

RURAL ECONOMY

**Cross Product Censoring in a Demand System with Limited
Dependent Variables: A Multivariate Probit Model Approach**

Kevin Chen and Chen Chen

Staff Paper 00-02

Staff Paper



Department of Rural Economy
Faculty of Agriculture, Forestry
and Home Economics
University of Alberta
Edmonton, Canada

Cross Product Censoring in a Demand System with Limited Dependent Variables: A Multivariate Probit Model Approach

Kevin Chen and Chen Chen

Staff Paper 00-02

The authors are, respectively, Assistant Professor and Research Associate, Department of Rural Economy, University of Alberta, Edmonton, Alberta.

The authors would like to thank Peter Boxall for several enlightening discussions during the course of preparing this manuscript.

Copyright 2000 by Kevin Chen and Chen Chen. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

The purpose of the Rural Economy 'Staff Papers' series is to provide a forum to accelerate the presentation of issues, concepts, ideas and research results within the academic and professional community. Staff Papers are published without peer review.

Cross Product Censoring in a Demand System with Limited Dependent Variables: A Multivariate Probit Model Approach

Introduction

The most challenging problem in cross sectional demand analyses is to deal with zero expenditure because OLS estimates tend to be biased towards zero in a regression model where a large proportion of the dependent variable is zero (Deaton 1986, Greene 1993). The problem arises as households participating in the survey do not report or consume all types of food products during the survey period. While several econometric approaches have been proposed to deal with zero expenditure problems, the most common strategy adopted in the food demand literature is to employ an extension of Tobin's (1958) limited dependent variable model for single equations, later generalized by Amemiya (1974) for systems of equations. In empirical applications, the Heckman (1979) two-step type estimation procedure for a demand system with limited dependent variables proposed by Heien and Wessells (1990) has become increasingly popular in applied demand analysis (Abdelmagid, Wohlgenant, and Safley 1996, Alderman and Sahn 1993, Gao and Spreen 1994, Gao, Wailes, and Cramer 1997, Han and Wahl 1998, Nayaga 1995, 1996, 1998, Salvanes and DeVoretz 1997, Wang et al 1996). The practice, however, has recently been questioned due to an apparent internal inconsistency problem associated with the Heien and Wessells procedure. Shonkwiler and Yen (1999) and Su and Yen (2000) therefore proposed a consistent two-step estimation of a censored demand system, which is also adopted here.

A remaining problem with the Heien and Wessells procedure is that consumers' market participation in each product is model as an independent process and estimated by

the univariate probit model. Though attractive because of the ease with which the model can be estimated, correction factors obtained from univariate probit equations do not capture cross-product censoring impacts in multiple equations. In the same spirit as the seemingly unrelated regression model, this could result in inefficient probit estimates when cross product censoring occurs. These inefficient probit estimates likely affect the estimation of the second stage censored demand system. The greater the cross product censoring (the correlation of disturbances), the greater the efficiency gain one would gain if using a multivariate probit model. However, in practice this has rarely been evaluated because the estimation of the multivariate probit model involves numerical integration of a multiple dimensional multivariate normal density function. This, combined with the necessity of using an iterative technique to maximize the likelihood function, has made the application of the multivariate probit model computationally difficult. With the development of various simulation techniques, estimation of the multivariate probit has recently become more feasible.

This paper applies a simulated maximum likelihood estimation (MLE), known as the GHK sampling method (Hajivassiliou 1993), to estimate the multivariate probit model of Alberta consumer participation in 1% milk, 2% milk, whole milk, and skim milk markets. The data is extracted from the 1996 Canadian food expenditure survey (FFES). About a third of Alberta households consumed only one type of milk, necessitating procedure to account for censored expenditure distributions. The fluid milk is chosen due to three further considerations. First cross censoring among various types of fluid milk is likely high. For example, Cornick, Cox, and Gould (1994) reported interdependence of skim, reduced-fat milk, and whole milk purchases in the context of

multivariate Tobit model. This could help highlight the importance of accounting for cross product censoring. Second, fluid milk is a frequently purchased food item. With two-week data, there are less problems of zero expenditures generated by infrequency of purchase. This makes a choice of the Heckman two-step type estimation appropriate for the data. Third, with increased health concerns about dietary fat intake, determining the underlying causes for changes in fluid milk consumption patterns are of interest. In particular, this study represents the first attempt to model 1% and 2% milk separately. It is interesting to see if the demand for these two types of milk are governed by similar factors. In previous studies of milk purchasing, 1% and 2% milks were counted as one category - reduced-fat milk.¹

In the remaining sections of the paper, a likelihood ratio test is used to investigate the significance of cross product censoring in fluid milk purchases. A censored almost ideal demand system is then specified and estimated to illustrate the potential impact of ignoring cross product censoring on the estimation of the censored demand system in the second stage. The Likelihood Dominance Criterion (LDC) developed by Pollak and Wales (1991) is used to rank the multivariate and univariate probit-based demand systems.

Two-Step Estimation of A Censored Demand System

Consider the following Amemiya's (1974) type of censored demand system

¹ Recently, Food and Drug Administration (FDA) in the United States has proposed a new guideline for labeling fluid milk. For example, 2% milk should be labeled as a reduced-fat milk while 1% milk can be labeled as a low-fat milk.

$$\begin{aligned}
y_{it}^* &= f(x_{it}, \beta_i) + \varepsilon_{it} \\
z_{it}^* &= w_{it}' \alpha_i + \mu_{it} \\
y_{it} &= z_{it} y_{it}^*, z_{it} = 1 \text{ if } z_{it}^* > 0 \text{ and } z_{it} = 0 \text{ if } z_{it}^* \leq 0
\end{aligned} \tag{1}$$

where y_{it} and z_{it} are the observed dependent variables for product i and household t , y_{it}^* and z_{it}^* are the unobserved (latent) variables, x_{it} and w_{it} are vectors of exogenous variables, α_i and β_i are conformable vectors of parameters, and ε_{it} and μ_{it} are random errors. These random errors are distributed as a multivariate normal with probability density function, $\phi(\varepsilon, \mu, 0, \Sigma)$, where Σ is a covariance matrix for $[\varepsilon_{it}, \mu_{it}]$. Direct MLE estimation of equation (1) is computationally difficult because of the need to evaluate multiple integrals in the likelihood function. In practice, the Heckman (1979) two-step type estimation procedure is applied to the above system to alleviate some of these computational difficulties. For example, Heien and Wessells (1990) propose that each equation in the system is augmented by a selectivity regressor derived from the univariate probit estimates in an earlier step, and the system of equations is estimated with seemingly unrelated regression in the second step. As the convenience of dropping zero observations is not possible when applying Heckman (1979) two step procedure to a demand system, Heien and Wessells (1990) redefine the model using all observations in the second stage. However, they fail to account for the unconditional mean as suggested by Maddala (1980, p. 222). Assuming for each i the error terms $[\varepsilon_{it}, \mu_{it}]$ are distributed as bivariate normal with $\text{cov}(\varepsilon_{it}, \mu_{it}) = \lambda_i$, Shonkwiler and Yen (1999) correct this inconsistency by defining the second stage regression for product i as

$$y_{it} = \Phi(w_{it}' \alpha_i) f(x_{it}, \beta_i) + \lambda_i \phi(w_{it}' \alpha_i) + \xi_{it} \tag{2}$$

where $\Phi(w'_i\alpha)$ is cumulative multivariate normal probability evaluated at $w'_i\alpha$ such that

$$P(y_{it}) = \int_{-\infty}^{\infty} d\varepsilon_{it} \int_{-\infty}^{y_i} d\varepsilon_{1t} \dots \int_{-\infty}^{y_i} d\varepsilon_{mt} \phi(\varepsilon_{1t}, \dots, \varepsilon_{mt}; 0, \Sigma), \quad (3)$$

$\phi(w'_i\alpha)$ is the multivariate normal probability density evaluated at $w'_i\alpha$, and ξ_{it} is a error term.

To calculate $\Phi(w'_i\alpha)$ and $\phi(w'_i\alpha)$, MLE of $\hat{\alpha}_i$ should be, in theory, obtained by multivariate probit model. A common practice, however, is to obtain univariate probit estimates using the binary outcome $z_{it} = 1$ and $z_{it} = 0$ for each i . This implies that $E(\varepsilon_{it}\varepsilon_{kt}) = 0$. This practice produces consistent but asymptotically inefficient parameter estimates for the probit model. The degree of the inefficiency, however, depends on the degree of the correlation among the ε_{it} 's. To evaluate the degree of such correlation, one must estimate the multivariate probit model.

Data

The two-step estimation of the above censored demand system is applied to consumer demand for four types of fluid milk in Alberta, Canada, including 1% milk, 2% milk, whole milk, and skim milk. The data is extracted from 1996 Canadian food expenditure survey (FFES) which covers 846 Alberta households in 1996. For each household, the survey was completed in two consecutive survey weeks. We aggregated the weekly data before estimation, as two-week data should exhibit less problems of zero expenditure generated by infrequency of purchase. Close to 50% of Alberta households purchased 2% milk, 30% purchased 1% milk, 20% purchased skim milk, and 15% purchased whole milk during the two-week survey period.

The FFES data also provide detailed socioeconomic and demographic information on the sampled households. The selection of socioeconomic and demographic variables in the models that are applied here is guided by previous studies on the demand for fluid milk (for example, Huang and Rauniker 1983, Reynolds 1991, Cornick, Cox, and Gould 1994, Gould 1996) as well as the availability of the data. In addition to the major demographic and socioeconomic variables such as age, ethnicity, gender, urban/rural area, household size, income, education levels, we also consider the effect on family milk expenditures of the employment status of female household members and the changing composition of families. For example, different household characteristics are included in the demand models postulated here to capture the influence of changing family structure in Canada. The various household characteristics in Alberta are summarized in Table 1.

Simulated MLE of Multivariate Probit Model

Estimating the multivariate probit model is challenging because the need to evaluate multi-dimensional integrals of probability density functions. Several simulation methods have been recently advanced in the literature to overcome this computational difficulty, including the frequency method (Lerman and Manski 1980), the important sampling method (McFadden 1989), Stern's method (Stern 1992), and the smooth recursive conditioning simulator or the Geweke-Hajivassiliou-Keane (GHK) simulator. We propose to use the GHK simulator in this paper, since this has several advantages as noted in Chen and Cosslett (1998). First, the resulting simulated probabilities are continuous in parameter space, and therefore estimation can be performed by using standard optimization packages. Secondly, Borsch-Supan and Hajivassiliou (1993) show

that the GHK simulator is unbiased for any given number of replications R , and that it generates substantially smaller variances than the frequency simulator and Stern's simulator. Based on the root-mean-square error criterion, Hajivassiliou, McFadden, and Ruud (1996) show that the GHK simulator is unambiguously the most reliable method for simulating normal probabilities, compared to twelve other simulators. Furthermore, to estimate the parameters by the simulated maximum likelihood estimation method, we need only replace the choice probabilities in the likelihood function by the simulated probabilities using the GHK simulator. Details of the computational steps required to simulate the probabilities can be found in Borsch-Supan and Hajivassiliou (1993) and Hajivassiliou (1999).

The simulated MLE estimates of multivariate probit model are reported in Table 1. The statistical significance of the model is examined using a likelihood ratio test of the null hypothesis that all slope coefficients are zero. The statistic of Chi-squared indicates rejection of this hypothesis. As our primary interest is with respect to interdependence of milk purchases, the estimated correlation coefficients are provided in Table 2. All estimated correlation coefficients, except that between whole and skim milk, are statistically significant. In order to formalize these results, a test was performed on the hypothesis that all estimated correlation coefficients are zero. This was accomplished using a likelihood ratio test. The value of the likelihood function for the multivariate probit model was compared with the value of the likelihood function for the univariate probit models. The test statistic is

$$-2 \left[\log(L_{\text{multivariate probit}}) - \log(L_{\text{univariate probit}}) \right] \quad (3)$$

The test statistic is distributed as χ^2 with 6 degrees of freedom. The log likelihood value of the multivariate probit model is $-1,647$, as compared with $-1,710$ for the univariate probit model. The resulting LR-statistic is 126, which is larger than the critical value of 14.45 at 5% significance level, indicated rejection of the null hypothesis that $E(\varepsilon_{it}\varepsilon_{kt})=0$. In other words, failure to account for the correlation of the disturbances (cross product censoring) could cause substantial efficiency loss. We now turn to examine how inefficient probit estimates in the first stage might influence the second stage estimation of the censored demand system.

A number of socioeconomic and demographic factors were found to significantly influence the probability of milk purchases. Immigrants from West Europe are more likely to purchase 2% milk than Canadian born households, Asian and other immigrants are more likely to purchase whole milk than Canadian born households, and Asian immigrants are less likely to purchase 1% milk. The frequency of purchasing 2% milk increases with age and decreases if the household head is male. Working women are more likely to purchase 1% milk. The likelihood to purchase skim milk increases if the household reside in urban area. As expected, the households with children are more likely to purchase 2% and whole milk and less likely to purchase 1% milk. University educated households have higher purchase probabilities of purchasing skim milk. Income has positive effect on the probabilities of purchasing skim milk but no influence on other types of milk. It is interesting to note that different factors affect the probability of 1% and 2% milk purchases. Together with founded negative correlation between 1% and 2% milk purchases, this indicates that previous practice to have aggregate reduced-fat milk (including both 1% and 2%) may be not appropriate.

A Censored Almost Ideal Demand System

We approximate $f(\beta, x)$ by Deaton and Muelbauer's (1980) almost ideal demand system augmented with socioeconomic and demographic effects:

$$y_{it} = f(x_{it}, \beta_i) = l_0 + \sum_{k=1}^K l_k \Psi_k + \sum_{j=1}^n \gamma_{ij} \ln p_j + \beta_i \ln \left(\frac{M}{P^*} \right) \quad (4)$$

where y_i is now interpreted as the budget share of i th commodity, Ψ_k is the k th socioeconomic and demographic variable, P_i is the price for i th commodity, M is the total expenditure of the household, and P^* is a price index defined by

$$\log P^* = \alpha_0 + \sum_k \alpha_k \log P_k + \frac{1}{2} \sum_j \sum_k \gamma_{k,j} \log P_k \log P_j \quad (5)$$

Adding up restrictions require that $\sum \alpha_i = 1$, $\sum \gamma_{ij} = 0$, and $\sum \beta_i = 0$, homogeneity requires $\sum \gamma_{ij} = 0$, and symmetry requires that $\gamma_{ij} = \gamma_{ji}$, $\forall i$ and j . In many practical situations, where prices are relatively collinear, P_i is approximated proportional to any appropriately defined price index, for example, the Stone index, by $\sum w_k \log p_k$. Such an index can be calculated before estimation so that equation (5) becomes straightforward to estimate. Replacing the price index with Stone's price index gives the linear approximation of the model (LAIDS), which is extensively used in demand analysis. Combining (2) and (5) yields the following censored almost ideal demand system:

$$y_{it} = \Phi(w_{it}, \alpha_i) \left[l_0 + \sum_{k=1}^K l_k \Psi_k + \sum_{j=1}^n \gamma_{ij} \ln p_j + \beta_i \ln \left(\frac{M}{P^*} \right) \right] + \lambda_i \phi(w_{it}, \alpha_i) + \xi_{it} \quad (6)$$

The linear version of equation (6) can be estimated using seemingly unrelated regression (SUR) or three stage least square (3SLS) with the usual demand restrictions imposed. 3SLS is adopted here to account for potential simultaneous bias (Alston et al

1994). To estimate demand share equations (6), one share equation needs to be dropped to avoid a singular variance-covariance matrix. Unfortunately, the presence of $\Phi(w_{it}'\alpha)$ and $\phi(w_{it}'\alpha)$ in the right hand side of equation (6) does not permit invariance to equation dropped because each equation will not now have identical regressors. One can avoid this problem by estimating some *ad-hoc* quantity-dependent demand systems rather than share-dependent systems such as AIDS (Shonkwiler and Yen 1999, Su and Yen 2000). Because which equation to drop didn't affect our primary conclusions reach in this paper, only the estimated coefficients with skim milk share equation dropped were provided in Table 4.

The adjusted R-square ranged from 0.04 to 0.12, which are not unusual for cross sectional data. All coefficients on the probability density are positively significant in all equations, suggesting that correcting selection bias is important. The estimated own price and expenditure coefficients from the LAIDS demand system for the fluid milk are statistically significant. Most of the estimated cross price coefficients are also statistically significant. It is interesting to observe that only a few coefficients of socioeconomic and demographic variables are statistically significant. One person family in Alberta appears to demand for 1% milk. Immigrants from Southern Europe, Asia, and Other drink more whole milk than Canadian born households. The households with children demand for whole milk than no-children household, while the demand for whole milk decreases with the increase in the size of households.

For comparison, the estimated coefficients based on univariate probit models are provided in Table 5. Though the sign and significance level of all estimated coefficients between multivariate and univariate probit models are similar, magnitudes and standard

errors of the coefficients are different. As expected, standard errors of the coefficients from multivariate probit based demand system are smaller than those from univariate probit based demand system. The differences between coefficient estimates for multivariate and univariate probit model-based demand system are indicative of the average bias that would occur without accounting for cross product censoring. Table 6 provides pair-wise t-test results for the differences of price and expenditure coefficients between multivariate and univariate probit model based demand systems. All the differences except one are statistically significant at 5%, indicating the importance of accounting for cross product censoring in the first stage estimation.

To select the best model, we need to apply a non-nested test as the univariate and multivariate demand systems are not nested with each other. Since we cannot estimate the composite to use the standard likelihood ratio test procedure to compare the two hypotheses with the composite, we used the Likelihood Dominance Criterion (LDC) introduced by Pollak and Wales (1991). This method provides a simple way to compare two non-nested specifications relying on the value of estimated likelihood functions. Unlike the non-nested J test and the Cox test, the LDC can be used to rank a pair of competing models *without* actually estimating the composite model. In a recent paper, Saha, Shumway, and Talpaz (1994), using a Monte Carlo approach, found that the LDC outperformed some widely used non-nested procedures in selecting the true model. Because two hypotheses contain the same number of parameters, the LDC always prefers the one with the higher likelihood (Pollak and Wales 1991, p 229). Since multivariate probit-based demand system has the higher log likelihood (-604.68) than univariate probit

based-demand system (-606.31), it is concluded that the multivariate probit-based demand system is preferred to the univariate probit-based demand system.

Conclusion

This paper estimates a censored demand system for fluid milk, including 1% milk, 2% milk, whole milk, and skim milk. Departure from previous studies on the two-step estimation of the censored demand system, we accounted for potential cross product censoring of milk purchases. The GHK sampling method is applied to obtain simulated MLE estimates of multivariate probit model. A likelihood ratio test indicated the significance of cross product censoring in fluid milk purchases. The empirical implication is that failure to account for the cross product censoring could cause substantial efficiency loss. The inefficient probit estimates in the first stage might influence the second stage estimation of the censored demand system.

The second stage demand system is approximated by an almost ideal demand system. This paper represents a first application of Shonkwiler and Yen's (1999) and Su and Yen's (2000) consistent two-step estimation of a censored demand system to an almost ideal demand system. The comparison of the second stage almost ideal demand system estimates based on multivariate and univariate probit models revealed significant biases of ignoring cross product censoring in the first stage. A nonnest test, the Likelihood Dominance Criterion (LDC) proposed by Pollak and Wales (1991), indicates that the multivariate probit-based demand system is a better model. These results together indicate the importance of accounting for the correlation of the disturbances in the first stage when cross product censoring is suspected. Instead of continuing to apply

univariate probit models, the multivariate probit model should be applied in the first stage estimation of the censored demand systems. With the recent development in computational capacity, this is no longer unachievable. Of course, cautions are still needed as the estimation of the four-equation system with 846 observations like ours took few hours using Pentium III 550 machine to complete with modest replications ($R=300$).

References

- Abdelmagil, B. D., M. K. Wohlgenant, and C. D. Safley. (1996). "Demand for Plants Sold in North Carolina Garden Centers." *Agr. And Resour. Econ. Rev.* 25: 28-37.
- Alderman, H. and D. E. Sahn. (1993) "Substitution between Goods and Leisure in a Developing Country." *Amer. J. Agr. Econ.* 75: 875-83.
- Alston, J. M., K.A.Foster, and R.D. Green. (1994) "Estimating Elasticities with the Linear Approximate Almost Ideal Demand System: Some Monte Carlo Results." *Rev. Econ. Statist.* 76: 351-56.
- Amemiya, T. (1974) "Multivariate Regression and Simultaneous Equation Models when the Dependent Variables are Truncated Normal", *Econometrica*, 42: 999-1012.
- Borsch-Supan, A. and V. A. Hajivassiliou. (1993) "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models", *Journal of Econometrics*, 58: 347-68.
- Chen, H. Z. and R. C. Cosslett. "Environmental Quality Preference and Benefit Estimation in Multinomial Probit Models: A Simulated Approach", *Amer. J. Agr. Econ.*, 80(1998); 512-520.
- Cornick, J., T. L. Cox, and B. W. Gould. (1994) "Fluid Milk Purchases: A Multivariate Tobit Analysis." *Amer. J. Agr. Econ.* 76: 74-82.
- Chiang, J. and L.F. Lee. (1992) "Discrete/continuous models of consumer demand with binding nonnegative constraints," *Journal of Econometrics* 54: 79-93.
- Deaton, A. (1986) "Demand Analysis", in *Handbook of Econometrics*, eds. Z. Griliches and M. D. Intriligator (North-Holland, Amsterdam), Ch. 30, Vol. III, pp. 1767-839.
- Deaton, A. and J. Muellbauer. (1980) *Economics and Consumer Behavior* New York; Cambridge University Press.
- Gao, X. M. and T. Spreen. (1994) "A Microeconomic Analysis of the U.S. Meat Demand." *Can. J. Agr. Econ.* 42: 397-412.
- Gao, X. M., E. J. Wailes, and G. L. Cramer. (1997) "A Microeconomic Analysis of Consumer Taste Determination and Taste Change for Beef." *Amer. J. Agr. Econ.* 79: 573-82.
- Greene, W.H. (1997) *Econometric Analysis* Third Edition. New Jersey: Prentice Hall Press.
- Hajivassiliou, W. A. "Some Practical Issues in Simulated Maximum Likelihood," in *Simulated-Based inference in Econometrics: Methods and Applications*, R. Mariaio, M. Weeks, and T. Schermann, eds, Cambridge University Press, 1999.
- Hajivassiliou, V. and D. McFadden. (1998) "The Method of Simulated Score for the Estimation of LDV Models," *Econometrica*, 66: 863-96.
- Hajivassiliou, V., McFadden, D., and Ruud, P. (1996) "Simulation of Multivariate Normal Rectangle Probabilities and their Derivatives: Theoretical and Computational Results", *Journal of Econometrics*, 72,: 85-134.
- Hajivassiliou, V. (1993) "Simulation Estimation Methods for Limited Dependent Variable Models," *Handbook of Statistics*, Vol, 11, *Econometrics*, G.S. Maddala, C. R. Rao, and .D. Vinod, eds., pp. 519-43. Amsterdam: North-Holland.
- Han, T. and T. Wahl. (1998) "China's Rural Household Demand for Fluid and Vegetables." *J. Agr. And Appl. Econ.* 30: 141-50.
- Heckman, J. (1979) "Sample Selection Bias as a Specification Error," *Econometrica* 47 (1) 153-161.

- Heien, D. and C. R. Wessells.(1990) "Demand Systems Estimation with Microdata: A Censored Regression Approach," *Journal of Business and Economic Statistics* 8: 365-371.
- Huang, C. L. and R. Rauniker. (1983) "Household Fluid Milk Expenditure Patterns in the South and United States." *S. J. Agr. Econ.* 15: 27-33.
- Lee, L. F. (1978), "Simultaneous Equations Models with Discrete and Censored Dependent Variables," in *Structural Analysis of Discrete Data with Econometric Applications*, eds., PL Manski and D. McFadden, Cambridge, MA: MIT Press, pp. 346-364.
- Lee, L. F. and Pitt, M. M. (1986) "Microeconomic Demand Systems with Binding Nonnegativity Constraints: the Dual Approach", *Econometrica*, 54: 1237-42.
- Lerman, S. R. and C. F. Manski. (1981) "On the Use of Simulated Frequencies to Approximate Choice Probabilities", *Structural Analysis of Discrete Choice Data with Econometric Applications*. C. F. Manski and D. McFadden, eds. Pp. 305-19. Cambridge MA: The MIT Press, 1981.
- Maddala, G. S. (1983) *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge, UK.
- McFadden, D. (1989) "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration", *Econometrica*, 58: 995-1026.
- Murphy, K. M. and Topel, R. H. (1983) "Estimation and Inference in Two-step Econometric Models", *Journal of Business and Econometric Statistics*, 3: 370-9.
- Nayga, R. M. (1995) "Microdata Expenditure Analysis of Disaggregate Meat Products." *Rev. Agr. Econ.* 17: 275-85.
- Nayga, R. M., Jr. (1996) "Wife's Labour Force Participation and Family Expenditures for Prepared Food, Food Prepared at Home, and Food Away from Home," *Agriculture and Resource Economic Review* 25(2): 179-86.
- Nayga, R. M. (1998) "Wife's Labor Force Participation and Family Expenditures for Prepared Food, Food Prepared at Home. And Food Away From Home." *Agr. And Resour. Econ. Rev.* 25: 179-86.
- Pollak, R.A. and T. J. Wales (1978) "Demographic Variables in Demand Analysis," *Econometrica* 49(6): 1533-1551.
- Pollak, R. and T. Wales. "The Likelihood Dominance Criterion." *J. Econometrics*. 47(1991): 227-42.
- Reynolds, A. (1991) *Modelling Consumer Choice of Fluid Milk*. Working Paper WP91/04, Department of Agricultural Economics and Business, University of Guelph.
- Saha, A., R. Shumway and H. Talpaz. *Performance of Likelihood Dominance and Other Nonnested Model Selection criteria: Some Monte Carlo Results*. Department of Agricultural Economics Working Paper. College Station: Texas A&M University, 1994.
- Salvanes, K. G. and D. J. DeVoretz. (1997) "Household Demand for Fish and Meat Products: Separability and Demographic Effects." *Marine Resour. Econ.* 12: 37-55.
- Shonkwiler, J.S. and Yen, S. T. (1999) "Two-step Estimation of a Censored System of Equations", *American Journal of Agricultural Economics*, 81: 972-82.
- Statistics Canada, Household Surveys Division, Family Expenditures Surveys Section (1996) *Consumer Expenditure Survey 1996*, Ottawa, Canada.

- Stern, S. (1992) "A Method for Smoothing Simulated Moments of Discrete Probabilities in Multinomial Probit Models," *Econometrica*, 60: 943-52.
- Su, S. J. and S. Yen (2000) "A Censored System of Cigarette and Alcohol Consumption", *Applied Economics* 32: 729-737.
- Tobin, J. (1958) "Estimation of Relationships for Limited Dependent Variables." *Econometrica*. 26: 24-36.
- Wang, J., X. M. Gao, E. J. Wailes, and G. L. Cramer. (1996) "U.S. Consumer Demand for Alcoholic Beverages: Cross-Section Estimation of Demographic and Economic Effects." *Rev. Agr. Econ.* 18: 477-88.

Table 1 Descriptive Statistics of Socioeconomic and Demographic Variables, 1996

Variable Definition and Code	Mean	St. Dev.
Birth Country of Household Managers		
Canada born*	0.7978	/
West Europe* (WEUROPE)	0.0839	/
Southeast Europe* (SEUROPE)	0.0355	/
Asian-pacific* (ASIA)	0.0626	/
Other nations*(OTHERN)	0.0201	/
Age of Household Head (AGE)	45.3676	15.4791
Male headed household (MALE)*	0.4775	/
Employment Status of Female Household Members		
Full-time employed* (FEWOMEN)	0.4267	/
Part time employed *(PEWOMEN)	0.1962	/
Household with only male members*(MONLY)	0.1382	/
Household Residing in Urban Area* (URBAN)	0.8782	/
Number of Children in Household (CHILDREN)	0.5366	/
Total Household Size (number of persons, HSIZE)	2.6536	1.3818
Social Assistance Recipient* (SAR)	0.0532	/
Household with University Educated Head* (UNIV)	0.4066	/
Household Income Before Tax (INCOME)	47,289.1	32,379.8
Survey Time (Quarterly Dummy Variables)		
Q1 (First quarter)*	0.2695	/
Q2 (Second quarter)*	0.2647	/
Q3 (Third quarter)*	0.2293	/
Q4 (Fourth quarter)*	0.2364	/
Family Composition		
FC1 (One person household)*	0.2210	/
FC2 (Married couple household, without children)*	0.2423	/
FC31 (Married couple household, with one child)*	0.1903	/
FC32 (Married couple household, with more than one child)*	0.1737	/
FC4 (HC2 with relative or non relative)*	0.0401	/
FC5 (Single parent household)*	0.0662	/
FC6 (Other household with relative)*	0.0283	/
FC7 (Other non married couple household)*	0.0378	/

Note: Variables with * are dummy variables which equals 1 for households belonging to this category and 0 otherwise.

Table 2 Simulated MLE of Multivariate Probit Model, 1996

	1% Milk	t-ratio	2% Milk	t-ratio	Whole Milk	t-ratio	Skim Milk	t-ratio
WEUROPE	-0.2015	-1.0630	0.3157*	1.7360	-0.2646	-0.9210	0.1825	0.9720
SEUROPE	-0.7202	-1.4510	0.2118	0.7750	0.4923	1.3130	-3.3510	-0.0020
ASIA	-0.5172*	-2.1830	0.0434	0.2100	0.7152*	3.0330	0.0308	0.1230
OTHERN	-0.6256	-1.1950	-0.5609	-1.3070	0.8019*	2.0340	0.3622	0.7050
AGE	0.0030	0.6690	0.0092*	2.3340	0.0087	1.5840	-0.0003	-0.0530
MALE	0.0752	0.6100	-0.2507*	-2.1560	0.1964	1.2750	0.0683	0.4840
FWOMEN	0.0073*	2.6530	-0.0003	-0.1010	-0.0005	-0.1400	-0.0060*	-1.8960
PWOMEN	0.0087*	2.3530	0.0014	0.3900	-0.0012	-0.2590	0.0023	0.5430
URBAN	-0.0954	-0.6250	-0.0960	-0.6590	-0.1425	-0.7800	1.1220*	3.5700
CHILDREN	-0.3438*	-2.1750	0.3028*	1.9650	0.3895**	1.7250	-0.1508	-0.7720
HSIZE	0.0599	0.6230	-0.0064	-0.0690	-0.0851	-0.7220	-0.0752	-0.6580
SASIFLAG	-0.1794	-0.6350	0.3259	1.4390	0.3413	1.3210	-0.0282	-0.0780
UNIV	-0.0003	-0.0020	-0.1186	-1.1580	-0.0479	-0.3210	0.3314*	2.7050
D1	0.1501	1.0430	-0.1279	-0.9450	-0.0520	-0.2720	0.1941	1.1470
D2	0.2963*	2.0290	0.0302	0.2180	-0.0056	-0.0290	-0.0354	-0.2040
D3	0.1236	0.8170	-0.0676	-0.4760	0.1344	0.6570	0.0637	0.3470
HC1	-0.4124	-1.5170	-0.4963**	-1.8270	-0.0405	-0.1050	-0.1692	-0.5460
HC2	-0.3606	-1.4420	0.2773	1.0980	0.0666	0.1860	-0.0632	-0.2190
HC31	-0.1483	-0.5740	0.2605	1.0380	0.2653	0.7840	0.0612	0.1950
HC32	0.3430	0.9990	0.0486	0.1460	0.0646	0.1550	0.0811	0.1970
HC4	-0.1310	-0.3700	0.5900	1.6220	0.0251	0.0460	0.1109	0.2510
HC6	-0.5017	-1.1230	0.6419	1.4230	0.0137	0.0270	-0.1101	-0.2470
HC7	-0.6021	-1.6050	-0.0430	-0.1190	0.0748	0.1370	0.0217	0.0520
INCOME	-0.0075	-0.3950	-0.0222	-1.1840	-0.0225	-0.8460	0.0643*	2.7240

*Statistically significant at either 5% critical level and ** at 10% critical level.

Table 3. Estimated Correlation Coefficients of the Disturbances in the Multivariate Probit Model

	Correlation Coefficients	T-ratios
R _{1% Milk, 2% Milk}	-0.4623*	-8.1300
R _{1% Milk, Whole Milk}	-0.3008*	-3.1870
R _{2% Milk, Whole Milk}	-0.1637*	-2.0870
R _{1% Milk, Skim Milk}	-0.2052*	-2.5810
R _{2% Milk, Skim Milk}	-0.2235*	-2.9050
R _{Whole Milk, Skim Milk}	-0.0875	-0.7990

*Statistically significant at either 5% critical level.

Table 4 Estimated Coefficients of the Second Stage Expenditure Share Equations with Symmetric Restrictions Imposed, Based on Multivariate Probit Model, 1996

	<i>1% Milk</i>	<i>t-ratios</i>	<i>2% Milk</i>	<i>t-ratios</i>	<i>Whole Milk</i>	<i>t-ratios</i>
WEUROPE	0.0319	0.4370	0.0427	0.7620	-0.0289	-0.5400
SEUROPE	0.0306	0.1370	0.0924	1.1460	0.0878**	1.7470
ASIA	-0.0631	-0.5070	0.0059	0.0900	0.1378*	3.4270
OTHERN	0.3163	1.3220	-0.2456	-1.3600	0.1797*	2.8210
AGE	-0.0011	-0.8310	0.0014	1.1630	0.0009	1.1640
MALE	0.0072	0.1720	-0.0383	-0.9390	0.0298	1.0870
FEWOMEN	0.0012	1.3070	-0.0003	-0.3020	0.0002	0.3360
PEWOMEN	0.0011	0.9540	-0.0010	-0.8260	-0.0003	-0.4240
URBAN	-0.0444	-0.9130	-0.0077	-0.1650	-0.0101	-0.3180
CHILDREN	0.0098	0.1660	0.0585	1.2230	0.0720*	2.2160
HSIZE	0.0146	0.4930	-0.0130	-0.4660	-0.0499*	-2.6380
ASSISTANCE	-0.0723	-0.5950	0.0448	0.6580	0.0509	1.1880
UNIVERSITY	-0.0355	-0.9530	-0.0568	-1.5960	0.0003	0.0110
Q1	0.0387	0.7080	-0.0453	-0.9760	0.0085	0.2670
Q2	0.0410	0.7540	-0.0254	-0.5690	-0.0404	-1.2960
Q3	0.0033	0.0570	-0.0189	-0.4020	0.0252	0.8120
FC1	0.2370*	2.5900	0.0919	1.0410	-0.0719	-1.2760
FC2	0.1007	1.3000	0.0604	0.8430	0.0100	0.1850
FC31	0.0125	0.1490	0.0132	0.1730	0.0346	0.6610
FC32	-0.0533	-0.4230	-0.0573	-0.5780	0.0111	0.1630
FC4	0.0292	0.2470	-0.0199	-0.1800	0.0326	0.3930
FC6	0.0107	0.0520	0.1079	1.0340	-0.0285	-0.3480
FC7	-0.0175	-0.1100	0.1500	1.4000	0.0719	0.8310
Log($P_{1\%milk}$)	-0.6587*	-6.8170	0.2487*	3.6080	0.0725	1.0170
Log($P_{2\%milk}$)	0.2487*	3.6080	-0.2159*	-2.7470	0.0541	0.9690
Log($P_{wholemilk}$)	0.0725	1.0170	0.0541	0.9690	-0.6546*	-5.6080
Log($P_{skimmilk}$)	0.3375*	4.4110	-0.0869	-1.3100	0.5280*	6.5390
Log(Expenditure)	0.0885*	4.8430	0.1118*	6.8030	0.0464*	4.0530
Probability Density	0.2123*	2.3190	0.5643*	4.3760	0.0795	0.9470
Adj. R-square	0.1166		0.0916		0.0352	

*statistically significant at the 5% critical level and ** at the 10% critical level.

Table 5 Estimated Coefficients of The Second Stage Expenditure Share Equations with Restrictions Imposed, Based on Univariate Probit Models, 1996

	<i>1% Milk</i>	<i>t-ratios</i>	<i>2% Milk</i>	<i>t-ratios</i>	<i>Whole Milk</i>	<i>t-ratios</i>
WEUROPE	0.0291	0.3820	0.0411	0.7190	-0.0300	-0.5760
SEUROPE	0.1599	0.4470	0.0926	1.1330	0.0906**	1.8080
ASIA	-0.0575	-0.4330	0.0070	0.1050	0.1378*	3.4540
OTHERN	0.2840	1.1950	-0.2496	-1.2750	0.1800*	2.8340
AGE	-0.0010	-0.7150	0.0013	1.0340	0.0009	1.1990
MALE	0.0052	0.1190	-0.0345	-0.8230	0.0312	1.1490
FEWOMEN	0.0011	1.1860	-0.0002	-0.2210	0.0002	0.3560
PEWOMEN	0.0010	0.7950	-0.0010	-0.8360	-0.0003	-0.4310
URBAN	-0.0469	-0.9290	-0.0016	-0.0350	-0.0072	-0.2300
CHILDREN	0.0318	0.5000	0.0549	1.1290	0.0706*	2.1840
HSIZE	0.0138	0.4450	-0.0144	-0.5090	-0.0494*	-2.6200
ASSISTANCE	-0.0704	-0.5450	0.0378	0.5490	0.0489	1.1510
UNIVERSITY	-0.0401	-1.0290	-0.0542	-1.4800	-0.0003	-0.0150
Q1	0.0334	0.5720	-0.0407	-0.8550	0.0075	0.2400
Q2	0.0366	0.6280	-0.0232	-0.5110	-0.0397	-1.2870
Q3	-0.0016	-0.0260	-0.0134	-0.2790	0.0254	0.8290
FC1	0.2490*	2.6480	0.1077	1.1940	-0.0720	-1.2830
FC2	0.1138	1.4000	0.0683	0.9300	0.0044	0.0820
FC31	0.0120	0.1360	0.0215	0.2770	0.0300	0.5790
FC32	-0.0863	-0.6390	-0.0488	-0.4860	0.0081	0.1210
FC4	0.0293	0.2400	-0.0222	-0.1980	0.0239	0.2940
FC6	0.0230	0.1020	0.1130	1.0760	-0.0306	-0.3830
FC7	-0.0058	-0.0320	0.1661	1.5170	0.0614	0.7530
Log(P _{1%<i>milk</i>})	-0.6794*	-6.7380	0.2360*	3.3060	0.0879	1.2400
Log(P _{2%<i>milk</i>})	0.2360*	3.3060	-0.2114*	-2.6350	0.0646	1.1690
Log(P _{whole<i>milk</i>})	0.0879	1.2400	0.0646	1.1690	-0.6681*	-5.8060
Log(P _{skim<i>milk</i>})	0.3555*	4.4220	-0.0892	-1.2950	0.5157*	6.4960
Log(Expenditure)	0.0919*	4.7810	0.1147*	6.8340	0.0474*	4.1830
Probability Density	0.2198*	2.4540	0.5541*	4.4520	0.0769	0.9060
Adj. R-square	0.11		0.09		0.04	

*statistically significant at the 5% critical level

**statistically significant at the 10% critical level.

Table 6 t-test Results for the Differences between Estimated Price and Expenditure Coefficients of the Second Stage Expenditure Share Equations, Based on Multivariate and Univariate Probit Models, respectively, 1996

	<i>1% Milk</i>	<i>t-ratio</i>	<i>2% Milk</i>	<i>t-ratio</i>	<i>Whole Milk</i>	<i>t-ratio</i>
Log(P _{1%<i>milk</i>})	0.0207*	4.2997	0.0127*	3.7118	-0.0154*	-4.4415
Log(P _{2%<i>milk</i>})	0.0127*	3.7118	-0.0045	-1.1621	-0.0105*	-3.8760
Log(P _{whole<i>milk</i>})	-0.0154*	-4.4415	-0.0105*	-3.8760	0.0135*	2.3885
Log(P _{skim<i>milk</i>})	-0.0180*	-4.7036	0.0023	0.6971	0.0123*	3.1507
Log(Expenditure)	-0.0034*	-3.7174	-0.0029*	-3.5712	-0.0010**	-1.7987

*statistically significant at the 5% critical level and ** at the 10% critical level.