Journal of Rural Development 17(1994) : 99-116

TWO-STAGE ESTIMATION OF A CENSORED DEMAND SYSTEM

CHOI JI-HYEON*

I. Introduction

Recently, economists have utilized systems to estimate the demand for particular food products using household level data. The use of household data for a detailed commodity analysis, however, creates a major estimation problem. This problem stems from the fact that households are observed to consume zero amounts of the various commodities under consideration (Heien and Wessells, 1990). According to the USDA's 1987/1988 National Food Consumption Survey (NFCS) data, 87.4 percent of the households surveyed consumed fresh vegetables, and only 66.2 percent and 34.6 percent of the households consumed canned and frozen vegetables, respectively, during the one week survey period.

For estimation of food demand systems with such a sample, the estimation method should allow zero expenditures to occur with positive probability. However, the standard estimation techniques, such as three stage least squares (3SLS) and seemingly unrelated regression (SUR), used in most studies assume that expenditure (or shares) follow a joint normal distribution. Without modifications, this does not allow for a positive probability of zero expenditures or shares. These conventional estimation techniques consequently yield inconsistent estimates of the parameters (Wales and Woodland, 1983).

Another problem in cross-sectional data sets is missing price information due to zero consumption. Nonconsuming household

---

* Research Associate, Korea Rural Economic Institute, Seoul, Korea. The author is deeply indebted to Dr. Vicki McCracken in Washington State University, who has supervised this research.

observations do not provide price information for commodities they do not consume. This constitutes a major problem when price is used as an explanatory variable in demand estimation. The reasons why households do not consume the food item during the survey week may be economic related or may indicate a unique preference towards (or against) the food.

The overall objective of this research is to assess the impacts of economic and socio-demographic variables on vegetable consumption. The specific objectives include: 1) developing an econometric framework for modeling zero consumption behavior, and 2) evaluating the impacts of economic and socio-demographic variables on fresh, canned, and frozen vegetable consumption.

II. Development of Model

In order to evaluate these impacts within the vegetable demand system, the Almost Ideal Demand System (AIDS) model was selected as the framework. Deaton and Muellbauer (1980a) proposed the AIDS model which is of comparable generality as are the Rotterdam and translog models. The AIDS model can be derived from the PIGLOG (expenditure) function which defines the minimum expenditure necessary to attain a specific utility level at given prices. The AIDS demand functions in budget share form are as follows:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln (y/p^*) \qquad (1)$$

where w_i is the budget share of good i, y is the expenditure on the group of commodities, p_i is the price of good i, and P* is Stone's index used as a proxy for the group price index, P. Stone's index can be written as:

$$\ln p^* = \sum_{j=1}^{n} w_k \ln p_k \qquad (2)$$

where w_k is the budget share of good k. The model that uses Stone's index is referred to as the "linear approximate AIDS" (LA/AIDS). LA/AIDS, like AIDS, provides an arbitrary first order approximation to any demand system (Deaton and Mueallbuer, 1980a). In addition, by using LA/AIDS, the nonlinearity problem in estimation can be avoided. In this study, the LA/AIDS model was chosen to estimate the parameters of a vegetable demand system.

The values of the parameters α_i, γ_{ij}, and β_i, required to satisfy the theoretical general restrictions of adding up, homogeneity, and symmetry can be achieved by imposing, respectively:

$$\sum_{i=1}^{n} \alpha_i = 1; \quad \sum_{i=1}^{n} \gamma_{ij} = 0; \quad \sum_{i=1}^{n} \beta_i = 0 \tag{3}$$

$$\sum_{j=1}^{n} \gamma_{ij} = 0 \tag{4}$$

$$\gamma_{ij} = \gamma_{ji} \tag{5}$$

When food demand systems are empirically modelled, economists typically face the problems caused by the large number of commodities consumed by households. The usual way to approach to this problem is to assume a priori some type of structure on consumers' preferences, the most common assumption being that of weak separability.

In empirical work, a popular application of weak separability is to systems of demand equations in order to limit the number of equations in the system. Separability has frequently been utilized in conjunction with the AIDS model to estimate conditional demand systems. Weak separability is a prerequisite for application of the two stage maximization procedure (Deaton and Muellbuer, 1980b).

For this study, it is assumed that vegetables, the food group of interest in this study, are weakly separable from other food commodities such as meats, dairy products, and fruits as well as nonfood commodities. This specification implies that the MRS between fresh vegetables and canned vegetables is independent of the amount consumed of other food commodities, say meats or fruits.

Under this assumption, at the first stage of the two stage budgeting, consumers allocate total income among vegetables, other food, and nonfood groups.[1] At the second stage, consumers allocate the vegetable expenditure among fresh, canned, and frozen vegetables. Thus, commodities other than vegetables are excluded from the analysis.

In this study, the socio-demographic variables and nutrition and dietary related-variables are incorporated into the AIDS demand functions as proxies for consumer preferences in order to control for their effects on the demand for vegetables. Thus, the intercept term, α_i, in equation (1) can be represented as:

$$\alpha_i = \rho_i + \sum_{k=1}^{n} \rho_{ik} d_k. \qquad i = 1, 2, 3 \tag{6}$$

where i indicates vegetable group, ρ_i is redefined as an intercept in equation (1), d_k is demographic variable or nutrition and dietary related variable k of which there are s, and ρ_{ik} are the coefficients associated with these variables.

The LA/AIDS demand function (1), incorporating the socio-demographic and nutrition and dietary related-variables as in equation (6), can be written as:

$$w_i = \rho_i + \sum_{k=1}^{s} \rho_{ik} d_k + \sum_{j=1}^{3} \gamma_{ij} \ln p_j + \beta_i \ln (y/p^*) \tag{7}$$

where the three vegetable groups considered are canned (i=1), frozen (i=2), and fresh (i=3) vegetables; w_i, p_j, ρ_i and y/p^* are as previously defined.

Equation (7) constitutes the empirical model of the LA/AIDS vegetable demand system.

After incorporating the socio-demographic variables into equation (7) the adding up conditions in equation (3) are rewritten as follows:

---

[1] In this study, the first-stage allocation was taken as given and was not estimated.

$$\sum_{i=1}^{3} \rho_i = 1 \qquad \sum_{i=1}^{3} \gamma_{ij} = 0 \qquad \sum_{i=1}^{3} \beta_i = 0 \tag{8}$$

$$\sum_{k=1}^{s} \rho_{ik} = 0 \qquad i,\ j = 1,\ 2,\ 3$$

III. Estimation Procedure

1. Two Stage Estimation Method

The application of the standard estimation methods such as SUR or 3SLS with no correction for sample selectivity would yield biased and inconsistent estimates because the error terms have expectations which are not zero and depend upon the exogenous variables.

Lee (1978) generalized the two stage methods proposed by Heckman (1976) and Amemiya (1977). Two stage estimation methods such as Heckman's two stage method has been applied in applied demand analyses (e.g., labor demand). In these applications, typically only the positive observations were used at the second stage estimation.

Heien and Wessells (1990) estimated the second stage model of the two stage procedure using all observations. They used different inverse Mill's ratios for the limit sample and non-limit sample observations. Maddala (1983) showed that the full sample, instead of the positive observations only, can be used in the second stage estimation of the two step estimation procedure. The expectation of the budget share, w_i, in the AIDS demand system for the full sample is:

$$\begin{aligned} E(w_i) &= \text{Prob}(w_i>0)\, E(w_i|w_i>0) + \text{Prob}(w_i=0)\, E(w_i|w_i=0) \\ &= \Phi_i\,(\beta' X_i + \sigma\phi_i/\Phi_i) + (1 - \Phi_i) \times 0 \\ &= \Phi_i\,\beta' X_i + \sigma\phi_i \end{aligned} \tag{9}$$

where Φ_i and ϕ_i are the cumulative normal distribution function and the standard normal density function, respectively, and σ is the standard deviation of the normal error term, X_i are explanatory variables in the AIDS model, and β along with σ is the vector of the coefficients to be estimated. In view of the econometric specification in this study, the second stage equation (9) is more plausible than the second stage equation suggested by Lee (1978) since the second stage equation is usually used for the purpose of the prediction.

In this study, based on equation (9) and using all observations, the second stage AIDS demand system is estimated. In addition, based on equation (9) and using all observations, the AIDS demand system is estimated by SUR at the first stage without the second stage, for purposes of comparing the results.

The two step estimation method consists of the first stage probit estimation and the second stage SUR estimation. The estimation procedure is as follows. The first step involves the estimation of a probit regression that determines the probability that a household consumes each of the vegetable groups (excluding fresh vegetables since all households included in the sample consumed fresh vegetables). These probit coefficients are then used to estimate Φ_i and ϕ_i to correct for selectivity bias, and then the estimated Φ_i and ϕ_i are used in the second stage estimation.[2]

2. Development of Empirical Model

In the first stage, the vegetable consumption decision is modelled by incorporating the socio-demographic and nutrition and dietary related variables into the binary choice problem. The probit equation was specified as follows:

$$\begin{aligned} I_i = {} & \delta_i + \delta_1 \text{FOODEXP} + \delta_2\text{CENCITY} + \delta_3\text{HHSIZE} \\ & + \delta_4\text{HPVEGI} + \delta_5\text{BHEAD} + \delta_6\text{COLLEGE} + \delta_7\text{SPRING88} \\ & + \delta_8\text{WINTER} + \delta_9\text{FALL} + \delta_{10}\text{BCOLLAR} + \delta_{11}\text{FOODSTMP} \\ & + \delta_{12}\text{WHITE} + \delta_{13}\text{INFOMED} + \delta_{14}\text{INFOCOM} \\ & + \delta_{15}\text{INFOPROF} + \delta_{16}\text{INFOREL} + \delta_{17}\text{RNVTRFE} \\ & + \delta_{18}\text{RNVTRVC} + \delta_{19}\text{NCENT} + \delta_{20}\text{WEST} + u_i \end{aligned} \tag{10}$$

where I_i is the binary choice variable which indicates whether

[2] In this chapter, the theoretical framework of the two step estimation procedure is not discussed. For details of the two step procedure, refer to Choi (1993)

a household consumed vegetable group i during the survey period. Specifically, I_i is 1 if the household consumesvegetable group i (i.e., $W_i>0$) and I_i is zero if the household does not consume vegetable group i (i.e., $W_i=0$). The explanatory variables included total food expenditure (FOODEXP), urbanization (CENCITY), household size (HHSIZE), presence of home grown vegetables (HPVEGI), presence of female and male household heads (BHEAD), education level of household meal planner (COLLEGE), season (SPRING88, WINTER, and FALL), occupation of household head (BCOLLAR), food stamp participation (FOODSTMP), race of the household head (WHITE), source of nutrition information (INFOMED, INFOREL, INFOPROF, and INFOCOM), dietary status of household (RNVTRFE, RNVTRVC), and location of residence (NCENT, WEST); and u_i is an error term. Equation (10) is estimated by the probit method for each of the two vegetable groups: canned vegetables and frozen vegetables.

The second step of the two-step estimation procedure involves the estimation of the AIDS vegetable demand model (7) based on equation (9). The AIDS model was initially specified as:

$$\begin{aligned} w_i = {} & \rho_i + \gamma_{i1}LPCANND^* + \gamma_{i2}LPFROZEN^* + \gamma_{i3}LPFRESH^* \\ & + \beta_i LEP^* + \rho_{i1}CENCITY^* + \rho_{i2}HHSIZE^* + \rho_{i3}HPVEGI^* \\ & + \rho_{i4}BHEAD^* + \rho_{i5}COLLEGE^* + \rho_{i6}SPRING88^* \\ & + \rho_{i7}WINTER^* + \rho_{i8}FALL^* + \rho_{i9}BCOLLAR^* \\ & + \rho_{i10}FOODSTMP^* + \rho_{i11}WHITE^* + \rho_{i12}INFOPROF^* \\ & + \rho_{i13}INFOCOM^* + \rho_{i14}RNVTRFE^* + \rho_{i15}RNVTRVC^* \\ & + \rho_{i16}NCENT^* + \rho_{i17}WEST^* + \sigma_i PDF + e_i \end{aligned} \tag{11}$$

where w_i is the budget share of vegetable group i (i=1,2,3). The set of explanatory variables differs between equation (10) and equation (11). These differences include: (1) logarithm of vegetable price variables (LPCANND, LPFROZEN, and LPFRESH), vegetable group expenditure deflated by price index (LEP), and ϕ_i (PDF) are only included in equation (11), (2) total food expenditure (FOODEXP) is only included in equation (10), and (3) among the nutrition information source variables, only INFOMED and INFOREL variables are included in equation (11).

Following equation (9), all explanatory variables except for

PDF in the second stage AIDS demand functions were transformed (e.g., LPCANND*, HHSIZE*, and WHITE*).

Equation (11) was estimated by SUR with the restrictions (4) and (5). The adding up condition (8) was used to recover the coefficients estimates of the deleted (fresh) vegetables budget share equations.

The data used in the empirical analysis are from the household component of the 1987-88 NFCS data, which was conducted by the Human Nutrition Information Service (HNIS), U.S. Department of Agriculture. The NFCS data were collected over five seasons from April, 1987 to June, 1988. The total sample consists of 4,495 households. In this study, two different types of data screening were performed: screening with respect to 1) abnormal food consumption patterns, and 2) abnormal prices.[3]

Another problem in cross-sectional data sets is missing price information due to zero consumption. In the prediction of missing prices for the vegetable groups, a household's nonconsuming behavior was modelled by the Type 2 Tobit model, which was estimated using Heckman's two-step procedure that corrects for sample selectivity bias.

IV. Estimation Results

Estimation results for the AIDS vegetable demand model were obtained from two different estimation procedures: 1) SUR without considering the censoring problem (one stage-SUR), and 2) SUR with a two stage estimation method (two stage-SUR).

The two step estimation procedure was applied to the AIDS vegetable demand model using SAS/IML software package. The two stage estimation method consists of a first stage probit estimation and a second stage SUR. The estimation results for the first stage probit model (10) are presented in Table 1. A large number of coefficients are statistically significant at the 5 percent level. Coefficients obtained from the first stage probit regression were used to estimate Φ_i and ϕ_i and the estimated ϕ_i and Φ_i are used in the following second stage estimation. The second stage of the two stage vegetable demand

3 Robust estimation was used for identifying price outliers in this study.

TABLE 1 Probit Estimation Results for Canned and Frozen Vegetables

	Canned Vegetables		Frozen Vegetables	
Variables[a]	Coefficient	t-value	Coefficient	t-value
INTERCEPT	- 0.8964	- 5.66	- 1.2556	- 7.85
FOODEXP	0.0032	2.29	0.0058	4.60
CENCITY	- 0.0363	- 0.56	0.0434	0.67
HHSIZE	0.1246	4.08	0.0034	0.12
HPVEGI	- 0.1058	- 1.89	- 0.1282	- 2.34
BHEAD	0.0041	0.06	0.1366	1.99
COLLEGE	- 0.1576	- 2.72	0.1137	2.01
SPRING88	0.2015	2.85	0.1576	2.24
WINTER	0.4267	5.98	0.1825	2.61
FALL	0.1988	2.43	0.0858	1.05
BCOLLAR	0.4759	0.85	0.1437	- 2.63
FOODSTMP	0.2833	2.45	0.0047	0.04
WHITE	0.2122	2.59	0.3324	3.98
INFOMED	0.1464	1.48	0.3157	3.18
INFOCOM	0.1063	0.92	0.1592	1.39
INFOPROF	0.0509	0.57	0.2455	2.69
INFOREL	- 0.0543	- 0.34	0.0348	0.21
RNVTRFE	0.3117	4.71	- 0.1616	- 2.55
RNVTRVC	- 0.0188	- 1.10	0.0295	1.76
NCENT	0.1448	2.13	- 0.1944	- 2.91
WEST	0.2639	4.03	- 0.0851	- 1.34
Number of Obervations	2653		2653	
Log L	- 1524.5		-1610.47	
R^2	0.0835[b]		0.05713[b]	

[a]The dependent variable for each model was the qualitative dummy variable which was one if canned(and frozen) vegetables were consumed, or zero otherwise.

[b]Squared correlation between observed and predicted values of y.

model (11) was estimated using all observations by SUR.

The results show that the variables included to correct for selectivity bias (PDF) were statistically significant for all budget share equations in the two stage-SUR model, indicating the significance of the omitted variables (table 2). For 62 out of 66 coefficients, the two stage-SUR coefficients were greater in absolute value than the one stage-SUR coefficients. These reflect the

TABLE 2 Estimation Results for Vegetable Demand Model by One Stage-SUR and Two Stage-SUR

Variable[a]	One Stage-SUR			TWO Stage-SUR		
	Canned	Frozen	Fresh	Canned	Frozen	Fresh
INTERCEPT	0.0536	0.0819	0.8645	-0.1676	-0.0274	1.1950
	(2.23)[b]	(4.58)	(32.52)	(-1.58)	(-0.92)	(10.86)
LPCANND	-0.0041	-0.0171	0.0212	-0.0062	-0.0222	0.0284
	(0.56)	(-3.72)	(2.92)	(-0.55)	(-2.39)	(2.47)
LPFROZEN	-0.0171	0.0741	-0.0570	-0.0222	0.2228	-0.2006
	(-3.72)	(13.23)	(9.72)	(-2.39)	(13.65)	(-12.52)
LPRESH	0.0212	-0.0569	0.0357	0.0284	-0.2006	0.1722
	(2.92)	(-9.72)	(3.70)	(2.47)	(-12.52)	(8.50)
LEP	0.0322	-0.0261	-0.0061	0.0492	-0.1027	0.0535
	(4.99)	(-5.44)	(-0.85)	(5.25)	(-7.38)	(3.54)
CENCITY	0.0037	0.0166	-0.0206	0.0168	0.0304	-0.0472
	(0.34)	(1.98)	(-1.65)	(1.01)	(1.31)	(-1.85)
HHSIZE	0.0155	0.0085	-0.024	0.0224	0.0247	-0.0470
	(4.02)	(3.01)	(-5.67)	(2.98)	(3.52)	(-4.91)
HPVEGI	-0.0559	-0.0199	0.0758	-0.0837	-0.0297	0.1134
	(5.81)	(-2.77)	(7.12)	(-5.91)	(-1.44)	(5.05)
BHEAD	-0.0336	0.0273	0.0063	-0.0457	0.0481	-0.0024
	(-2.82)	(3.08)	(0.48)	(-2.64)	(1.70)	(-0.08)
COLLEGE	-0.0252	0.0079	0.0174	-0.0376	0.0067	0.0309
	(-2.52)	(1.06)	(1.57)	(-2.43)	(0.34)	(1.38)
SPRING88	0.0376	0.0082	-0.0458	0.0567	-0.0093	-0.0474
	(3.05)	(0.90)	(-3.37)	(3.06)	(-0.35)	(-1.62)
WINTER	0.1011	-0.0019	-0.0992	0.1483	-0.0299	-0.1184
	(8.32)	(-0.21)	(-7.38)	(6.99)	(-1.15)	(-3.87)
FALL	0.0681	-0.0002	-0.0679	0.1034	-0.0302	-0.0733
	(4.77)	(-0.20)	(-4.30)	(4.84)	(-0.93)	(-2.09)
BCOLLAR	0.0254	-0.0080	-0.0174	0.0438	-0.0011	-0.0427
	(2.66)	(-1.13)	(-1.64)	(3.27)	(-0.05)	(-1.88)

TABLE 2 Continued

Variable[a]	One Stage-SUR			TWO Stage-SUR		
	Canned	Frozen	Fresh	Canned	Frozen	Fresh
FOODSTMP	0.0921	-0.0051	-0.0870	0.1283	-0.0271	-0.1012
	(4.95)	(-0.37)	(-4.23)	(5.08)	(-0.58)	(-2.09)
WHITE	0.0440	0.0345	-0.0786	0.0667	0.0764	-0.1432
	(3.16)	(3.33)	(-5.09)	(3.39)	(2.18)	(-3.84)
INFOPROF	-0.0052	0.0049	0.0003	-0.0108	0.0065	0.0043
	(-0.52)	(0.67)	(0.02)	(-0.76)	(0.32)	(0.20)
INFOCOM	0.0124	0.0101	-0.0226	0.0198	0.0070	-0.0269
	(0.79)	(0.88)	(-1.31)	(0.92)	(0.21)	(-0.75)
RNVTRFE	0.0530	-0.0066	-0.0464	0.0811	0.0022	-0.0833
	(5.63)	(-0.94)	(-4.45)	(4.63)	(0.10)	(-3.18)
RNVTRVC	-0.0215	0.0019	0.0196	-0.0331	0.0068	0.0263
	(-7.02)	(0.83)	(5.79)	(-7.34)	(1.13)	(3.88)
NCENT	0.0522	-0.0372	-0.0151	0.0706	-0.0527	-0.0179
	(4.39)	(-4.17)	(-1.15)	(4.05)	(-2.04)	(-0.64)
WEST	0.0593	-0.0617	0.0024	0.0778	-0.1254	0.0477
	(5.14)	(-6.85)	(0.19)	(4.40)	(-5.05)	(1.75)
PDF				0.6164	0.3524	-0.9688
				(2.77)	(3.27)	(-3.93)
Number of Observations	2653	2653	2653	2653	2653	2653
Mean of Dependent Variable	0.24278	0.09932	0.6579	0.24278	0.09932	0.6579
R^2	0.1154	0.1043	c	0.1122	0.0890	c

NOTE : Blank space indicates variable was not included in the equation.

[a]The dependent variables for each model were vegetble budget shares.

[b]The numbers in parentheses below the coefficients are the t-vales.

[c]R^2-s for the omitted(fresh) vegetable budget share equation are not available.

downward asymptotic bias of OLS in a limited dependent variable model (Greene, 1981). The two stage-SUR coefficients generally appear reasonable with respect to signs (e.g., LEP).

All of the own price coefficients were statistically significant, and most socio-demographic variables, in the canned and fresh vegetable budget share equations, were significant. Among the various socio-demographic variables, household size (HHSIZE), the presence of home grown-vegetables (HPVEGI) and white headed households (WHITE) variables were consistently significant.

The variables associated with the dietary status of the household (RNVTRFE, RNVTRVC) overall had significant impacts on the vegetable budget shares except for the frozen vegetable budget share.

The R^2's in the second stage estimation are around 0.1. The the reason for low R^2's in the cross sectional data analysis is due to the inherent heterogeneity of demand at the household level.

1. Demand Elasticities

The formula used for the final calculation of the Marshallian price elasticities for the two stage-SUR model is

$$E^*_{ij} = -\delta_{ij} + \Phi_i \frac{\gamma_{ir}}{w_i}, \qquad i, j = 1, 2, 3 \tag{12}$$

where Eij* is the Marshallian price elasticities for the two stage-SUR model. The Φ_i were evaluated as the sample mean of the Φ_i (across observations).

All own price elasticities were negative and were generally inelastic, ranging from -0.26 to -1.02 (Table 3). The own price

TABLE 3 Marshallian Price and Expenditure Elasticities for the AIDS Vegetable

Item	Canned Vegetables	Frozen Vegetables	Fresh Vegetables	Expenditure
Canned Vegetables	-1.08181[a] (-1.017)	-0.0631 (-0.0705)	0.08071 (0.0874)	1.1398 (1.1323)
Frozen Vegetables	-0.0746 (-0.1710)	-0.2575 (-0.2592)	-0.6684 (-0.5698)	0.6578 (0.7389)
Fresh Vegetables	0.0433 (0.0323)	-0.3053 (-0.0867)	-0.7379 (-0.9456)	1.08141 (0.9908)

NOTE : Elasticities are evaluted at the sample means. The numbers in parentheses are the elasticities calulated from the one stage-SUR model.

[a]The numbers above the parentheses are the elasticities calculated from the two stag-SUR model.

elasticity for fresh vegetables was more inelastic in the two stage-SUR model relative to that in the one stage-SUR model. The own price elasticity for frozen vegetables was inelastic relative to canned and fresh vegetables. This might be related to the relatively small budget share (0.1) for frozen vegetables compared to the relatively large budget share for fresh vegetables (0.66) and canned vegetables (0.24). The cross price elasticities imply that canned and frozen vegetables, and fresh and frozen vegetables are complements, but fresh and canned vegetables are substitutes in the Marshallian sense.

The expenditure elasticities, as expected, were positive for all vegetable groups. The expenditure elasticity for frozen vegetables was 0.66, implying that the demand for frozen vegetables is inelastic relative to canned (1.14) and fresh vegetables (1.08).

All own price Hicksian elasticities were negative (Table 4). Compared to the uncompensated own price elasticities reported in Table 3 for vegetables, the compensated own price elasticities were more inelastic due to the positive income effect for vegetables. The relation of the compensated and uncompensated price elasticities, given the positive income effects, can easily be seen from the Slutsky equation. In the two stage-SUR model, based on the Hicksian cross price elasticities, frozen

TABLE 4 Hicksian Price Elasticities for The AIDS Vegetable Demand Model : One Stage-SUR vs. Two Stage-SUR Models

Item	Canned Vegetables	Frozen Vegetables	Fresh Vegetables
Canned Vegetables	-0.7406[a] (-0.7418)	0.0508 (0.0428)	0.8295 (0.8313)
Frozen Vegetables	0.0858 (0.6086)	-0.1917 (-0.1852)	-0.2362 (-0.0843)
Fresh Vegetables	0.3060 (0.2730)	-0.1971 (0.0123)	-0.0274 (-0.2946)

NOTE : Elasticities are evaluted at the sample means. The numbers in parentheses are the elasticities calculated from the one stage-SUR model.

[a]The numbers above the parentheses are the elasticities calculated from the two stag-SUR model.

and canned vegetables are substitutes, and fresh and canned vegetables are complements. The concept of substitutes in the Hicksian sense is different from that of substitutes in the Marshallian. Specially this difference arises in that in the calculation of a Hicksian cross price elasticity only the substitution effect is considered, while in calculation of a Marshallian cross price elasticity, the net price effect, taking account of the income effect, is considered. The Marshallian and Hicksian cross price elasticities obtained in this study, except for canned and frozen vegetables, are consistent with the findings of Cox and Wohlgenant (1986). It is difficult to judge the appropriateness of the finding of substitutability or complementarity between the high aggregate vegetable groups analyzed since the basic concepts of substitutability and complementarity are developed intuitively at the individual item level.

2. The Impacts of the Socio-Demographic Variables

If the estimation involves censored regression models such as the standard Tobit or probit model, then the coefficients themselves do not directly provide meaningful interpretations as marginal effects. In this study, the marginal effects are used to examine the impacts of socio-demographic variables on vegetable consumption.

From equation (9), the marginal effects of socio-demographic, nutrition information, and dietary status variables, d_k, can be derived as follows:

$$\begin{aligned}\frac{\partial E(w_i)}{\partial d_k} &= \beta_j \Phi_i + d_k \beta_i \frac{\partial \Phi_i}{\partial d_k} + \sigma_i \frac{\partial \phi_i}{\partial d_k} \\ &= \beta_j \Phi_i + d_k \gamma_j \phi_i (\beta_j - \gamma_j \sigma_i)\end{aligned} \qquad (13)$$

where d_k is socio-demographic variable k (k=1, ..., s) which is a subset of the X variables, γ_k and β_k are the coefficients associated with variable k obtained from the first stage probit and the second stage AIDS vegetable demand model, respectively. For these discrete dummy variables, a marginal effect is interpreted as the difference in the dependent variable value (vegetable budget share) as the category of the dummy variables shifts from the omitted group to the group under consideration, everything else unchanged.

The marginal effects of socio-demographic variables were calculated based on the formula given in equation (13). For the calculation of the marginal effects, Φ_k and ϕ_k for each vegetable group are evaluated for a reference household.[4] The marginal effects converted to dollar terms are presented in Table5. The results show that, ceteris paribus, two person households are also expected to spend more money on canned (4.6 cents) and frozen vegetables (7.5 cents) and less money on fresh vegetables (23.8 cents) relative to one person households.

Households who grew vegetables for home consumption are likely to spend less money on canned (30.5 cents) and frozen vegetables (12.2 cents) and more money (57.0 cents) on fresh vegetables than the households who did not grow vegetables at home. Households with both male head and female heads are likely to spend

TABLE 5 Differences in Vegetable Expenditures between A Selected Household and A Reference Household

Variable	Canned Vegetables	Frozen Vegetables	Fresh Vegetables
		(cents)	
CENCITY	8.23	10.22	-23.71
HHSIZE	4.58	7.47	-23.76
HPVEGI	-30.46	-12.20	56.97
BHEAD	-18.89	17.92	-1.20
COLLEGE	-9.64	4.96	15.54
SPRING88	15.83	1.31	-23.81
WINTER	45.14	-4.02	-59.50
FALL	35.11	-6.80	-36.81
BCOLLAR	16.21	-4.06	-21.44
FOODSTMP	42.19	-7.97	-50.85
WHITE	19.57	31.49	-71.93
INFOPROF	-6.32	8.32	2.18
INFOCOM	4.22	62.30	-13.49
RNVTRFE	21.80	3.53	-41.86
RNVTRVC	-0.67	2.82	13.20
NCENT	23.63	-20.10	-8.99
WEST	22.20	-39.73	23.95

[4] See Choi (1993) for more detail on the reference household.

more money on frozen vegetables (17.9 cents) but spend less on canned (18.9 cents) and fresh vegetables (1.2 cents) than female or male only headed households.

The dietary status of a household (RNVTRFE and RNVTRVC) had opposite impacts on the expected vegetable budget shares. The results showed that a one unit increase in the level of food energy availability (relative to the RDA) would reduce expected fresh vegetable expenditure by 41.9 cents, but would increase expected canned and frozen vegetable expenditures by 3.5 cents and 21.8 cents, respectively.

V. Conclusions

The estimation results for the AIDS vegetable demand model showed that the two stage-SUR method was more appropriate than the one stage-SUR method in terms of the magnitudes and signs of the coefficient estimates. The results indicated that the coefficient estimates for the one stage-SUR model were downward biased due to the existence of selectivity bias in the censored vegetable demand system. The estimation results showed that fresh and canned vegetables, and canned and frozen vegetables can be considered as substitutes in the Hicksian sense. These results are consistent with the findings of previous studies. This study also found that the price elasticities and income elasticities varied across the vegetable groups and in general, were inelastic. The Hicksian compensated own price elasticities were more elastic due to the positive income effect for vegetables. The estimation results showed that many of the socio-demographic variables significantly impacted vegetable consumption and their marginal effect varied across the vegetable groups. The level of food energy availability of the household significantly impacted canned and fresh vegetable expenditures.

This study utilized the recent NFCS data to examine the impact of economic and socio-demographics on vegetable consumption using a censored regression approach. It appears to be the first attempt to estimate a censored demand system for detailed vegetable products. This censored regression approach can be applied to the cross-sectional food demand analysis in Korea.

REFERENCES

Amemiya, T. "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables are Truncated Normal." *Econometrica*, 42(1974).

______. "Tobit Models: A survey." *J. of Econometrics*, 24(1984):3-61.

Choi, J. "Estimation of A vegetable Demand System: A Censored Regression Aproach." Ph.D. Dissertation, Dept. of Ag Econ., Washington State Univ. 1993.

Cox, T.L., and M.K. Wohlgenant. "Prices and Quality Effects in Cross-Sectional Demand Analysis." *Amer. J. of Ag. Econ.*, 68 (1986):908-917.

Cox, T.L., R.F. Ziemer, and J. Chavas. "Household Demand for Fresh Potatoes: A Disaggregated Cross-Sectional Analysis." *Western J. of Ag. Econ.*, 9(1984):40-56.

Deaton, A., and J. Muellbauer. "An Almost Ideal Demand System." *Amer. Econ. Rev.*, 70(1980a):312-326.

______. *Economics and Consumer Behavior*. Cambridge: Cambridge University Press, 1980b.

Green, R., and J.M. Alston. "Elasticities in AIDS Models." *Amer. J. of Ag. Econ.*, 72(1990):442-445.

Greene, W.H. *Econometric Analysis*. New York: Macmillan Publishing Co., 1990.

______. "On the Asymptotic Bias of the Ordinary Least Squares Estimator of the Tobit Model." *Econometrica*, 49(1981):505-513.

Heckman, J. "Sample Selection Bias As a Specification Error." *Econometrica*, 47(1979):153-161.

______. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica*, 42(1974):679-694.

Heckman, J. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Econ. and Social Measurement*, 5(1976):475- 492.

Heien, D., and G. Pompelli. "The Demand for Beef Products: Cross-Section Estimation of Demographic and Economic Effects." *Western J. of Ag. Econ.*, 13(1988):37-44.

Heien, D., and C.R. Wessells. "Demand Systems Estimation with Microdata: A Censored Regression Approach." *J. of Business & Econ. Stat.*, 8(1990):365-371.

_____. "The Demand for Dairy Products: Structure, Prediction, and Decomposition." *Amer. J. of Ag. Econ.*, 70(1988):219-228.

Heien, D., and L.S. Willett. "The Demand for Sweet Spread: Demographic and Economic Effects for Detailed Commodities." *NE. J. of Ag. Res. Econ.*, 15(1986): 160-167.

Lee, L.F. "Simultaneous Equations Models with Discrete and Censored Dependent Variables." *Structural Analysis of Discrete Data with Econometric Applications*, eds. P. Manski and D. McFadden, 346-364. Cambridge, MA: MIT Press, 1978.

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

McCracken, V.A., H. Shi, et al. *The Importance of Socio-Economic, Demographic, Personal and Health-Related Factors and Food Product Attributes on U.S. Nutrient Intake and Implications for the Dietary Status of the Household*, Final Report submitted to U.S. Dept. of Agr., Washington State University, Pullman, 1993.

Phlips, L. *Applied Consumption Analysis*, Amsterdam: North-Holland Publishing Co., 1990.

Pudney, S. *Modelling Individual Choice: The Econometrics of Corners, Kinks, and Holes*. London: Basil Blackwell,1989.

Wales, T.J., and A.D. Woodland. "Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints." *J. of Econometrics*, 21(1983):263-285.