European Review of Agricultural Economics Vol **29 (1)** (2002) pp. 85–103

Beverage consumption among US children and adolescents: full-information and quasi maximum-likelihood estimation of a censored system

Steven T. Yen

University of Tennessee, Knoxville, TN, USA

Biing-Hwan Lin

US Department of Agriculture, Washington, DC, USA

Received December 2000; final version accepted November 2001

Summary

Milk, soft drink and juice consumption is investigated for children and adolescents in the USA. Full-information maximum likelihood estimator (FIML) and a parsimonious quasi maximum-likelihood (QML) alternative are used to estimate a censored system of beverage equations. The QML estimator is found to be an acceptable alternative to the FIML estimator for the empirical application considered. We find displacement of milk by soft drinks as a child or adolescent grows older. Income, TV watching, gender, race, and other demographic variables also play significant roles in determining beverage consumption.

Keywords: censoring, microdata, milk, soft drinks

JEL classification: C34, D12, Q18

1. Introduction

Data from surveys conducted by the US Department of Agriculture (USDA) indicate major changes in beverage consumption among US children and adolescents during the past two decades. The popularity of soft drinks rose whereas milk consumption declined (Table 1). Of the female adolescents aged 12–17, for instance, the proportion consuming carbonated soft drinks on a given day increased from 47 per cent in 1977–1978 to 62 per cent in 1994–1996, whereas the proportion consuming milk declined from 74 per cent to 52 per cent. For the same age and gender group, average milk consumption dropped from 11 oz[1] to 7 oz per day, whereas average soft drink consumption rose from 8 oz to 13 oz per day, during the same period of time.

1 Oz = ounce(s). 1 ounce = 28.35 g.

Published by Oxford University Press

Table 1. Milk and carbonated soft drink consumption among children and adolescents in the USA: 1977–1978 to 1994–1996

	Milk		Soft drinks	
	1977–1978	1994–1996	1977–1978	1994–1996
Percentage consuming (%)				
All children	85	71	44	49
Children age 2–5	88	82	42	33
Children age 6–11	89	76	41	45
Male age 12–17	84	62	47	67
Female age 12–17	74	52	47	62
Average daily consumption (oz)				
All children	14	10	6	10
Children age 2–5	13	11	5	3
Children age 6–11	15	11	6	7
Male age 12–17	17	11	8	20
Female age 12–17	11	7	8	13

Sources: USDA's Nationwide Food Consumption Survey (NFCS), 1977–1978 and Continuing Survey of Food Intakes by Individuals (CSFII), 1994–1996, 1 day intake data.

Milk is a rich source of calcium, vitamins A and D, and other nutrients, whereas soft drinks have been called liquid candy (Jacobson, 1999). It has been recognised that a notable proportion of US children and especially adolescents fail to meet the recommended calcium intake.[2] For instance, during 1994–1996 only 13 per cent of female adolescents met the calcium recommendation. Higher intake of dietary calcium during adolescence and early adulthood increases peak bone mass and delays the onset of bone fracture later in life (USDHHS, 1988). It has been estimated that improved diets might save $5.1 billion to $10.7 billion each year in medical care costs, missed work, and premature deaths associated with osteoporosis-related hip fractures (Barefield, 1996). Therefore, the rapidly rising consumption of soft drinks and declining consumption of milk have been a cause of concerns about the diets of children and adolescents.

High soft drink consumption might also lead to excessive energy intake, which may contribute to childhood obesity, a growing problem in the USA. The number of overweight children 6–17 years of age has doubled during the last three decades (Troiano *et al.*, 1995; NIDDKD, 1996; Troiano and Flegal, 1998). Approximately one in every five children in the USA is now overweight. Although obesity-related morbidity and mortality rates are lower among children than adults, overweight children have a greater chance of developing into adult obesity and the risk increases with age and severity of

2 Reference is made to the Institute of Medicine's (1997) Dietary Reference Intake (DRI) for calcium.

the overweight (Whitaker *et al.*, 1997; Dietz, 1998). Obese children are also at higher risk of heart disease and diabetes and are more likely to suffer from social, emotional, and psychological stresses because they appear 'different' from peers (Gortmaker *et al.*, 1993). Each year, over $68 billion are spent on direct health care related to obesity, representing 6 per cent of total US health-care expenditure (Wolf, 1998).

Despite the growing concerns about displacement of milk and other more nutritious beverages (e.g. fruit juice) by soft drinks in the diets of children and adolescents, to our knowledge only one study has examined beverage choices among children and adolescents in the USA. Harnack *et al.* (1999) analyse 1994 USDA food consumption survey data to predict the odds of consuming soft drinks among children and adolescents as well as the association among the consumption of milk, soft drinks, and fruit juice. Soft drinks are found to displace milk and fruit juice, particularly at high levels of soft drink consumption.

Drawing on findings by Harnack *et al.* (1999) that soft drinks displace milk and fruit juice, we turn our attention to quantifying the choice among these beverages within a system of equations. We recognise the fact that some children and adolescents do not consume certain beverages, especially during a short survey period, resulting in zero values in observed consumption. As statistical estimation procedures not accounting for zero consumption produce biased and inconsistent parameter estimates, we address such issues of censoring. There has been long-standing interest among empirical analysts in the estimation of demand relationships with censored dependent variables. Full-information maximum-likelihood (FIML) estimation is known to be complicated in the estimation of censored systems as a result of the need to evaluate multiple probability integrals (Amemiya, 1974; Lee, 1993). As conjectured by Heyde (1997: 10), however, 'everything that can be done via likelihoods has a corresponding quasi-likelihood generalisation'. The current study attempts one such quasi-likelihood adventure. Our empirical analysis calls for estimation of a system of three equations. The FIML estimation, made possible by the small number of equations considered, provides a convenient benchmark to evaluate the performance and potential of the quasi maximum-likelihood (QML) procedure.[3]

The specific objectives of the current study are threefold: (1) to present the QML procedure for censored systems and assess its potential by comparing it with the FIML alternative; (2) to investigate the role of age in beverage consumption, which may shed some light on the displacement of milk by soft drinks; (3) to quantify the effects of other socioeconomic variables on beverage consumption by children and adolescents. We proceed with the econometric specification in the next section. The USDA data are then described and sample statistics are provided, followed by presentation of empirical results and concluding remarks.

3 Monte Carlo simulation would be a prohibitively computer-intensive alternative for comparison of the two estimators.

2. Econometric specification

To motivate the econometric specification, consider an individual's choice set (quantity vector) q with prices p. Also let c be a vector of personal characteristics. Optimal levels of q are determined by solving the constrained utility maximisation problem

$$\max_{q}\{U(q,c)|\,p'q=m\} \tag{1}$$

where m is income. Assuming the utility function $U(q,c)$ is continuous, increasing, and quasi-concave in q, optimal levels of quantities can be expressed as a function of prices, income and personal characteristics $h_i(p,m,c)$. In practice, the levels of consumption are also subject to non-negativity constraints, and therefore the amount consumed of each good can be either zero or positive, that is,

$$q_i = \max\{0, h_i(p,m,c)\}. \tag{2}$$

With a single cross-section, prices are assumed constant and are therefore absorbed into the constant term. Using a vector x to represent explanatory variables, a linear function to approximate the deterministic demand function, and a random error e_i to capture the unobservable for each equation, we consider a system of censored equations (Amemiya, 1974)

$$\begin{aligned} q_i &= x'\beta_i + e_i \quad \text{if } x'\beta_i + e_i > 0 \\ &= 0 \qquad\qquad \text{if } x'\beta_i + e_i \le 0, \quad i = 1,2,\ldots,n \end{aligned} \tag{3}$$

where β_i $(i = 1,2,\ldots,n)$ are vectors of parameters. We assume random error vector $[e_1,e_2,\ldots,e_n]'$ is n-dimensional normal with zero mean vector and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & & & \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & & \\ \vdots & \vdots & \ddots & \\ \rho_{n1}\sigma_n\sigma_1 & \rho_{n2}\sigma_n\sigma_2 & \cdots & \sigma_n^2 \end{bmatrix} \tag{4}$$

where σ_i are standard deviations of e_i $(i = 1,2,\ldots,n)$ and ρ_{ij} are correlation coefficients between e_i and e_j for all $i > j$. We denote the joint probability density function (pdf) of $[e_1,e_2,\ldots,e_n]'$ as $f(e_1,e_2,\ldots,e_n)$, and consider, without loss of generality, a regime (for an individual) in which the first l goods are consumed, with observed n-vector $q = [q_1,q_2,\ldots,q_l,0,0,\ldots,0]$. The contribution of this regime to the likelihood function is

$$L_c = \int_{-\infty}^{-x'\beta_n} \cdots \int_{-\infty}^{-x'\beta_{l+2}} \int_{-\infty}^{-x'\beta_{l+1}} f(e_1,e_2,\ldots,e_l,u_{l+1},u_{l+2},\ldots,u_n) \times \mathrm{d}u_{l+1}\,\mathrm{d}u_{l+2}\cdots\mathrm{d}u_n \tag{5}$$

where $e_i = q_i - x'\beta_i$ $(i = 1, 2, \ldots, l)$.[4] To simplify the likelihood function (5), we partition the error vector into

$$[e_1, e_2, \ldots, e_n]' = [e_1, e_2, \ldots, e_l | e_{l+1}, e_{l+2}, \ldots, e_n]'$$

with error covariance

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

partitioned conformably such that Σ_{11} is $l \times l$, Σ_{21} is $(n-l) \times l$, and Σ_{22} is $(n-l) \times (n-l)$. By conditioning, the joint pdf of random errors can be rewritten as

$$\begin{aligned} & f(e_1, e_2, \ldots, e_l, e_{l+1}, e_{l+2}, \ldots, e_n) \\ & \quad = h(e_{l+1}, e_{l+2}, \ldots, e_n | e_1, e_2, \ldots, e_l) g(e_1, e_2, \ldots, e_l) \end{aligned} \tag{6}$$

where $g(e_1, e_2, \ldots, e_l)$ is the marginal pdf of $e_1, e_2, \ldots, e_l$, distributed as l-dimensional normal with zero mean vector and covariance matrix Σ_{11}, and $h(e_{l+1}, e_{l+2}, \ldots, e_n | e_1, e_2, \ldots, e_l)$ is the conditional pdf of $[e_{l+1}, e_{l+2}, \ldots, e_n]'$ given $[e_1, e_2, \ldots, e_l]'$, distributed as $(n-l)$-dimensional normal with mean vector and covariance matrix, respectively (Kotz *et al.*, 2000),

$$\mu_{2.1} = \Sigma_{21}\Sigma_{11}^{-1}[e_1, e_2, \ldots, e_l]' \tag{7}$$

$$\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{21}'. \tag{8}$$

Using (6), the likelihood contribution (5) can be rewritten as

$$\begin{aligned} L_c = g(e_1, e_2, \ldots, e_l) \int_{-\infty}^{-x'\beta_n} \cdots \int_{-\infty}^{-x'\beta_{l+2}} \int_{-\infty}^{-x'\beta_{l+1}} h(u_{l+1}, u_{l+2}, \ldots, u_n | e_1, e_2, \ldots, e_l) \\ \times \mathrm{d}u_{l+1}\, \mathrm{d}u_{l+2} \cdots \mathrm{d}u_n. \end{aligned} \tag{9}$$

Using the mean vector (7) and covariance matrix (8), the multi-level integral in (9) can be evaluated as an $(n-l)$-dimensional standard normal cdf. Then, the likelihood function for a sample of T observations is

$$L = \prod_{t=1}^{T} \prod_{c} [L_c(q_t)]^{I_t(c)} \tag{10}$$

where $I_t(c)$ is a dichotomous indicator for observation t such that $I_t(c) = 1$ if observed consumption vector q_t lies in demand regime c and $I_t(c) = 0$ otherwise. FIML proceeds by maximising the likelihood function (10). The unknown parameters are β_i, σ_i $(i = 1, 2, \ldots, n)$ and ρ_{ij} for $i > j$.

4 With all goods consumed there is no integration and the likelihood contribution is simply $L_c = f(e_1, e_2, \ldots, e_n)$. At the other extreme, when all goods are zeros, integration is over all u_is and

$$L_c = \int_{-\infty}^{-x'\beta_n} \cdots \int_{-\infty}^{-x'\beta_2} \int_{-\infty}^{-x'\beta_1} f(u_1, u_2, \ldots, u_n)\, \mathrm{d}u_1,\, \mathrm{d}u_2 \cdots \mathrm{d}u_n.$$

For a large system with many censored dependent variables, evaluation of the multiple probability integral in equation (9) numerically can be prohibitively expensive.[5] A long list of empirical studies have used a two-step estimation procedure initiated by Heien and Wessells (1990). The procedure involves probit estimation in the first step and seemingly unrelated regression with selectivity regressors included in the system of equations in the second step. Shonkwiler and Yen (1999) point out that the augmented second-step equations in Heien and Wessells (1990) are inconsistent with the unconditional means of the dependent variables, and suggest an alternative estimation procedure which also involves probit estimation in the first step but with different sets of augmented equations in the second step. More recently, Perali and Chavas (2000) suggest estimating the unrestricted form of each demand equation using the jackknife technique; restricted demand parameters are then recovered by minimum distance estimation in the second step. The Shonkwiler and Yen (1999) and Perali and Chavas (2000) estimators are both consistent but suffer in efficiency.

Numerical difficulties with the multiple probability integrals in FIML can be overcome by a procedure parallel to the QML method in the multivariate probit literature (Avery *et al.*, 1983; Avery and Hotz, 1985). For the multivariate probit model

$$\begin{aligned} y_i &= 1 \quad \text{if } z_i'\alpha_i + v_i > 0 \\ &= 0 \quad \text{if } z_i'\alpha_i + v_i \le 0, \quad i = 1, 2, \ldots, n \end{aligned}$$

where y_i are binary dependent variables, z_i are vectors of explanatory variables, α_i are conformable parameter vectors, and v_i are random errors with error correlation matrix $[\tau_{ij}]$, Avery *et al.* (1983) suggest first obtaining univariate maximum-likelihood probit estimates $\hat{\alpha}_i$ for all i. Conditional on $\hat{\alpha}_i$, the correlation coefficients τ_{ij} for all $i > j$ are then estimated based on the quasi-likelihood (for a single observation)

$$\prod_{i=2}^{n} \prod_{j=1}^{i-1} \Psi(\kappa_i z_i'\hat{\alpha}_i, \kappa_j z_j'\hat{\alpha}_j, \kappa_i \kappa_j \tau_{ij}) \tag{11}$$

where $\kappa_i = 2y_i - 1$, $\kappa_j = 2y_j - 1$, and $\Psi(\cdot, \cdot, \cdot)$ is the standard bivariate normal cumulative distribution function (cdf). The sample quasi-likelihood function is the product of (11) over the sample. A more efficient procedure is to jointly estimate α_i $(i = 1, 2, \ldots, n)$ and τ_{ij} for all $i > j$, based on the quasi-likelihood (Avery and Hotz, 1985)

$$\prod_{i=2}^{n} \prod_{j=1}^{i-1} \Psi(\kappa_i z_i'\alpha_i, \kappa_j z_j'\alpha_j, \kappa_i \kappa_j \tau_{ij}).$$

5 Direct numerical evaluation of trivariate normal cdfs is possible in conventional software packages (e.g. Gauss, GQOPT), and cdfs of higher dimensions can be evaluated with recursive simulation techniques (e.g. Hajivassiliou, 1993).

The properties of the QML multivariate probit estimator are investigated by Kimhi (1999), and an application of the procedure is seen in Kimhi (1994).

This quasi-likelihood literature for the multivariate probit, with likelihood function constructed from pairwise bivariate normal probabilities, each of which accommodates an error correlation, lends direct support to the censored system problem, as both involve evaluations of multivariate normal probabilities. It is surprising that the quasi-likelihood approach has not been fully exploited in the censored system literature, the only exception being Harris and Shonkwiler (1997), who investigate the interdependence of retail businesses in Nevada and Utah, USA.[6] To apply the quasi-likelihood procedure to multivariate Tobit, the quasi-likelihood function is specified as the product of a sequence of bivariate Tobit likelihoods L_{ij} for all $i > j$:

$$L = \prod_{i=2}^{n} \prod_{j=1}^{i-1} L_{ij} \tag{12}$$

where, using (3), the bivariate Tobit likelihood for q_i and q_j is

$$\begin{aligned} L_{ij} = & \prod_{q_i=0, q_j=0} \Psi(-x'\beta_i/\sigma_i, -x'\beta_j/\sigma_j, \rho_{ij}) \\ & \times \prod_{q_i=0, q_j>0} \sigma_j^{-1}\phi[(q_j - x'\beta_j)/\sigma_j]\Phi\left[\frac{-x'\beta_i/\sigma_i - \rho_{ij}(q_j - x'\beta_j)/\sigma_j}{(1-\rho_{ij}^2)^{1/2}}\right] \\ & \times \prod_{q_i>0, q_j=0} \sigma_i^{-1}\phi[(q_i - x'\beta_i)/\sigma_i]\Phi\left[\frac{-x'\beta_j/\sigma_j - \rho_{ij}(q_i - x'\beta_i)/\sigma_i}{(1-\rho_{ij}^2)^{1/2}}\right] \\ & \times \prod_{q_i>0, q_j>0} \sigma_i^{-1}\sigma_j^{-1}\psi[(q_i - x'\beta_i)/\sigma_i, (q_j - x'\beta_j)/\sigma_j, \rho_{ij}]. \end{aligned} \tag{13}$$

In equation (13), $\psi(\cdot,\cdot,\cdot)$ is the bivariate standard normal pdf, and $\phi(\cdot)$ and $\Phi(\cdot)$ are univariate standard normal pdf and cdf, respectively. QML estimation proceeds by maximising (12), with respect to the same set of parameters as in FIML. The QML estimator differs from equation-by-equation Tobit estimators in that, as in FIML, cross-equation error correlations are accommodated, thus improving efficiency. Relative to FIML, the strength of the QML estimator is its ease of implementation, requiring only evaluations of bivariate normal cdfs, whereas the weakness is a loss in efficiency. Obviously, in a two-equation system ($n = 2$), the quasi-likelihood (12) reduces to $L = L_{12}$, which, using (13), is equivalent to the full-information likelihood contribution (9).

The small number of equations considered in the current application allows estimation of the censored system by FIML (as well as QML), which provides benchmark results for comparison with those of the QML alternative.

6 The application in Harris and Shonkwiler (1997) has little behavioural appeal in that a censored estimator is used not because the dependent variables ('pull factors' in retail business) are censored but because there are not sufficient observations (46 counties) to fit different statistical models.

3. Data

This study uses data from the 1994–1996 USDA Continuing Survey of Food Intakes by Individuals (CSFII). Children and adolescents (henceforth, children) who were in school and between the ages of 5 and 18 years are included in the analysis. Complete 2 day dietary recalls are available for 2,414 children. To avoid clustering effects as a result of multiple children from the same households, only one child is selected randomly from each household, reducing the sample size to 1,582. Further, it is hypothesised that the educational level of a meal planner affects children's food choices and therefore a small number of households with missing information on meal planners' education are excluded. In all, 1,546 children remain in the final sample.

Several categories of beverages are considered in this analysis, including milk, carbonated soft drinks, fruit drinks and ades,[7] fruit juice, and vegetable juice. By considering the nutritional profiles of these beverages, they are aggregated into three categories: milk, soft drinks (carbonated soft drinks, fruit drinks, and ades), and juice (fruit and vegetable juice). Fruit drinks contain no more than 10 per cent of juice, whereas juice has to be 100 per cent juice. Sample statistics, presented in Table 2, suggest that over a 2-day period, 83, 87, and 45 per cent of the sample consumed milk, soft drinks, and juice, respectively. As expected, these percentages of consuming children are higher than those suggested by the 1-day intake data (Table 1). The consumption of soft drinks was the highest, averaging about 33 oz over 2 days for the sample, followed by 20 oz for milk and 7 oz for juice. Among the consuming children, the corresponding means are 37 oz, 24 oz, and 15 oz, respectively.

Socioeconomic and demographic data are also collected for the sample households and their members. The explanatory variables include income (per capita), age, number of hours spent on TV watching, number of dietary recall days falling on weekends; as well as dummy variables indicating gender, meal planner's educational level (high school, college), urbanisation (central city, suburban), race (white, black), ethnicity (Hispanic), and region (Northeast, Midwest, and South).

4. Results

FIML estimation is carried out by Newton's method, and QML estimation by the quadratic hill-climbing algorithm (Goldfeld *et al.*, 1966), with standard errors for parameter estimates calculated from White's (1982) robust covariance matrix.[8] Results are presented in Table 3.

In comparing the FIML estimates with corresponding QML estimates, all coefficients are qualitatively consistent (in signs and significance) and are extremely close (in reference to their standard errors) between the two sets

7 Such as lemonade and orangeade.

8 Gauss codes for FIML estimation (using maxlik), FORTRAN codes for QML estimation (using GQOPT), and analytic gradients of the quasi-likelihood are available from the authors.

Table 2. Sample statistics of beverage consumption and explanatory variables

Variable	Definition	Mean	S.D.
Milk	Quantity consumed (oz/2 days)		
	Whole sample	20.22	17.94
	Consuming (82.7% of sample)	24.44	16.91
Soft drinks	Quantity consumed (oz/2 days)		
	Whole sample	32.55	32.22
	Consuming (87.3% of sample)	37.28	31.82
Juice	Quantity consumed (oz/2 days)		
	Whole sample	6.86	10.95
	Consuming (44.7% of sample)	15.35	11.74
Age	Individual age in years	10.59	3.81
Income	Per capita household income (USD$1,000)	10.79	7.35
TV	Number of hours watching TV over 2 days	5.55	4.08
Weekend	Number of survey days falling on weekend	0.56	0.56
Dummy variables (yes = 1; no = 0)			
Meal planner's education			
High sch.	High-school educated	0.37	
College	College-educated or higher	0.50	
No high sch.	Less than high-school (reference)	0.13	
Individual characteristics			
Male	Gender is male	0.51	
Hispanic	Ethnicity is Hispanic	0.13	
Urbanisation			
City	Resides in central city	0.27	
Suburban	Resides in suburban area	0.48	
Rural	Resides in rural area (reference)	0.25	
Region			
Northeast	Resides in the Northeast	0.17	
Midwest	Resides in the Midwest	0.25	
South	Resides in the South	0.36	
West	Resides in the West (reference)	0.22	
Race			
White	Race is White	0.76	
Black	Race is Black	0.14	
Other	Other race (reference)	0.10	

Source: Compiled from the CSFII, 1994–1996, 2 day intake data. Sample size 1,546.

Table 3. FIML and QML estimates of censored system of equations

	FIML			QML		
	Milk	Soft	Juice	Milk	Soft	Juice
Constant	23.006**	−19.626**	−9.711**	22.941**	−19.549**	−9.852**
	(3.263)	(6.043)	(3.835)	(3.256)	(6.037)	(3.829)
Age	−0.760**	3.238**	−0.060	−0.757**	3.238**	−0.052
	(0.144)	(0.276)	(0.158)	(0.144)	(0.276)	(0.160)
Income	−0.072	0.296**	0.172**	−0.074	0.295**	0.174**
	(0.087)	(0.146)	(0.088)	(0.087)	(0.145)	(0.088)
TV	−0.259*	0.634**	−0.213	−0.259*	0.632**	−0.219
	(0.141)	(0.223)	(0.155)	(0.141)	(0.223)	(0.156)
Weekend	−2.354**	3.792**	2.158**	−2.363**	3.782**	2.171**
	(0.946)	(1.545)	(1.047)	(0.946)	(1.542)	(1.048)
Male	7.892**	8.722**	0.958	7.924**	8.711**	1.028
	(1.056)	(1.684)	(1.145)	(1.055)	(1.682)	(1.146)
City	3.402**	−1.329	5.909**	3.447**	−1.316	5.994**
	(1.549)	(2.333)	(1.704)	(1.547)	(2.333)	(1.706)
Suburban	2.376*	0.537	3.005**	2.384*	0.540	3.043**
	(1.352)	(2.221)	(1.463)	(1.349)	(2.224)	(1.463)
White	1.415	3.940	−2.369	1.411	3.913	−2.352
	(2.047)	(2.998)	(2.113)	(2.037)	(2.995)	(2.124)
Black	−7.533**	5.607	−2.872	−7.548**	5.582	−2.853
	(2.457)	(3.801)	(2.589)	(2.454)	(3.798)	(2.602)
Hispanic	0.930	−1.409	2.7794	0.910	−1.388	2.814
	(1.797)	(2.465)	(1.849)	(1.807)	(2.468)	(1.857)
Northeast	−0.315	−4.140*	6.818**	−0.351	−4.110*	6.776**
	(1.614)	(2.444)	(1.864)	(1.761)	(2.440)	(1.880)
Midwest	1.637	8.556**	0.371	1.662	8.519**	0.372
	(1.677)	(2.782)	(1.776)	(1.712)	(2.781)	(1.810)
South	−3.228**	0.932	0.286	−3.227**	0.911	0.233
	(1.455)	(2.224)	(1.660)	(1.497)	(2.232)	(1.674)
High school	1.504	−2.889	0.605	1.539	−2.911	0.583
	(1.698)	(2.904)	(1.938)	(1.688)	(2.906)	(1.971)
College	1.576	−4.129	5.092**	1.609	−4.125	5.054**
	(1.726)	(2.910)	(2.027)	(1.717)	(2.917)	(2.045)
σ	19.920**	32.382**	19.679**	19.909**	32.364**	19.680**
	(0.565)	(1.949)	(0.705)	(0.563)	(1.947)	(0.705)
$\rho_{SOFT,MILK}$	−0.234**			−0.235**		
	(0.029)			(0.029)		
$\rho_{JUICE,MILK}$	−0.028			−0.031		
	(0.032)			(0.031)		
$\rho_{JUICE,SOFT}$	−0.141**			−0.141**		
	(0.035)			(0.034)		
Log-likelihood		−16221.746			−32495.932	

Asymptotic standard errors are given in parentheses. ** and * denote significance at the 5 per cent and 10 per cent levels, respectively.

of estimates. As expected, the more notable differences between the two sets of estimates occur with variables that are not significant, that is, those parameters that are estimated with less precision (larger standard errors). Of particular interest are results of the three error correlation coefficients, which are extremely close between the FIML and QML estimates, all up to three decimal places. These findings are remarkable, because FIML estimates the error correlations jointly across the system whereas the QML procedure obtains the system estimates by pairwise bivariate analyses. Similarity of the two sets of estimates provides empirical evidence that, for the current application, the QML estimator is a good alternative to the FIML estimator.

The error correlations are significant between milk and soft drinks, and between soft drinks and juice, but not significant between milk and juice. Significance of these correlation coefficients suggests that the system estimators are appropriate and that the FIML and QML estimates are preferable to the (less efficient) equation-by-equation Tobit estimates (not reported).

According to both FIML and QML results, age has a (significantly) negative effect on milk but a (significantly) positive effect on soft drinks. Thus, milk is displaced by soft drinks as a child grows older. This result is also reported by Harnack *et al.* (1999) based on probit analysis. Furthermore, our results can be used, as shown below, to quantify the amount of milk displaced by soft drinks as a child grows older. As to other socioeconomic variables, children's life style also affects beverage choices. While spending more time watching TV, a child consumes more soft drinks and less milk and juice, whereas during weekends, a child consumes less milk but more juice and soft drinks. This may be because children are watching more TV on weekends than weekdays. Another plausible explanation is that milk is a mandated item in the National School Lunch Program and children are more likely to eat school lunches during weekdays than weekends. Moreover, children are more likely to skip breakfast during weekends, reducing milk consumption.

Income has a positive effect on soft drink and juice consumption but is insignificant in the milk equation, raising questions about the effectiveness of government income support programs (e.g. food stamp) in promoting milk consumption and calcium intake. Meal planner's education, reflected by dummy variable 'college', has a positive effect on juice consumption but insignificant effects on soft drinks and milk. The effects of gender, race and ethnicity are consistent with expectations. In particular, although being male is not significant in the juice equation, it is significant and positive in both the milk and soft drinks equations. Thus, boys tend to consume more milk and soft drinks than girls. As the black population is known to be less lactose tolerant than others, it is not surprising that black children consume less milk than their white counterparts (but consume about the same amounts of soft drinks and juice).

Among the other variables considered, urbanisation (central city and suburban) and regions (Northeast, Midwest and South) also play important roles in determining beverage consumption. In particular, children living in the city and suburban areas tend to consume more juice and milk than those living in rural areas. Compared with children living in the Western states, children in the Southern states consume less milk; children in the Midwestern states consume more soft drinks; and children in the Northeast consume less soft drinks but more juice. It is worth noting that, because prices are not available in the cross-sectional data we use, these regional dummies might capture the effects of prices.[9]

As in other limited dependent variable models, the effects of variables can be quantified further by calculating the elasticities (McDonald and Moffitt, 1980). These elasticities also facilitate further comparisons between the FIML and QML results. For each equation, the probability of a positive observation is

$$\Pr(q_i > 0) = \Phi(x\beta_i/\sigma_i) \tag{14}$$

and the conditional mean of the dependent variable is

$$E(q_i|q_i > 0) = x'\beta_i + \sigma_i\phi(x'\beta_i/\sigma_i)/\Phi(x'\beta_i/\sigma_i). \tag{15}$$

Using (14) and (15), the unconditional mean is

$$E(q_i) = \Phi(x'\beta_i/\sigma_i)x'\beta_i + \sigma_i\phi(x'\beta_i/\sigma_i). \tag{16}$$

Elasticities with respect to explanatory variables are derived by differentiating (14), (15) and (16), respectively, and are calculated at the sample means of all variables.[10] In addition, standard errors for elasticities are approximated by the delta method (Spanos, 1999: 493). The conditional elasticities measure the percentage changes in beverage consumption among those who are already consuming the beverage, whereas the unconditional elasticities add the effects of switching between consumption and non-consumption to the conditional elasticities.

These elasticities provide a better quantitative assessment of the impacts of each variable on relevant components of the dependent variable which are otherwise masked in the parameter estimates. Even though it does not occur in this study, it is possible that a particular variable may have no significant effect on the probability of consumption, but have significant effects on conditional and unconditional consumption. For instance, boys may not be more likely to consume soft drinks than girls but boys who consume soft drinks may drink more than girls who also consume soft drinks; that is, the dummy variable for gender could be associated with a significant, positive conditional elasticity.

Table 4 presents the elasticities with respect to continuous variables, along with their standard errors. Only those calculated from the FIML estimates are shown because, as both sets of parameter estimates are extremely close, so too are the corresponding elasticities. Soft drinks have the highest elasticities with respect to age. As age increases by 1 per cent, all else equal, the probability of

9 An attempt to include seasonal dummies was unsuccessful, resulting in statistical insignificance.

10 Elasticity formulae are available from the authors.

Table 4. Elasticities with respect to continuous variables (FIML)

	Milk			Soft drinks			Juice		
	Probability	Cond. level	Uncond. level	Probability	Cond. level	Uncond. level	Probability	Cond. level	Uncond. level
Age	−0.130**	−0.198**	−0.327**	0.333**	0.518**	0.851**	−0.029	−0.014	−0.043
	(0.026)	(0.039)	(0.064)	(0.029)	(0.034)	(0.057)	(0.075)	(0.038)	(0.113)
Income	−0.013	−0.019	−0.032	0.031**	0.048**	0.079**	0.083**	0.042**	0.125**
	(0.015)	(0.023)	(0.038)	(0.015)	(0.024)	(0.039)	(0.043)	(0.021)	(0.064)
TV	−0.023*	−0.035*	−0.058*	0.034**	0.053**	0.087**	−0.053	−0.027	−0.080
	(0.013)	(0.019)	(0.032)	(0.012)	(0.019)	(0.031)	(0.039)	(0.020)	(0.058)
Weekend	−0.021**	−0.032**	−0.053**	0.021**	0.032**	0.052**	0.054**	0.027**	0.081**
	(0.009)	(0.013)	(0.021)	(0.008)	(0.013)	(0.022)	(0.026)	(0.013)	(0.040)

Asymptotic standard errors are given in parentheses. ** and * denote significance at the 5 per cent and 10 per cent levels, respectively.

consuming soft drinks increases by 0.33 per cent and, conditional on consumption, the level increases by 0.52 per cent. Overall, the elasticity of unconditional level suggests that, on average, as age increases by 1 per cent, the level of soft drinks consumed increases by 0.85 per cent.

The displacement of milk by soft drinks is also seen and quantified in the elasticities. The negative elasticities of milk with respect to age indicate that as children become older, milk consumption drops and fewer children drink milk. As to magnitudes, according to the elasticity of unconditional level, each percentage increase in age results in a total decline of milk consumption by 0.33 per cent and an increase of 0.85 per cent in soft drink consumption. The mean milk consumption is 20 oz for 2 days, whereas soft drink consumption is 33 oz at the mean. Using these mean consumption levels and unconditional elasticities, each 1 oz decline in milk consumption is accompanied by about a 4.2 oz increase in soft drinks. The CSFII data contain the nutrient profiles of all foods reported. The nutrient database indicates that each 1 oz of soft drinks and milk contains an average of 11 and 15 calories, respectively.[11] Therefore, the displacement relationship suggests a gain of 31 calories for each 1 oz of milk displaced. Each 1 oz of milk contains 34 mg of calcium, whereas calcium in soft drinks is negligible.

Children from higher-income households are more likely to consume and also consume more of juice and soft drinks, but not milk. As income increases by 1 per cent, all else equal, total juice consumption increases by 0.13 per cent whereas total soft drink consumption increases by 0.08 per cent. TV watching is associated with both a higher probability and higher level of soft drink consumption, whereas its effects on milk are opposite. During weekends, children are more likely to consume and also consume more of juice and soft drinks, whereas they are less likely to consume milk and their consumption level also declines.

The effects of binary variables are calculated differently from the elasticities with respect to the continuous variables. In particular, the effects of each discrete variable are calculated as changes in the probability (14), conditional level (15), and unconditional level (16) resulting from a simulated change in the variable from zero to one, while holding all other variables constant. As a result of the very close similarity between the elasticities based on the FIML and the QML estimates, only the former are presented in Table 5. Compared with girls, boys are more likely to consume soft drinks and milk and at the same time they consume larger quantities of both beverages. Black children not only are less likely to consume milk, they also tend to consume smaller quantity of milk than other children.

Meal planner's education has positive effects on the probability and levels of juice consumption, but insignificant effects on milk and soft drink consumption. Household income is likely to correlate with education, making it difficult to separate the effects of the two variables. Consistent with the

11 Because there are various types of soft drinks and milk with different caloric contents, the averages are weighted by their amounts consumed by US children.

Table 5. Effects of discrete variables (FIML)

	Milk			Soft drinks			Juice		
	Probability	Cond. level	Uncond. level	Probability	Cond. level	Uncond. level	Probability	Cond. level	Uncond. level
Male	0.051**	2.332**	3.166**	0.032**	2.695**	3.601**	0.009	0.145	0.184
City	0.046**	2.035**	2.769**	−0.011	−0.802	−1.089	0.118**	1.977**	2.628**
Suburban	0.033*	1.404*	1.918*	0.004	0.329	0.444	0.059**	0.960**	1.249**
White	0.005	0.197	0.271	0.007	0.584	0.787	−0.011	−0.173	−0.217
Black	−0.104**	−3.398**	−4.771**	0.036	3.031	4.044	−0.047	−0.722	−0.891
Hispanic	0.011	0.467	0.643	−0.010	−0.739	−1.002	0.048	0.764	0.989
Northeast	−0.004	−0.149	−0.206	−0.029*	−2.040*	−2.782*	0.113**	1.884**	2.500**
Midwest	0.017	0.710	0.975	0.046**	4.073**	5.409**	−0.005	0.085	0.107
South	−0.031**	−1.145**	−1.590**	0.005	0.363	0.489	−0.004	0.056	0.070
High sch.	0.013	0.546	0.751	−0.015	−1.092	−1.483	0.007	0.116	0.148
College	0.011	0.457	0.628	−0.017	−1.241	−1.687	0.050**	0.809**	1.048**

Asymptotic standard errors are given in parentheses. ** and * denote significance at the 5 per cent and 10 per cent levels, respectively.

parameter estimates reported in Table 3, children living in rural areas consume less juice and milk and, additionally, they are less likely to consume juice and milk than children living in cities and suburban areas. There are some significant regional variations in the probability and consumption level of these three beverages.

5. Concluding remarks

There has been emerging interest among policy makers in calcium deficiency and displacement of milk and the more nutritious juice by soft drinks among children and adolescents in the USA. Building on previous findings of such displacement, the current study attempts to quantify such displacement and investigate the effects of other factors in the consumption of milk, soft drinks, and juice—the primary beverages consumed by children and adolescents. Even though the CSFII prevents examination of the role of relative prices in a demand system, it allows investigation of other important issues that are not possible in aggregate time series data. As in other applications with microdata, censoring in the dependent variables presents a challenging statistical problem. Although a Monte Carlo simulation is not feasible for comparing the QML and the computer-intensive FIML estimators used in this study, our comparison between the two sets of empirical estimates offers useful insights into the usefulness of the QML estimator. We provide some evidence that, for the current application, the QML estimator performs as well as the FIML estimator in that the two procedures produce remarkably close empirical estimates in terms of signs and magnitudes of parameter estimates and elasticities. We conjecture that the estimator is worthy of further consideration in other applications, and can be valuable in large systems with many censored dependent variables when FIML estimation is prohibitive. We also note that the QML procedure can be equally applicable in theoretically plausible (linear and nonlinear) demand systems.[12]

Echoing previous findings of displacement in beverages among children and adolescents, we find that consumption of soft drinks increases and consumption of milk decreases as a child becomes older. This finding suggests great potential in promoting milk consumption by campaigning among the adolescents. On average, for each 1 oz reduction in milk consumption, a child consumes 4.2 oz of soft drinks, resulting in a net gain of 31 calories and a loss of about 34 mg of calcium. Therefore, the changing beverage consumption among children may have contributed to the increased prevalence in children's overweight and obesity. Further, girls are less likely to consume and also consume less milk than boys. This finding is consistent with the observation among policy makers that calcium deficiency is particularly severe among female adolescents, which triggered a series of campaigns by government agencies. The examples include the joint 'national bone health campaign'

12 For a demand system, deterministic components $x'\beta_i$ in the system can be replaced with $f_i(x, \beta)$ for all i, where β is a vector of all parameters. Cross-equation restrictions such as homogeneity and symmetry can be imposed by regular means.

by the Center for Disease Control and Prevention (CDC), Department of Health and Human Services' Office on Women's Health, and National Osteoporosis Foundation (CDC, 2001); the 'milk matters campaign' by the National Institute of Child Health and Human Development (NICHD, 2001); and the 'crash course on calcium' and 'milk mustache' campaigns by the National Institute of Health (NIH, 1997). The milk mustache campaign even targets teenage girls by using role models such as young female Olympic athletes. Our finding suggests that more of these campaigns could be valuable in improving the nutritional wellbeing of female adolescents and children in general.

The results of this study suggest a more active role for parents in improving children's food choices. Soft drink consumption is positively related to TV watching, over which parents can exert more control. During weekends, children tend to consume more soft drinks and less milk. As children age, they tend to consume more soft drinks and less milk so that nutrition education by the parents and the public should start early in childhood.

Acknowledgements

Research for this paper was supported by USDA-ERS Cooperative Agreement No. 43-3AEM-0-80042. The views in this paper are those of the authors and do not necessarily reflect the views or policies of the US Department of Agriculture.

References

Amemiya, T. (1974). Multivariate regression and simultaneous equation models when the dependent variables are truncated normal. *Econometrica* 42(6): 999–1012.

Avery, R. B. and Hotz, V. J. (1985). *HotzTran User's Manual*. Old Greenwich, CT: CERA Economic Consultants.

Avery, R. B., Hansen, L. P. and Hotz, V. J. (1983). Multi-period probit models and orthogonality condition estimation. *International Economic Review* 24(1): 21–35.

Barefield, E. (1996). Osteoporosis-related hip fractures cost $13 billion to $18 billion yearly. *FoodReview* 19(January–April): 31–36.

CDC (2001). Powerful girls have powerful bones. National Institute of Health, Center for Disease Control and Prevention. http://www.cdc.gov/powerfulbones/ [Accessed 3 October 2001].

Dietz, W. H. (1998). Health consequences of obesity in youth: childhood predictors of adult disease. *Pediatrics* 101(3) (Suppl.): 518–525.

Goldfeld, S. M., Quandt, R. E. and Trotter, H. F. (1966). Maximization by quadratic hill-climbing. *Econometrica* 34(3): 541–551.

Gortmaker, S. L., Must, A., Perrin, J. M., Sobol, A. M. and Dietz, W. H. (1993). Social and economic consequences of overweight in adolescence and young adulthood. *New England Journal of Medicine* 329(14): 1008–1012.

Hajivassiliou, V. A. (1993). Simulation estimation methods for limited dependent variable models. In Maddala, G. S., Rao, C. R. and Vinod, H. D. (eds), *Handbook of Statistic, Vol. 19*. Amsterdam: North-Holland, 519–543.

Harnack, L., Stang, J. and Story, M. (1999). Soft drink consumption among US children and adolescents: nutritional consequences. *Journal of the American Dietetic Association* 99(4): 436–441.

Harris, T. R. and Shonkwiler, J. S. (1997). Interdependence of retail businesses. *Growth and Change* 28(4): 520–533.

Heien, D. and Wessells, C. R. (1990). Demand systems estimation with microdata: a censored regression approach. *Journal of Business and Economic Statistics* 8(3): 365–371.

Heyde, C. C. (1997). *Quasi-Likelihood and its Application: a General Approach to Optimal Parameter Estimation*. New York: Springer.

Institute of Medicine, National Academy of Sciences (1997). *Dietary Reference Intake: Calcium, Phosphorus, Magnesium, Vitamin D, and Fluoride*. Washington, DC: National Academy Press.

Jacobson, M. F. (1999). Liquid candy: how soft drinks are harming Americans' health. Center for Science in the Public Interest. (http://www.cspinet.org/sodapop/liquid _candy.htm) [Accessed 30 November 2000].

Kimhi, A. (1994). Quasi maximum likelihood estimation of multivariate probit models: farm couples' labor participation. *American Journal of Agricultural Economics* 76(4): 828–835.

Kimhi, A. (1999). On the tradeoff between computational simplicity and asymptotic properties in multivariate probit. *Computational Economics* 13(1): 93–101.

Kotz, S., Balakrishnan, N. and Johnson, N. L. (2000). *Continuous Multivariate Distributions, Vol. 1: Models and Applications*, 2nd edn. New York: John Wiley.

Lee, L.-F. (1993). Multivariate tobit models in econometrics. In Maddala, G. S., Rao, C. R. and Vinod, H. D. (eds), *Handbook of Statistic, Vol. 11*. Amsterdam: North-Holland, 145–173.

McDonald, J. F. and Moffitt, R. A. (1980). The uses of tobit analysis. *Review of Economics and Statistics* 62(2): 318–321.

NICHD (2001). Milk matters campaign. National Institute of Child Health and Human Development. http://www.nichd.nih.gov/milk/milk.cfm [Accessed 28 September 2001].

NIDDKD (1996). Statistics related to overweight and obesity. NIH Publication 96-4158. National Institute of Diabetes and Digestive and Kidney Diseases. Rockville, MD: National Institute of Health.

NIH (1997). Secretary Shalala announces partnership to increase teen calcium consumption. NIH News Alert, November 12. National Institute of Health.

Perali, F. and Chavas, J. P. (2000). Estimation of censored demand equations from large cross-section data. *American Journal of Agricultural Economics* 82(4): 1022–1037.

Shonkwiler, J. S. and Yen, S. T. (1999). Two-step estimation of a censored system. *American Journal of Agricultural Economics* 81(4): 972–982.

Spanos, A. (1999). *Probability Theory and Statistical Inference: Econometric Modeling with Observational Data*. Cambridge: Cambridge University Press.

Troiano, R. P. and Flegal, K. M. (1998). Overweight children and adolescents: description, epidemiology, and demographics. *Pediatrics* 101(3) (Suppl.): 497–504.

Troiano, R. P., Flegal, K. M., Kuczmarski, R. J., Campbell, S. M. and Johnson, C. L. (1995). Overweight prevalence and trends for children and adolescents. The NHANES Surveys, 1963–1991. *Archives of Pediatric and Adolescent Medicine* 149(10): 1085–1091.

USDHHS (1988). The Surgeon General's Report on Nutrition and Health. USDHHS Publication 88-50210. Washington, DC: US Department of Health and Human Services.

Whitaker, R. C., Wright, J. A., Pepe, M. S., Seidel, K. D. and Dietz, W. H. (1997). Predicting obesity in young adulthood from childhood and parental obesity. *New England Journal of Medicine* 337(13): 869–873.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* 50(1): 1–25.

Wolf, A. M. (1998). What is the economic cost for treating obesity? *Obesity Research* 6(Suppl. 1): 2S–7S.

Corresponding author: Steven T. Yen, Department of Agricultural Economics, University of Tennessee, Knoxville, TN 37996-4518, USA. E-mail: syen@utk.edu.