(0.623)	(0.633)	
Summer		!0.832	!0.678	
		(0.612)	(0.623)	
Fall		!1.521*	!1.497*	
		(0.609)	(0.618)	
θ	0.053*		0.048*	
	(0.003)		(0.002)	
Log-likelihood		!24422.435		!24483.061

Note: Asymptotic standard errors in parentheses. Asterisks indicate significance at the 0.10 level.

TABLE 4.
Elasticities With Respect to Continuous Variables

Variable	IHS double-hurdle			IHS Tobit		
	Prob-ability	Cond. level	Uncond. level	Prob-ability	Cond. level	Uncond. level
Members # 18	!0.060*	!0.084*	!0.144*	!0.070*	!0.089*	!0.159*
	(0.008)	(0.012)	(0.014)	(0.008)	(0.012)	(0.014)
Members 19-64	!0.012	0.236*	0.224*	0.066*	0.265*	0.331*
	(0.032)	(0.042)	(0.057)	(0.030)	(0.039)	(0.055)
Members $ 65	!0.011	0.041*	0.030	!0.004	0.045*	0.041*
	(0.012)	(0.014)	(0.021)	(0.011)	(0.012)	(0.018)
Income	0.208*	0.132*	0.340*	0.203*	0.151*	0.353*
	(0.034)	(0.014)	(0.035)	(0.014)	(0.011)	(0.024)
Age	!0.417*	!0.345*	!0.762*	!0.568*	!0.423*	!0.991*
	(0.049)	(0.049)	(0.096)	(0.044)	(0.033)	(0.077)

Note: Asymptotic standard errors in parentheses. Asterisks indicate significance at the 0.10 level.

TABLE 5.
Effects of Binary Variables

	IHS double-hurdle			IHS Tobit		
Variable	Prob-ability	Cond. level	Uncond. level	Prob-ability	Cond. level	Uncond. level
Northeast	0.126	$4.88	$5.06	0.099	$3.34	$3.54
Midwest	0.084	2.93	3.07	0.069	2.27	2.37
South	0.072	2.67	2.67	0.057	1.84	1.92
West	0.115	4.23	4.46	0.099	3.35	3.54
High school	0.082	2.83	2.98	0.082	2.59	2.67
College	0.116	3.97	4.33	0.124	4.11	4.31
Married	!0.078	!3.47	!3.52	!0.061	!2.16	!2.32
Homeowner	0.016	0.47	0.60	0.027	0.94	1.00
Male	0.082	2.45	2.99	0.107	3.65	3.87
White	0.087	3.30	3.39	0.066	2.19	2.31
Spring	!0.040	!1.56	!1.69	!0.036	!1.26	!1.35
Summer	!0.017	!0.68	!0.74	!0.013	!0.47	!0.51
Fall	!0.031	!1.22	!1.33	!0.029	!1.03	!1.10

Endnotes

i.For the rest of the paper, "consumption" refers to the "expenditure" made on the commodity in question. We expect the expenditure decision to coincide with the consumption decision. We do not address explicitly the "durable" component of alcohol expenditures.

ii.For the IHS Tobit model because the conditional density of E is skewed it is no longer tied as closely to the probability of a non-zero outcome as in the standard Tobit model.

iii.Expenditures consist of the transaction costs, including excise and sales taxes, of the expenditures on alcohol acquired during the interview period for at-home or away-from-home use. They also include expenditures for gifts, but exclude purchases or portions of purchases directly assignable to business purposes.

iv.As one reviewer pointed out, state excise tax rates are often used as proxies for prices in alcohol demand studies. In addition, state laws that control the sale of alcohol are also potential explanatory variables. Unfortunately, households in the Diary Surveys are not identifiable by states.

v.The idea to accommodate infrequency of purchases in survey data is appealing, but the infrequency-of-purchase model is also used at a cost, by ruling out other plausible causes of zero such as abstention ! the major motivation behind all double-hurdle models. In addition, for computational tractability the infrequency-of-purchase model is specified with very restrictive assumptions regarding the specification of the purchase probability and the consumption-purchase relationship (Pudney 1989, pp. 179-180). Our test result may reflect the rejection of these assumptions, not the idea of infrequency of purchases.

vi.Results are not strictly comparable because the analysis by Blaylock and Blisard (1993) was based on individual intake data, and for women only.

REFERENCES

Amemiya, T. and J.L. Powell (1981), "A Comparison of the Box-Cox Maximum Likelihood Estimator and the Non-Linear Two Stage Least Squares Estimator," *Journal of Econometrics*, 17 (December): 351-381.

Arabmazar, A. and P. Schmidt (1981), "Further Evidence on the Robustness of the Tobit Estimator to Heteroskedasticity," *Journal of Econometrics*, 17 (November): 253-258.

Arabmazar, A. and P. Schmidt (1982), "An Investigation of the Robustness of the Tobit Estimator to Non-normality," *Econometrica*, 50 (July): 1055-1063.

Atkinson, A.B., J. Gomulka, and N.H. Stern (1990), "Spending on Alcohol: Evidence From the Family Expenditure Survey 1970-1983," *Economic Journal*, 100 (September): 808-827.

Blaylock, J.R. and W.N. Blisard (1993), "Women and the Demand for Alcohol: Estimating Participation and Consumption," *Journal of Consumer Affairs*, 27 (Winter), 319-334.

Blundell, R.W. and C. Meghir (1987), "Bivariate Alternatives to the Univariate Tobit Model," *Journal of Econometrics* 33 (January/February): 179-200.

Burbidge, J.B., L. Magee, and A.L. Robb (1988), "Alternative Transformations to Handle Extreme Values of the Dependent Variable," *Journal of the American Statistical Association*, 83 (March): 123-127.

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

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

Haines, P., D. Guilkey, and B. Popkin (1988), "Modeling Food Consumption Decisions as A Two-Step Process," *American Journal of Agricultural Economics*, 70 (August): 543-552.

Heien, D. and G. Pompelli (1989), "The Demand for Alcoholic Beverages: Economic and Demographic Effects," *Southern Economic Journal*, 55 (January): 759-770.

Hilton, M.E. (ed.) (1993), *Economics and the Prevention of Alcohol-Related Problems*, Research Monograph No. 25, National Institute on Alcohol Abuse and Alcoholism, Rockville, Maryland, 1993.

Horowitz, J.L., and G.R. Neumann (1989), "Specification Testing in Censored Regression Models: Parametric and Semiparametric Methods," *Journal of Applied Econometrics* 4: S61-S86.

Johnson, N.L., and S. Kotz (1970). *Distributions in Statistics: Continuous Univariate Distributions - 1*, New York: Wiley.

Jones, A.M. and J. Possnet (1991), "Charitable Donations by UK Households: Evidence From the Family Expenditure Survey," *Applied Economics*, 23 (February): 343-351.

Leung, S.F. and C.E. Phelps, "'My Kingdom for a Drink...?': A Review of the Estimates of the Price Sensitivity of Demand for Alcoholic Beverages," in M.E. Hilton, ed., *Economics and the Prevention of Alcohol-Related Problems*, Research Monograph No. 25, National Institute on Alcohol Abuse and Alcoholism, Rockville, Maryland, 1993.

MacKinnon, J.G. and L. Magee (1990), "Transforming the Dependent Variable in Regression Models," *International Economic Review*, 31 (May): 315-339.

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

McDonald, J.F. and R.A. Moffitt (1980), "The Uses of Tobit Analysis," *Review of Economics and Statistics*, 62 (May): 318-321.

Ornstein, S.I. (1980), "Control of Alcohol Consumption Through Price Increases," *Journal of Studies on Alcohol*, 41 (9): 807-818.

Ornstein, S.I. and D. Levy (1983), "Price and Income Elasticities and the Demand for Alcoholic Beverages," in M. Galanter ed., *Recent Development in Alcoholism*, Vol. I, New York: Plenum.

Pompelli, G. and D. Heien (1991), "Discrete/Continuous Consumer Demand Choices: An Application to the U.S. Domestic and Imported White Wine Markets," *European Review of Agricultural Economics*, 18 (1): 117-130.

Pudney, S.E. (1989), *Modelling Individual Choice: The Econometrics of Corners, Kinks, and Holes*. Oxford: Basil Blackwell.

Reynolds, A. (1990), "Analyzing Fresh Vegetable Consumption From Household Survey Data," *Southern Journal of Agricultural Economics*, 22 (December): 31-38

Reynolds, A. and J.S. Shonkwiler (1991), "Testing and Correcting for Distributional Misspecifications in the Tobit Model: An Application of the Information Matrix Test," *Empirical Economics*, 16 (3): 313-323.

Selvanathan, E.A. (1991), "Cross-Country Alcohol Consumption Comparison: An Application of the Rotterdam Demand System," *Applied Economics*, 23 (October): 1613-1622.

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

U.S. Department of Commerce, Bureau of Labor Statistics (1989), *Consumer Expenditure Survey: Diary Survey, Documentation of Public-Use Tape*, Washington, D.C.

U.S. Department of Commerce, Bureau of Labor Statistics (1990), *Consumer Expenditure Survey: Diary Survey, Documentation of Public-Use Tape*, Washington, D.C.

Vuong, Q.H. (1989), "Likelihood Ratio Tests for Model Selection and Nonnested Hypotheses," *Econometrica*, 57 (March): 307-333.

Wang, Q. and H.H. Jensen (1994), "Do Consumers Respond to Health Information in Food Choices? Models and Evaluation of Egg Consumption," In T.A. Mauldin ed., *Consumer Interests Annual: 40th Annual Conference of the American Council on Consumer Interests, March 23-26, Minneapolis, Minnesota*, pp. 57-64.

Yen, S.T. (1994), "Cross-Section Estimation of U.S. Demand for Alcoholic Beverage," *Applied Economics*, 26 (April): 381-392.