FS-00-07 July 2000

SELECTIVITY BIAS IN THE DEMAND FOR FOOD OF THE HISPANIC COMMUNITY IN THE U.S.

Bruno A. Lanfranco, Glenn W.C. Ames, and Chung L. Huang

FS-00-07 July, 2000

Selectivity Bias in the Demand for Food of the Hispanic Community in the U.S.

Bruno A. Lanfranco, Glenn W.C. Ames, and Chung L. Huang(*)

Paper presented at

American Agricultural Economics Association

Annual Meeting, Tampa, FL

July 30 – August 2, 2000

(*) Graduate Research Assistant, Professors respectively, Department of Agricultural and Applied Economics, The University of Georgia. 315 Conner Hall. Athens, GA. 30602-7509; Ph:(706)542-0757; Fax:(706)542-0739; email: games@agecon.uga.edu

Dept. of Agricultural & Applied Economics
College of Agricultural & Environmental Sciences
University of Georgia

SELECTIVITY BIAS IN THE DEMAND FOR FOOD OF THE HISPANIC COMMUNITY IN THE U.S.

Bruno A. Lanfranco, Glenn W.C. Ames, and Chung L. Huang

Department of Agricultural and Applied Economics

University of Georgia

Athens, GA 30602-7509

games@agecon.uga.edu

ABSTRACT

The presence of selectivity bias in the demand for food was examined for a sample of U.S. Hispanic households. Engel curves for ten food groups were estimated correcting for selectivity bias. Statistical evidence of selectivity bias was found in eight of the ten food groups analyzed in this study.

Key words: Selectivity bias, Engel curves, food demand, Hispanic population

SELECTIVITY BIAS IN THE DEMAND FOR FOOD OF THE HISPANIC COMMUNITY IN THE U.S.

by

Bruno A. Lanfranco, Glenn C. W. Ames, and Chung L. Huang (*)

American Agricultural Economics Association

Annual Meeting, Tampa, FL

July 30 – August 2, 2000

(*) Graduate Research Assistant, Professors respectively, Department of Agricultural and Applied Economics, The University of Georgia. 315 Conner Hall. Athens, GA. 30602-7509; Ph:(706)542-0757; Fax:(706)542-0739; email: games@agecon.uga.edu

SELECTIVITY BIAS IN THE DEMAND FOR FOOD OF THE HISPANIC COMMUNITY IN THE U.S.

Abstract
The presence of selectivity bias in the demand for food was examined for a sample of U.S. Hispanic households. Engel curves for ten food groups were estimated correcting for selectivity bias. Statistical evidence of selectivity bias was found in eight of the ten food groups analyzed in this study.

Key words: Selectivity bias, Engel curves, food demand, Hispanic population

Introduction

A problem that commonly arises in the estimation of food demand patterns is that the dependent variable can take values that are limited in some way. For example, some households can report zero consumption during the week they are surveyed, for one or more specific food groups, such as beef, vegetables, or fruits. In this case there could be a sort of selectivity bias and estimation of this model by least squares gives inconsistent estimates of the parameters (Maddala, 1983). In addition, it has been argued that the determinants of the decision whether to consume from a particular food group are often not the same as the determinants of how much to consume, in particular when we refer to highly specified food groups (Haines, Guilkey, and Popkin, 1988). As a consequence, ignoring the two-step decision process would miss the true behavioral patterns, leading to erroneous results in the estimation process. To address this problem, various estimation methods based on a two-step decision process have been utilized.

Heckman's two-step procedure has been widely used to take into account these two problems (Leung and Yu, 1996). However, Johnston and Dinardo (1997) pointed out that there is no consensus among analysts on the value of selectivity bias correction methods or when their use is appropriate. Kennedy (1998) refers to Heckman's two-step procedure as a second best alternative to the full information maximum likelihood (FIML) approach. Moreover, some authors recommend the use of the two-step procedure only to test for the presence of selectivity bias and, in case it is detected, they consider that the FIML should be used (Davidson and MacKinnon, 1993).

Objectives

The primary objective of this paper is to determine the presence of selectivity bias in the demand for food among a sample of the U.S. Hispanic population, the fastest growing population in the U.S. The correct estimation of food consumption patterns for Hispanic consumers could have important implications for the demand for food in the U.S. and the response of food processors.

Estimating Models with Selectivity Bias in a Two-step Food Consumption Process

Haines, Guilkey, and Popkin (1988) questioned the use of one-step decision methods for examining food consumption decisions, such as the Tobit model, where for a certain food the decision to consume and the decision about the amount to consume are the same. They argued that food consumption decisions should be regarded as a two-step decision process. People first decide whether or not to consume from a particular food group and then the quantity to consume. Moreover, the two-step process implies that two dependent variables are analyzed (Guilkey, Haines, and Popkin, 1990): (a) a dichotomous variable that indicates whether or not an individual consumes a nonzero amount from a particular food group, and (b) the actual quantity consumed for those who chose to consume. Consequently, a model assuming a two-step decision process for consumption of the i^{th} person can be described by the following equations:

$$y_i^* = \gamma_0 + \gamma_1 z_i + e_i^* \qquad \textit{Dichotomous or Decision Equation} \qquad (1)$$

$$q_i^* = \beta_0 + \beta_1 x_i + u_i^* \qquad \textit{Regression or Level Equation} \qquad (2)$$

The dependent variable y_i^* in equation (1) is a reservation value, and it is unobserved. Instead, we observe the binary realization y_i, which takes the value $y_i = 1$ (yes) when $y_i^* > 0$, and $y_i = 0$ (no) when $y_{1i}^* \leq 0$. The dependent variable in equation (2) contains the consumption information of those individuals for which the realization variable $y_i = 1$ (yes), that is $q_i = q_i^*$ when $y_i^* > 0$, being otherwise their information unobservable ($q_i = 0$). The vectors z_i and x_i represent the regressors included in the decision and level equation, respectively.

Estimating equation (1) by OLS, as a linear probability model, has two major weaknesses: (a) it does not constrain the predicted value to lie between zero and one, as expected; and (b) it is *heteroscedastic*, which violates one of the assumptions of the Classical Linear Regression (CLR) model (Johnston and Dinardo, 1997). In this case the error term is assumed to be normally distributed such that $e_i^* \sim N(\mu, \sigma_e^2)$. Thus, although the unobservable variable y_i^* is distributed normally, its realization y_i is not. Equation (1) can be regarded as a simple binary probit (BP), so it is also called the *probit equation*, and it can be estimated using Maximum Likelihood (MLE) method. Dropping subscript 1 for simplicity, it can be shown that,

$$\Pr(y_i = 1) = \Pr(y_i^* > 0) = \Pr\left(\frac{e_i^*}{\sigma_e} > -z_i'\frac{\gamma}{\sigma_e}\right)\Pr\left(\frac{e_i^*}{\sigma_e} < z_i'\frac{\gamma}{\sigma_e}\right) = \Phi\left(z_i'\frac{\gamma}{\sigma_e}\right), \tag{3}$$

where Φ is the standard normal cumulative distribution function (cdf). It follows that,

$$\Pr(y_i = 0) = 1 - \Pr(y_i = 1) = 1 - \Phi\left(z_i'\frac{\gamma}{\sigma_e}\right). \tag{4}$$

Under the assumption of independent identical distributed sampling (IID), the corresponding likelihood function can be derived as the product of the probability of each observation, as shown in Johnston and Dinardo (1997, p. 420). The likelihood function, denoted by L is then:

$$L = \prod_{i=1}^{T} \Phi\left(z_i' \frac{\gamma}{\sigma_e} \right)^{yi} \left[1 - \Phi\left(z_i' \frac{\gamma}{\sigma_e} \right) \right]^{1-yi} . \tag{5}$$

The parameters γ and σ_e go together, implying that the numerical scale of the latent variable is unobservable (γ and σ_e are not separately identified). The standard deviation of the disturbance term, σ_e, can be normalized to one, to be able to get γ. Taking logs, we obtain the following probit log-likelihood function, which is the specification used to estimate equation (1) by MLE.

$$\ln L = \sum_{i=1}^{T} \{ y_i \ln[\Phi(z_i'\gamma)] + (1 - y_i) \ln[1 - \Phi(z_i'\gamma)] \} \tag{6}$$

This likelihood function is globally concave in γ. Local and global maxima will be the same and the Newton-Raphson estimation method provides a straightforward estimation method (Amemiya 1994, p. 335). The correct specification of the likelihood function means that we have the asymptotic properties of MLE: consistency, asymptotic efficiency and asymptotic normality (Johnston and Dinardo 1997, p. 143).

Concerning to the level equation, at least in principle, u_i^* is not necessarily normally distributed, and up to this moment it does not involve any sample selection or selectivity bias. Thus, equation (2) could be estimated by OLS using only the observations for which $y_i = 1$ (yes), making corrections for heteroscedasticity, when necessary.

When the level of consumption, given any, is conditionally independent of the decision of consumption, the model is known in the literature as a *Two-Part* model (TP), as pointed out by Leung and Yu (1996). While this approach looks appealing to model two-step decisions where the level stage can be expressed by means of a conditionally independent equation, it does not seem appropriate to represent situations where selectivity bias is present. Following Davidson and MacKinnon (1993, p. 542), in many practical cases the truncation is based not on the value of the dependent variable but rather on the value of another variable that is correlated with it. If the dependent value in equation (2) is correlated with some reservation value described by equation (1) there may be a problem of selection sample often referred as *sample selectivity bias*.

In this case, another family of models called generically as *Sample Selection* models (SS) can be used. In the SS models, equation (2) is no more a conditional equation, but an unconditional one. Now, the disturbances of equations (1) and (2) are assumed to be correlated through a correlation coefficient ρ (rho), and following a bivariate normal distribution:

$$\left(e^*, u^*\right) \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{eu} \\ \sigma_{ue} & \sigma^2 \end{pmatrix}\right], \tag{7}$$

where $\sigma_e^2 = 1$ for identification purposes, so we can call $\sigma_u^2 = \sigma^2$ for simplicity. This extension provided by SS models is precisely the separation of the decision and the level process into two different but related equations. Amemiya (1985, p. 385) calls this sample selection model as a Type II Tobit model, which is a generalization of the original Tobit, widely used to estimate models with censored or truncated samples (Kennedy, 1998). The SS model connects the probit equation (1) that stands for the decision problem with the original regression equation (2) through the possible correlation between their disturbances, assumed to be joint normally distributed as denoted in (7).

There are some alternative ways to deal with this new situation. Holcomb, Park and Capps (1995) chose the two-step Heckman procedure to circumvent this problem because they considered this method to be less restrictive than the Tobit estimation technique and easy to implement. Heckman (1976; 1979) noted that when there exists self-selectivity, there is an omitted variable bias in the OLS estimates of the regression equation, with a magnitude given by the so-called inverse Mills ratio. If this omitted variable were included in the regression, then OLS is consistent. The two steps are as follows: first, run the probit model in (1), which deals with the decision problem and compute the inverse Mills ratio, as:

$$\hat{\lambda}_i = \lambda(z_i'\hat{\gamma}) = \frac{\phi(z_i'\hat{\gamma}/\sigma_e)}{\Phi(z_i'\hat{\gamma}/\sigma_e)}, \tag{8}$$

with $\sigma_e = 1$ as noted before, and ϕ is the standard normal probability distribution function (pdf). Second, estimate the model by OLS using equation (2), including the estimated inverse Mills ratio as one of the regressors (Heckman, 1979). The coefficient associated with λ_i in this regression is $\rho\sigma$, where ρ (rho) and σ (sigma) are defined as before. Since σ is always positive by definition, if this coefficient is significantly different from zero, it means that ρ is also significantly different from zero. In the same way, if this coefficient is not significantly different from zero, then ρ is not statistically significant.

Although the use of the Heckman's two-step procedure gained wide popularity, Johnston and Dinardo (1997) remark that there is no consensus among analysts on the value of selectivity bias methods and when their use is appropriate. Kennedy (1998) refers to Heckman's two-step procedure as a second best alternative to maximum likelihood. Davidson and MacKinnon (1993) suggest the use of a Full Information Maximum Likelihood method (FIML) which makes use of all the information about the covariance between the residuals of the probit and the regression equations, giving more efficient estimates. They also recommend the use of the two-step procedure only to test for the presence of selectivity bias, and in case it is detected, they suggest that the FIML should be used. In this case, equations (1) and (2) are estimated simultaneously.

This method, based on the maximization of a likelihood function, uses starting values from probit and OLS that in practice give better results than the Heckman procedure (TSP 4.5 Reference Manual, 1999, p. 239). The likelihood function of the SS model can be stated as:

$$L = \prod_0 f\left(q_i = 0, y_i^* \le 0\right) \prod_1 f\left(q_i = q_i^*, y_i^* > 0\right). \qquad (9)$$

Following Amemiya (1985), equation (9) can be rewritten as:

$$L = \prod_0 \Pr\left(q_i = 0 \mid y_i^* \le 0\right) \Pr\left(y_i^* \le 0\right) \prod_1 f\left(q_i \mid y_i^*\right) \Pr\left(y_i^* > 0\right)$$

$$= \prod_0 \Pr\left(y_i^* \le 0\right) \prod_1 \int_0^\infty f\left(y_i^*, q_i\right) . dy_i^*$$

$$= \prod_0 \Pr\left(y_i^* \le 0\right) \prod_1 \int_0^\infty f\left(y_i^* \mid q_i\right) f\left(q_i\right) . dy_i^*$$

It can be noted that expression $f(.,.)$ is the joint density of y_i^* and q_i, so it can be written as the product of a conditional density and a marginal density. Also, the conditional distribution of y_i^* given $q_i^* = q_i$ is normal with mean $z_i'\gamma + \sigma_{eu}\sigma_u^{-2}(q_i - x_i'\beta)$ and variance $(\sigma_e^2 - \sigma_{eu}^2\sigma_u^{-2})$. Thus,

$$L = \prod_0 \left[1 - \Phi\left(z_i'\gamma / \sigma_e\right)\right]$$
$$\times \prod_1 \Phi\left\{\left[z_i'\gamma / \sigma_e + \sigma_{eu}\left(q_i - x_i'\beta\right) / \sigma_e\sigma_u^2\right] \cdot \left[1 - \sigma_{eu}^2 / \sigma_e^2\sigma_u^2\right]\right\} \phi\left[\left(q_i - x_i'\beta\right) / \sigma_u\right] / \sigma_u$$

Defining the correlation coefficient *rho* as $\rho = \sigma_{eu}/\sigma_e.\sigma_u$ and letting $\sigma_e = 1$ and $\sigma_u = \sigma$ as before, we can take logs to write the final expression of the likelihood function used in the estimation with MLE:

$$\ln L_i = (1 - y_i)\ln \Phi(-z_i'\gamma) + y_i\left\{-\ln(\sigma) + \ln\phi\left(x_i'\frac{\beta}{\sigma}\right) + \ln\Phi\left[\left(z_i\gamma + \rho x_i'\frac{\beta}{\sigma}\right)(1-\rho^2)^{-0.5}\right]\right\} \qquad (10)$$

Leung and Yu (1996) pointed out that a fundamental distinction between the SS and the TP models lies in the assumptions on the error terms of equations (1) and (2). The SS model assume that $E(u_i^*) = 0$, so that $E(u_i^*|y_i^*>0) = \rho\sigma\lambda_i(z_i'\gamma)$. On the other hand, the TP model assumes that $E(u_i^*|y_i^*>0) = 0$. This difference is the core of the vigorous debate between the advocates of one model and the other. Leung and Yu (1996) considered that although the SS models have been vigorously criticized in favor of the TP models, the merits of the latter have been grossly exaggerated in the literature. They offer a more balanced account for the merits of the two models. "Our results do not support the contention that the two-part models dominate the sample selection model, nor do we find that the sample selection model is superior to the two-part model" (Leung and Yu 1996, p. 200).

A major weakness of the sample selection model is that it is sometimes affected by collinearity problems. In that sense, Leung and Yu (1996, p. 201) explain that "models with few exclusion restrictions, a high degree of censoring and a low variability among the regressors, or a large error variance in the choice equation, can all contribute to near collinearity between the regressors and the inverse Mills ratio, rendering the two-step estimator ineffective." In addition, TSP International reports in its software manual that some problems appear to occur when the probit equation dominates the likelihood function. In this case the intercepts become distorted, the estimated correlation coefficient of the residuals is slightly less than one in absolute value, and the residual covariance matrix is nearly singular. The standard error or the correlation coefficient and its covariance with other parameters is set to zero, and in these cases it is not clear how to interpret the model (TSP 4.5 Reference Manual, 1999. p.257). The following analysis tests for the presence of selectivity bias in the demand for food among the U.S. Hispanic population.

Data Set and Methodology

The cross-sectional data set used in this research was constructed from USDA's 1994-96 *Continuing Survey of Food Intakes by Individuals* (CSFII94-96). Only households of Hispanic origin that participated in the 1994-96 two-day survey and provided information about food consumption were selected for analysis. The total sample consisted of 643 households. Ten main food groups were considered in the analysis: grains, vegetables, fruits, milk, beef, pork, fats, sugar, legumes, and beverages. Engel curves for all the groups were estimated using both a double-logarithmic and a semi-logarithmic functional form.

Several demographic and socio-economic variables that were hypothesized to influence the demand for food were also included in the analysis of Hispanic consumers. Table 1 provides a complete description of the variables used in the estimation of the models.

Heckman's two-step procedure (HP) and Amemiya's Type II Tobit, also known as Sample Selection (SS) were used in this paper to test for sample selection bias. If there is selectivity bias, the disturbance terms of the decision and level equations are correlated through the correlation coefficient ρ (rho). We assumed that under the null hypothesis, there was no selectivity bias in the demand of Hispanic households for a specific food group (H_o: $\rho = 0$). The alternative hypothesis, on the other hand, recognizes the presence of selectivity bias (H_a: $\rho \neq 0$).

As it was discussed earlier, the coefficient associated with the inverse Mills ratio (λ_i) included in the level equation for the HP method is $\rho\sigma$. If this parameter is significantly different from zero, it means that the correlation coefficient ρ (rho) is also significantly different from zero. This, the null hypothesis for each food group was rejected if this parameter was significantly different from zero.

An equivalent test was performed using the SS method. In this case, the correlation coefficient ρ (rho) is estimated directly from the likelihood function, so that the statistical significance of the ρ (rho) led to the rejection of the null hypothesis for a specific food group. In both cases we chose the 10% level to establish the rejection region.

Results and Discussion

The estimated regression coefficients for all the food categories using the double-logarithmic (DL) and the semi-logarithmic (SL) model, estimated by both the HP and SS procedures, are presented in Tables 2 to 7. In all cases the coefficient estimates correspond to the level equation only. Coefficient estimates from the decision equation as well as the estimated income and household size elasticities are not presented here but they are available from the authors upon request. In general, the regression coefficients of household size indicated more statistical significance in the level equation than the estimated income coefficients, for most of the food groups. Household income appeared to have more effect on the decision equation.

Statistical evidence of selectivity bias in the demand for food was found for most of the food groups, sometimes in association with the specific functional form used in the estimation and in the estimation method used. The coefficient of the inverse Mills ratio used as a regressor in the HP method was statistically significant for grains and milk, regardless to the functional form, allowing for the rejection of the null hypothesis. In the case of beef, this coefficient was statistically significant only when using the SL model, whereas for sugar and beverages it was significantly different from zero when estimated using the DL model.

On the other hand, the correlation coefficient ρ (rho) obtained with the SS method was statistically different from zero for fruits, beef, pork, sugar, and legumes, although only when the

DL model was used instead of the SL model. Under this circumstance, the null hypothesis was rejected in favor of the alternative hypothesis, and the SS method provided evidence of selectivity bias in the demand for food. In addition to this, the SS estimation method provided more efficient coefficient estimates than the HP procedure (Davidson and MacKinnon, 1993). In all the other cases, the estimated correlation coefficient was slightly less than one and the corresponding residual covariance matrix was nearly singular. As discussed earlier, this near-singularity occurs when the probit equation strongly dominates the likelihood function, and consequently, the interpretation of the results is not clear.

A summary of the results of this paper is presented in Table 7. For each food group, the numbers in the second and third column show respectively, the total quantity and the percentage of households with positive observations; that is, those observations for which variable $y_i = 1$ (yes) in the probit equation. In other words, each of these numbers indicate how many households consumed a positive amount of the specific food during the survey period, so that $q_i = q_i^*$ in the level equation, as well as the percentage they represent in the total quantity of households in the sample. For example, 517 out of 643 of the Hispanic households (80.4%) reported that some positive amount of legumes was consumed by the household during the period of the survey. The next two columns indicate if any statistical evidence of selectivity bias was found, using both the DL and the SL functional form tested with the HP method. The last two columns show the same information when the SS method was used to test for the presence of selectivity bias.

The collinearity problems, as discussed earlier, prevented the SS method from producing any useful result with the SL model. The SS method was incapable of providing any evidence to accept or reject the null hypothesis of no selectivity bias with this model. On the other hand, we observed that the indeterminacy in the results of the test performed with the SS method, when the DL functional form was used, appeared to have some relationship with the proportion of households with positive consumption. With the noteworthy exception of total fats, the SS estimation method performed satisfactorily well with all the food groups in which the percentage of positive observations in the level equation was less than 90%. In those cases, not only the SS model made a good estimation of all the parameters of the likelihood function, but also showed statistical evidence of selectivity bias.

With regard to the HP method, we observed that the only two food groups, indicating selectivity bias with both functional forms, were those having the highest proportion of households consuming positive amounts of the good: grains (99.7%) and milk (97.5%). The remaining food groups did not show any distinctive behavior related to the percentage of positive observations in the level equation.

Conclusions

The results derived from this study suggest that selectivity bias is an important issue to be considered when modeling food consumption patterns of U.S. Hispanic households, at least for some specific food groups. Statistical evidence of the presence of selectivity bias was found for most of the food groups considered in this study: grains, fruits, milk, beef, pork, sugar, legumes, and beverages. It should be noted that beef and sugar were the only two food groups for which evidence of selectivity bias was found using both the HP and the SS estimation methods. On the other hand, only vegetables and fats did not show any evidence of selectivity bias at all.

As pointed out by many researchers, when selectivity bias is detected in the data, the use of estimation methods that consider this issue is more appropriate than methods where the regression or level equation is formulated by means of a conditional equation. Although considered by many researchers as a better estimation technique, when compared to the HP, the SS method appears to be very sensitive to the type of data set utilized. The SS method performed poorly in many of the cases considered in this paper, particularly with the SL model, due to collinearity problems between the inverse Mills ratio and the other regressors. This kind of problem appears when the probit equation dominates the likelihood function, and in this case it is not clear how to interpret the model. In all the cases we could not reject the null hypothesis of no selectivity bias using this method, it was due solely to the collinearity problems we faced with this particular data set and not due to the statistical insignificance of the correlation coefficient between the disturbance terms.

References

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

Amemiya, T. *Introduction to Statistics and Econometrics*. Cambridge, MA: Harvard University Press, 1994.

Davidson, R., and MacKinnon, J. *Estimation and Inference in Econometrics*. New York: Oxford University Press,, 1993.

Guilkey, D.K., Haines, P.S., and Popkin, B.M. "The Distribution of Food Consumption over a year: A Longitudinal Analysis." *Amer. J. Agr. Econ. 72* (November 1990):891-162.

Haines, P.S., Guilkey, D.K., and Popkin, B.M. "Modeling Food Consumption Decisions as a Two-Step Process." *Amer. J. Agr. Econ. 70* (August 1988):543-552.

Heckman, J.J. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1979):153-162.

Heckman, J.J. "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimation for Such Models." *Annals of Economic and Social Measurements*. 5 (1976):475-592.

Holcomb, R.B., Park, J.L., and Capps, O., Jr. "Revisiting Engel's law: Examining Expenditure Patterns for Food at Home and Away from Home." *Journal of Food Distribution Research* (September 1995):1-8.

Johnston, J., and Dinardo, J. *Econometric Methods.* Fourth Edition. New York: McGraw-Hill, 1997.

Kennedy, P. *A Guide to Econometrics*. Cambridge, MA: The MIT Press, 1998.

Leung, S.F., and Yu, S. "On the Choice between Sample Selection and Two-Part Models." *Journal of Econometrics*, 72 (1996):197-229.

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

TSP 4.5. TSP Reference Manual, TSP International, Version 4.5, 1999.

U.S Department of Agriculture. CSFII/DHKS 1994-96. *Continuing Survey of Food Intakes by Individuals and Diet and Health Knowledge Survey*. CD-ROM. Food Surveys Research Group, Beltsville Human Nutrition Research Center, Agricultural Research Center. Washington, D.C., 1998.

Table 1 - Variables used in the Estimation of Expenditure Patterns of Hispanic Population

Variables	Name	Description of the variable
Weekly Income	LINCWK	Natural logarithm of total before-tax income of the household, in dollars per week.
Household Size	LHHSIZE	Natural logarithm of household size using the Amsterdam Scale, in adult equivalents.
Age of Household Head	AGE	Natural logarithm of age of the head or the reference person of the household, in years.
Sex (Binary)	S_FEM	Sex of the household head (1-Female; 0-Male).
National Origin (Binary) (default: Other Spanish/Hispanics)	O_MEX	Household members identified as Mexican, Mexican-American, or Chicano.
	O_PRI	Household members identified as of Puerto Rican origin.
	O_CUB	Household members identified as of Cuban origin.
Geographic Region (Binary) (default: West Region)	R_NEAST	Households located in the Northeast region of the U.S.
	R_MWEST	Households located in the Midwest region of the U.S.
	R_SOUTH	Households located in the South region of the U.S.
Urbanization Status (Binary) (default: outside MSA)	U_MSAINC	Households located in central city, inside the Metropolitan Statistical Area (MSA).
	U_MSAOUT	Households located outside the Central City, inside the Metropolitan Statistical Area (MSA).
Education (Binary) (default: no formal education)	G_ELEM	Household head completed or attended one or more years of elementary school.
	G_HIGH	Household head with some years, or completed high school, or a General Education Degree.
	G_COLL	Household head with one to four of college education.
	G_GRAD	Household head with five or more years of college.
Tenure Status (Binary)	F_STAMP	Tenure status of the household dwelling (1-Owner; 0-Other).
Food Stamps (Binary)	T_OWNER	Any household member receiving benefits Food Stamps (1-Yes; 0-No).
WIC Program (Binary)	W_YES	Any household member under the Women, Infant and Children program (1-Yes; 0-No).
Year of the Survey (Binary) (default: Year 1994)	Y_95	Household surveyed in 1995 (1-Yes; 0-No).
	Y_96	Household surveyed in 1996 (1-Yes; 0-No).
Inverse Mills Ratio	I. Mills R.	Included as a regressor in the level equation for the two-step Heckman's procedure.
Additional Variables (Sample Selection method)	SIGMA	Standard deviation of the error term of the regression or level equation.
	RHO	Correlation coefficient between the disturbances of the decision (probit) and level equations.

Table 2. Parameter Estimates of Food Consumption for Hispanic Households in the U.S., 1994-96

Variables	GRAINS				VEGETABLES			
	Double Logarithmic model		Semi Logarithmic model		Double Logarithmic model		Semi Logarithmic model	
	HP	SS	HP	SS	HP	SS	HP	SS
Constant	0.916016	3.99260***	-2146.36***	-810.561*	0.623384	3.52411***	-2380.77	-16.5073
	(1.32945)	(0.696735)	(756.789)	(454.534)	(5.73497)	(0.887234)	(1777.84)	(231.200)
LINCWK	1.03855***	0.090978	437.747**	35.4760	0.335741	0.072920	251.897	43.9440**
	(0.323989)	(0.067870)	(171.356)	(44.0775)	(0.542640)	(0.086729)	(167.151)	(22.1312)
LHHSIZE	-1.62204**	0.521542***	-502.383	408.236***	0.546529	0.360100***	300.759**	124.335***
	(0.727296)	(0.092613)	(405.546)	(60.1663)	(0.404545)	(0.124791)	(131.105)	(30.9809)
LAGE	2.01232***	0.090047	960.505**	141.443*	0.217056	0.215356	-39.6299	-12.6744
	(0.662534)	(0.123194)	(386.439)	(80.4463)	(0.176447)	(0.164192)	(53.8145)	(39.8463)
S_FEM	0.139226*	0.105552	75.6727	67.5940	0.055268	0.036976	26.5514	4.64033
	(0.080696)	(0.079650)	(57.1912)	(51.9074)	(0.095813)	(0.095306)	(27.4888)	(17.0647)
O_MEX	0.070659	0.050525	25.7997	28.1843	0.093929	0.041470	18.7211	12.1162
	(0.088416)	(0.073211)	(52.7718)	(56.8011)	(0.115805)	(0.088915)	(31.0777)	(22.8138)
O_PRI	0.017510	0.018263	-40.4433	-21.5342	-0.288330*	-0.304624**	-107.862***	-79.4264*
	(0.136181)	(0.137685)	(89.4049)	(89.3122)	(0.153812)	(0.129945)	(38.5992)	(47.0071)
O_CUB	0.275562	0.227200	136.939	130.585	0.334242	0.318625**	252.013	72.1028**
	(0.170812)	(0.238257)	(164.921)	(154.809)	(0.304899)	(0.160818)	(200.214)	(28.2660)
R_NEAST	0.284721**	0.325246***	149.331	156.235*	0.469381***	0.367760***	83.5631*	68.8342**
	(0.124335)	(0.121376)	(97.9935)	(84.1165)	(0.148227)	(0.134115)	(44.7242)	(32.2064)
R_MWEST	-0.093562	-0.071837	-87.0817	-71.4815	0.155838	0.160171	36.3322	48.6508**
	(0.153931)	(0.147412)	(76.5272)	(95.6624)	(0.195110)	(0.163710)	(47.1080)	(23.7229)
R_SOUTH	-0.079727	-0.065641	-75.0416	-61.8853	0.264157**	0.206214	26.3859	30.0345
	(0.092811)	(0.090423)	(50.1748)	(59.8276)	(0.108895)	(0.136541)	(32.3233)	(24.6560)
U_MSAINC	0.222615*	0.187223*	144.272**	115.365	-0.203577	-0.213760**	-13.0881	-44.5684
	(0.117506)	(0.112903)	(62.2586)	(74.0275)	(0.134599)	(0.108079)	(36.8068)	(29.1144)
U_MSAOUT	0.245915**	0.238640**	174.971***	161.904**	-0.109055	-0.200505**	-1.85182	-17.9314
	(0.112243)	(0.104941)	(53.1525)	(68.4778)	(0.120341)	(0.100179)	(35.3886)	(31.4321)
G_ELEM	0.388641	0.364672	134.187	121.165	0.049940	0.134557	-8.02027	-9.25787
	(0.423453)	(0.338241)	(153.716)	(219.551)	(0.430017)	(0.457211)	(132.946)	(127.312)
G_HIGH	0.612555	0.597896*	275.002*	260.085	0.131376	0.234837	5.80557	-13.4230
	(0.423670)	(0.338110)	(159.458)	(219.423)	(0.426136)	(0.451842)	(132.119)	(126.139)
G_COLL	0.788516*	0.784254**	385.778**	375.673*	0.171565	0.279348	6.32230	-12.9990
	(0.426863)	(0.343424)	(160.258)	(224.155)	(0.434661)	(0.455312)	(134.842)	(128.237)
G_GRAD	0.620381	0.603334	319.306*	313.310	0.111847	0.125163	-34.7865	-25.7594
	(0.454535)	(0.370421)	(181.245)	(240.686)	(0.474713)	(0.493410)	(142.275)	(131.378)
T_OWNER	0.065034	0.042223	-54.5989	-62.9079	-0.040608	0.027132	-11.8844	-18.6236
	(0.091769)	(0.080765)	(57.6595)	(57.0855)	(0.103564)	(0.112969)	(27.9172)	(23.4577)
F_STAMP	0.179175	0.144656	64.9564	48.7397	0.203256*	0.231509*	21.5410	20.2386
	(0.111161)	(0.105771)	(80.4304)	(68.7812)	(0.111882)	(0.127942)	(37.6994)	(27.2646)
W_YES	-0.065309	-0.102075	108.240	79.0721	0.024425	0.008794	-4.14194	-10.6133
	(0.130468)	(0.098015)	(96.8824)	(68.7259)	(0.133556)	(0.132913)	(40.4764)	(25.3008)
Y_95	-0.257912***	-0.251454***	-103.347*	-101.066*	-0.088595	0.013862	-14.7978	7.70188
	(0.086810)	(0.087218)	(59.3587)	(58.5674)	(0.112193)	(0.117131)	(29.9413)	(27.0941)
Y_96	-0.136214	-0.143564	-7.30553	-19.6238	0.141003	0.178637	58.8969	40.2094*
	(0.095705)	(0.093004)	(62.2172)	(60.2350)	(0.114625)	(0.108555)	(36.3851)	(23.5830)
I. Mills R.	-789.277**	-	-333421.**	-	13.8452	-	12485.7	-
	(259.875)		(144472.)		(30.3539)		(9441.85)	
SIGMA	-	0.930146***	-	603.592***	-	1.17250***	-	328.595***
		(0.025223)		(16.6924)		(0.038900)		(9.27312)
RHO	-	1.000000	-	1.000000	-	1.000000	-	1.000000
		(~0.00000)		(~0.00000)		(~0.00000)		(~0.00000)

Note: Std. errors in parenthesis. Significance of 2-tail t-test: *** - 1% level; ** - 5% level; * - 10% level.

Table 3 - Parameter Estimates of Food Consumption for Hispanic Households in the U.S., 1994-96

Variables	FRUITS				MILK			
	Double Logarithmic model		Semi Logarithmic model		Double Logarithmic model		Semi Logarithmic model	
	HP	SS	HP	SS	HP	SS	HP	SS
Constant	5.22290***	7.07270***	363.124	581.478*	7.73595***	8.22627***	1369.05***	1691.70***
	(1.57736)	(0.825361)	(792.167)	(302.701)	(0.719999)	(0.807156)	(528.785)	(500.222)
LINCWK	0.068082	-0.106462	26.6966	15.4447	0.212330	-0.083789	179.212*	-2.08089
	(0.151008)	(0.079736)	(75.8886)	(33.4299)	(0.152864)	(0.078602)	(94.7608)	(46.1162)
LHHSIZE	0.719105*	0.130801	325.593	246.695***	1.03699***	0.757884***	690.228***	490.649***
	(0.727296)	(0.114306)	(217.996)	(51.4621)	(0.172641)	(0.109158)	(102.810)	(62.4965)
LAGE	-0.298794*	-0.169796	-158.249*	-113.700*	-1.16448***	-0.587555***	-760.821***	-368.799***
	(0.174965)	(0.149713)	(89.9148)	(66.5232)	(0.328645)	(0.141220)	(179.426)	(80.4463)
S_FEM	-0.071715	-0.025810	-25.5351	-20.8900	0.097778	0.039286	0.827190	4.34871
	(0.093218)	(0.084510)	(44.7602)	(36.6288)	(0.089267)	(0.092434)	(56.0814)	(49.3386)
O_MEX	-0.006143	-0.002872	15.9921	-6.93222	0.006867	-0.010202	87.1821	57.9951
	(0.098248)	(0.093409)	(51.0629)	(29.0603)	(0.100204)	(0.093636)	(60.2052)	(47.5587)
O_PRI	-0.130882	-0.165407	-46.6093	-56.3258	-0.460900**	-0.367851**	-136.356*	-104.874
	(0.149768)	(0.148527)	(64.4251)	(63.8855)	(0.178953)	(0.147476)	(75.5738)	(99.7771)
O_CUB	0.002715	0.063619	106.749	-100.493	0.113858	0.051850	174.419	54.2901
	(0.398653)	(0.227526)	(153.056)	(98.3956)	(0.276007)	(0.289065)	(194.174)	(135.086)
R_NEAST	0.211987	0.128785	45.7633	37.9757	0.333589**	0.233956	46.9835	45.0035
	(0.142771)	(0.137794)	(64.4303)	(66.5911)	(0.132025)	(0.150997)	(81.4063)	(66.2194)
R_MWEST	0.284847**	0.224199	51.8259	64.3986	-0.247096	-0.174732	-117.459	-67.6416
	(0.143435)	(0.158759)	(76.5941)	(41.0844)	(0.198253)	(0.166069)	(96.0409)	(107.128)
R_SOUTH	0.016715	0.017156	-20.9607	-0.801004	-0.090887	-0.062537	-155.930*	-63.1472
	(0.106113)	(0.100134)	(51.1231)	(33.9422)	(0.114243)	(0.099598)	(60.7279)	(66.8930)
U_MSAINC	0.177818	0.208145*	102.657*	48.3115	0.286934**	0.299903**	143.641**	142.073*
	(0.123407)	(0.118192)	(62.2586)	(49.7243)	(0.126820)	(0.129554)	(71.2300)	(74.0275)
U_MSAOUT	0.218716*	0.272825**	144.520***	37.7069	0.274625**	0.320733***	151.854**	130.798*
	(0.120489)	(0.111959)	(54.3586)	(49.8929)	(0.121444)	(0.106754)	(70.0568)	(78.0708)
G_ELEM	-0.296640	-0.130264	-61.7864	-116.891***	-0.627453**	-0.579317	-228.187	-246.315
	(0.377532)	(0.390823)	(270.710)	(33.3774)	(0.294313)	(0.401398)	(258.074)	(247.427)
G_HIGH	-0.261037	-0.156691	-77.9519	-120.850***	-0.437687	-0.358076	161.198	-228.506
	(0.373896)	(0.338136)	(267.373)	(24.8089)	(0.294164)	(0.401411)	(260.181)	(247.371)
G_COLL	-0.122580	-0.051575	-43.8757	-67.9840**	-0.316857	-0.311526	-68.9267	-170.123
	(0.378326)	(0.393290)	(268.379)	(224.155)	(0.304984)	(0.410347)	(265.573)	(249.144)
G_GRAD	-0.025996	0.081425	59.2251	-27.0390	-0.587269	-0.492280	-186.365	-240.715
	(0.409418)	(0.417484)	(283.528)	(88.5556)	(0.360187)	(0.413582)	(282.514)	(270.976)
T_OWNER	0.081539	0.047874	13.5664	-15.0828	0.083272	0.127868*	-17.7778	20.8597
	(0.100442)	(0.091425)	(52.1893)	(15.3729)	(0.104502)	(0.071447)	(65.0999)	(59.8987)
F_STAMP	-0.040440	-0.089370	-73.0699	-18.2649	0.249857**	0.242112**	118.137	136.562**
	(0.113260)	(0.113676)	(55.7440)	(36.0411)	(0.111250)	(0.112420)	(80.8169)	(67.8179)
W_YES	0.219003*	0.241577**	147.183**	68.1081***	0.343824***	0.292344**	201.119**	156.953**
	(0.130468)	(0.107518)	(71.8602)	(26.3139)	(0.099263)	(0.126146)	(90.9506)	(71.6758)
Y_95	-0.091190	-0.087246	-53.5658	-25.0044	-0.207144**	-0.206691**	-73.3981	-76.2403
	(0.107249)	(0.094775)	(54.3965)	(58.5674)	(0.100865)	(0.088815)	(67.3483)	(56.4721)
Y_96	-0.063939	-0.122205	-84.6058	-17.4081	-0.144587	-0.176632*	-37.7361	-15.8704
	(0.101382)	(0.098441)	(55.7632)	(29.5803)	(0.109249)	(0.097170)	(70.8156)	(49.8848)
I. Mills R.	2.47990	-	975.273	-	9.34725*	-	7786.74***	-
	(2.56884		(1263.22)		(5.24860)		(2710.50)	
SIGMA	-	1.12657***	-	541.988***	-	1.11167	-	682.328***
		(0.040361)		(17.7474)		(~0.00000)		(19.3579)
RHO	-	-0.931261***	-	1.000000	-	1.000000	-	1.000000
		(0.031271)		(~0.00000)		(~0.00000)		(~0.00000)

Note: Std. errors in parenthesis. Significance of 2-tail t-test: *** - 1% level; ** - 5% level; * - 10% level.

Table 4 - Parameter Estimates of Food Consumption for Hispanic Households in the U.S., 1994-96

Variables	BEEF				PORK			
	Double Logarithmic model		Semi Logarithmic model		Double Logarithmic model		Semi Logarithmic model	
	HP	SS	HP	SS	HP	SS	HP	SS
Constant	-24.8058	5.81335***	-5789.22**	-156.859*	-146.516	6.88781***	-700.851	21.4169
	(21.3805)	(1.24020)	(2735.79)	(93.2325)	(177.995)	(2.15051)	(7969.90)	(72.6893)
LINCWK	0.460717	0.064339	97.5839**	11.8562	6.21678	-0.413558**	25.2189	3.03272
	(0.324463)	(0.119749)	(42.3979)	(9.57597)	(7.71837)	(0.169171)	(348.609)	(7.08812)
LHHSIZE	4.813.81	-0.115956	965.837**	76.0756***	5.87034	0.425296	55.3345	18.3180
	(3.39090)	(0.177875)	(434.047)	(17.2478)	(6.29569)	(0.287754)	(279.790)	(11.8355)
LAGE	1.91885	-0.273010	387.133*	14.0007	2.89013	0.166397	11.9197	10.0920
	(1.53077)	(0.238419)	(192.451)	(21.1776)	(3.09026)	(0.348409)	(133.306)	(15.6569)
S_FEM	-0.108552	-0.074721	-8.81173	0.132988	-0.214613	-0.193330	-14.0126	-2.62980***
	(0.127256)	(0.118816)	(14.9771)	(3.54908)	(0.204307)	(0.186951)	(11.3964)	(0.025518)
O_MEX	-0.013410	0.090200	14.6894	-0.221178	0.006254	-0.001170	0.925912	-0.380914***
	(0.140155)	(0.122399)	(15.3252)	(7.09971)	(0.222875)	(0.187702)	(15.8679)	(0.026807)
O_PRI	0.185998	0.070556	7.26112	-3.03520	0.124326	0.045327	-5.31755	5.52730***
	(0.197780)	(0.218440)	(22.8051)	(15.9609)	(0.246699)	(0.310966)	(14.1937)	(0.039512)
O_CUB	0.400572	0.387292	46.6112	28.8856	0.368701	0.387260	11.3696	-1.80938***
	(0.334983)	(0.304711)	(39.9380)	(22.9697)	(0.611493)	(0.532412)	(21.0356)	(0.021987)
R_NEAST	0.272785	0.300101	38.8260*	9.35843	0.167530	0.234627	16.3876	-1.02558***
	(0. 187796)	(0.185981)	(22.6313)	(10.7096)	(0. 318456)	(0.368645)	(17.8237)	(0.036041)
R_MWEST	0.192282	0.129582	7.93609	14.0370	-0.016703	0.076092	1.99342	-1.13378***
	(0.196558)	(0.217245)	(19.4438)	(17.4489)	(0.335860)	(0.323832)	(16.7377)	(0.030164)
R_SOUTH	0.034555	0.087069	10.9063	-4.02403	-0.322611	-0.253349	-5.28293	-1.14155***
	(0.155284)	(0.125534)	(16.7033)	(7.22632)	(0.226725)	(0.205962)	(11.3753)	(0.029828)
U_MSAINC	-0.291800	-0.249893	-28.6220*	-12.6503	0.643298**	0.648134**	32.0774*	5.25864***
	(0.179195)	(0.152355)	(16.6949)	(12.0183)	(0.260819)	(0.254045)	(18.6045)	(0.029897)
U_MSAOUT	0.124620	0.069116	8.61386	-0.060409	0.400615	0.357767	10.9333	7.19648***
	(0.158945)	(0.141494)	(16.3348)	(9.43177)	(0.254422)	(0.234560)	(15.8785)	(0.027625)
G_ELEM	-0.069625	0.173096	6.78490	-36.8087**	-0.617908	-0.499962	-51.9124	-149.917***
	(0.438808)	(0.446561)	(28.1281)	(17.1381)	(0.411571)	(1.17157)	(33.3966)	(0.141871)
G_HIGH	0.035880	0.241328	17.6307	-27.3483**	-0.781067*	-0.676628	-61.1804**	-150.106***
	(0.441003)	(0.448515)	(29.8820)	(12.6815)	(0.403811)	(1.16887)	(29.6010)	(0.148933)
G_COLL	0.003132	0.034877	-7.20791	-20.2999	-0.528426	-0.463589	-55.3050	-145.810***
	(0.443043)	(0.458684)	(29.1643)	(13.8338)	(0.433333)	(1.17978)	(34.2839)	(0.196958)
G_GRAD	-0.161619	-0.031134	-15.8635	-18.6252	-1.65486**	-1.42891	-68.3875	-150.783***
	(0.527348)	(0.499306)	(35.6638)	(18.0587)	(0.694929)	(1.22396)	(50.2218)	(.155669)
T_OWNER	0.063958	0.044018	5.48684	-5.99714	-0.055666	0.035978	2.43402	-4.83287***
	(0.139093)	(0.129624)	(13.3737)	(11.0594)	(0.189422)	(0.200116)	(11.7646)	(0.039684)
F_STAMP	0.075435	0.040441	-1.96109	10.6900	-0.574159**	-0.508893*	-18.8045	-4.65897***
	(0.158974)	(0.158900)	(16.4596)	(13.8844)	(0.242897)	(0.261686)	(14.7885)	(0.034213)
W_YES	-0.046188	-0.033787	-8.26283	-7.68966	0.188560	0.240033	23.5736	5.33767***
	(0.156719)	(0.151537)	(18.3099)	(11.0380)	(0.285874)	(0.259581)	(21.9988)	(0.041653)
Y_95	-0.161923	-0.127654	-12.3441	-2.99794	-0.023048	0.106311	16.7416	-4.47172***
	(0.139456)	(0.135471)	(14.6050)	(7.77611)	(0.237659)	(0.205741)	(14.2438)	(0.025384)
Y_96	-0.100746	-0.088120	-3.12602	-5.08801	0.183375	0.162591	8.42302	4.57313***
	(0.151085)	(0.132139)	(15.0510)	(6.98460)	(0.201637)	(0.205652)	(13.1680)	(0.034402)
I. Mills R.	19.8022	-	3992.91**	-	85.8293	-	484.519	-
	(14.5505)		(1853.90)		(101.369)		(4542.55)	
SIGMA	-	1.52662***	-	152.688***	-	1.77229***	-	106.052
		(0.090293)		(6.18301)		(0.243782)		(~0.00000)
RHO	-	-0.934203***	-	1.000000	-	-0.880197***	-	1.000000
		(0.024523)		(~0.00000)		(0.075685)		(~0.00000)

Note: Std. errors in parenthesis. Significance of 2-tail t-test: *** - 1% level; ** - 5% level; * - 10% level.

Table 5 - Parameter Estimates of Food Consumption for Hispanic Households in the U.S., 1994-96

Variables	FATS				SUGAR			
	Double Logarithmic model		Semi Logarithmic model		Double Logarithmic model		Semi Logarithmic model	
	HP	SS	HP	SS	HP	SS	HP	SS
Constant	-9.01640	2.53342**	-157.275	-55.2273***	34.2014**	3.14498**	955.526	-110.993
	(12.8411)	(1.02024)	(244.130)	(16.8342)	(16.7376)	(1.41103)	(826.874)	(68.1305)
LINCWK	0.799357	-0.072949	13.9616	7.83446***	-0.263908	0.124497	-2.43926	5.73283
	(0.938646)	(0.101554)	(18.1170)	(1.82987)	(0.214861)	(0.124908)	(9.99288)	(5.07573)
LHHSIZE	-0.126048	0.251060	-1.84873	-0.569780	-0.007229	0.775150***	11.8544	36.2587***
	(0.532902)	(0.157712)	(10.2756)	(2.67892)	(0.486613)	(0.180487)	(25.5445)	(8.30124)
LAGE	0.980984	-0.140947	14.0184	5.48740	-5.48877**	-0.427467*	-168.523	17.8972
	(1.12714)	(0.213150)	(21.1970)	(3.63735)	(2.79218)	(0.243684)	(140.676)	(10.9774)
S_FEM	0.123658	0.097465***	2.52387	0.692298	0.034533	-0.054894	3.26094	0.984263
	(0.118932)	(0.030111)	(2.51806)	(0.809908)	(0.139030)	(0.127544)	(7.86340)	(5.30220)
O_MEX	0.009352	0.079189***	0.237469	0.136370	0.019021	0.053444	5.50002	0.569691
	(0.132692)	(0.026237)	(2.66592)	(0.190001)	(0.154469)	(0.137893)	(8.23651)	(3.52768)
O_PRI	-0.289454	-0.108735	-6.87065	-1.48096	-0.122270	-0.272457	-17.0529*	-3.51035
	(0.215406)	(0.113158)	(4.66810)	(2.22145)	(0.249956)	(0.223945)	(9.52558)	(10.4754)
O_CUB	-0.293874	-0.312678***	-11.7330*	-1.28093	0.035010	0.017780	4.01780	0.605144
	(0.353804)	(0.070099)	(6.35446)	(59.8567)	(0.374718)	(0.353128)	(25.8911)	(5.99900)
R_NEAST	0.546762***	0.208739***	7.94635*	1.99516	0.289675	0.421520*	18.0414	1.76859
	(0.184834)	(0.005515)	(4.28349)	(2.56499)	(0. 241375)	(0.215570)	(11.8335)	(4.45707)
R_MWEST	-0.016588	-0.398362***	-3.24735	-1.21678	0.158920	0.151065	8.58058	2.19575
	(0.236132)	(0.093630)	(4.65325)	(1.34900)	(0.249086)	(0.215630)	(10.8726)	(6.61630)
R_SOUTH	0.063911	-0.053283*	-1.55020	-0.374870	0.015666	0.144164	9.37270	1.33625
	(0.142525)	(0.027865)	(2.79261)	(0.861032)	(0.163213)	(0.147536)	(9.60024)	(4.75791)
U_MSAINC	-0.137111	0.074267**	-1.79483	-1.66393	-0.369457	-0.327610*	-13.4088	-1.81030
	(0.172352)	(0.037955)	(3.58646)	(1.11782)	(0.199415)	(0.181634)	(10.9519)	(4.89381)
U_MSAOUT	0.057857	0.087792***	-3.53339	-0.857420	-0.228309	-0.154837	-0.438327	-1.26643
	(0.153894)	(0.013794)	(3.20120)	(1.39057)	(0.179127)	(0.167875)	(10.2582)	(4.67182)
G_ELEM	0.816917	1.39649***	10.1578**	2.81620***	0.145785	0.784303	22.7145**	8.65108
	(0.718045)	(0.150009)	(4.43553)	(0.424892)	(0.362042)	(0.784128)	(10.8455)	(43.6112)
G_HIGH	0.934031	1.65738***	14.4799***	2.87221	0.218806	0.777220	18.2193	8.52398
	(0.711412)	(0.134352)	(4.40363)	(1.75676)	(0.365226)	(0.784275)	(12.3332)	(43.7039)
G_COLL	1.07662	1.79248***	15.5484***	3.45725	0.386474	0.959962	24.8902*	8.72387
	(0.716337)	(0.080033)	(5.00786)	(2.64199)	(0.391104)	(0.791591)	(14.4235)	(43.8510)
G_GRAD	1.43638*	1.99269***	26.9018***	5.54727***	0.325278	0.910920	22.3005	12.1350
	(0.750759)	(0.088937)	(8.03110)	(1.56866)	(0.456457)	(0.815597)	(24.5004)	(44.0402)
T_OWNER	-0.115853	0.126268***	-2.09623	-0.372583	0.139321	0.006373	-5.18843	1.05365
	(0.137634)	(0.026530)	(2.81147)	(0.664163)	(0.166318)	(0.146827)	(9.59730)	(5.49147)
F_STAMP	0.077498	0.285203**	1.26553	-0.635114	0.267598	0.341058**	16.7733	1.70664
	(0.167378)	(0.128934)	(3.03583)	(1.57702)	(0.187607)	(0.165026)	(11.4757)	(6.50965)
W_YES	-0.216232	-0.337188***	-5.31683*	-1.11633***	-0.160550	-0.169592	-7.12927	-1.37825
	(0.167123)	(0.125232)	(2.80788)	(0.314925)	(0.188274)	(0.161605)	(10.9363)	(6.99914)
Y_95	-0.227680*	-0.630133***	-7.90553***	-2.09418**	-0.319413*	-0.294569**	-12.6739	-2.26183
	(0.136842)	(0.035084)	(2.76442)	(0.911620)	(0.165618)	(0.140409)	(8.00905)	(5.29988)
Y_96	0.074110	-0.175212***	0.077070	-0.890270	-0.153160	-0.190172	-9.95405	-1.30419
	(0.148989)	(0.028770)	(3.13467)	(1.04627)	(0.155132)	(0.149810)	(9.19108)	(5.47146)
I. Mills R.	5.21443	-	79.6277	-	-26.9929*	-	-847.204	-
	(8.14607)		(151.086)		(14.0294)		(693.381)	
SIGMA	-	1.71659***	-	29.2378	-	1.79645***	-	86.2594
		(0.063756)		(~0.00000)		(0.078042)		(~0.00000)
RHO	-	-1.000000	-	1.000000	-	-0.947851***	-	1.000000
		(~0.00000)		(~0.00000)		(0.027516)		(~0.00000)

Note: Std. errors in parenthesis. Significance of 2-tail t-test: *** - 1% level; ** - 5% level; * - 10% level.

Table 6 - Parameter Estimates of Food Consumption for Hispanic Households in the U.S., 1994-96

Variables	LEGUMES, NUTS, AND SEEDS				BEVERAGES			
	Double Logarithmic model		Semi Logarithmic model		Double Logarithmic model		Semi Logarithmic model	
	HP	SS	HP	SS	HP	SS	HP	SS
Constant	2.54406** (1.16941)	1.60113 (1.06210)	-134.700 (240.094)	202.273 (167.838)	16.0192*** (5.45423)	3.62013*** (0.787624)	6554.07 (6334.37)	-2582.60** (68.1305)
LINCWK	-0.140577 (0.204885)	0.145999 (0.104824)	0.387283 (43.6641)	-17.0202 (17.1810)	0.843977*** (0.302228)	0.145281* (0.078306)	679.355** (345.363)	167.022* (98.5917)
LHHSIZE	1.24668** (0.267891)	0.240847 (0.150383)	202.278 (139.111)	189.445*** (30.2475)	0.275909* (0.149354)	0.563699*** (0.117874)	597.577*** (176.042)	724.395*** (142.962)
LAGE	0.037527 (0.323699)	0.487084** (0.196453)	-2.82867 (71.6438)	-44.5924 (36.3381)	-2.83208** (1.34878)	0.251850 (0.162359)	-1895.06 (1539.86)	372.376** (187.937)
S_FEM	0.002243 (0.112798)	0.034315 (0.109265)	3.55375 (26.7516)	3.46648 (6.49657)	0.036001 (0.108701)	0.061230 (0.073438)	38.5899 (127.901)	39.0632 (105.207)
O_MEX	0.323571** (0.130381)	0.300246** (0.117468)	41.2066 (28.5064)	7.57542 (18.6845)	-0.012902 (0.105614)	0.016676 (0.090943)	15.6741 (126.607)	-86.3234 (94.6552)
O_PRI	-0.022915 (0.227781)	0.001423 (0.208886)	39.1282 (37.0741)	-3.14534 (32.7995)	-0.037785 (0.182233)	0.252145** (0.127486)	-151.528 (201.004)	-189.026 (137.458)
O_CUB	0.744370** (0.369086)	0.619191* (0.318068)	223.580* (117.342)	53.8285 (59.8567)	-0.024566 (0.263890)	0.078636 (0.092983)	-237.357 (254.745)	-210.011 (374.497)
R_NEAST	0.032239 (0.193974)	-0.143988 (0.185780)	-80.7893** (36.3058)	-6.55583 (38.9977)	-0.018454 (0. 159139)	-0.210940* (0.215570)	-73.0354 (168.947)	-106.916 (160.909)
R_MWEST	-0.517945** (0.249914)	-0.456068** (0.199810)	-58.5551 (49.0259)	-7.39171 (9.37286)	0.059030 (0.192214)	0.136827 (0.139529)	38.8223 (220.560)	-21.1165 (148.680)
R_SOUTH	0.250705* (0.133116)	0.178114 (0.124442)	19.4589 (30.0245)	8.09779 (9.40701)	-0.150714 (0.122339)	-0.179822* (0.094752)	-80.4367 (136.229)	-5.13379 (97.9164)
U_MSAINC	0.047813 (0.162785)	0.140877 (0.153151)	22.3325 (33.2791)	-1.35730 (18.2850)	-0.383596*** (0.147477)	-0.270927** (0.111679)	-437.831** (179.613)	-14.7726 (126.153)
U_MSAOUT	-0.101062 (0.147769)	-0.016061 (0.147012)	-6.58331 (30.9907)	-0.184731 (20.8897)	-0.392548*** (0.138079)	-0.386146*** (0.088344)	-476.492*** (173.009)	-120.291* (66.9972)
G_ELEM	0.596302 (0.479751)	0.499650 (0.417812)	44.4777 (108.501)	-2.16841 (125.7657)	1.11796*** (0.305377)	1.24967*** (0.190950)	1024.73*** (214.395)	990.547* (521.416)
G_HIGH	0.640843 (0.476849)	0.584190 (0.419323)	75.8204 (107735)	1.15206 (24.2437)	1.14988*** (0.300020)	1.39278*** (0.208539)	1156.88*** (222.032)	1067.49** (523.680)
G_COLL	0.482565 (0.493516)	0.456030 (0.427515)	54.0479 (110.941)	-4.94423 (22.3407)	1.21834*** (0.314469)	1.40227*** (0.224962)	1148.11*** (231.610)	968.880* (527.018)
G_GRAD	0.542817 (0.543928)	0.502545 (0.461953)	39.9438 (115.429)	-33.2512 (57.0117)	1.11277*** (0.373449)	1.59087*** (0.264684)	1272.57*** (374.056)	1080.88** (541.639)
T_OWNER	-0.267652** (0.127974)	-0.311446*** (0.115504)	-78.7940*** (30.3449)	-13.8263 (9.83528)	-0.090054 (0.122954)	-0.043480 (0.074003)	-182.642 (145851)	-88.0214 (65.2919)
F_STAMP	0.002594 (0.148361)	0.032591 (0.143411)	-9.20042 (35.2093)	-6.17879 (21.7411)	0.006316 (0.157004)	-0.108723 (0.091808)	32.2523 (165.001)	53.5677 (129.794)
W_YES	0.141293 (0.155117)	0.167535 (0.138928)	62.2530 (39.8680)	19.2653 (17.0201)	-0.262734* (0.149206)	-0.141623 (0.104434)	-213.083 (193.548)	-187.192* (105.609)
Y_95	-0.178842 (0.135686)	-0.162881 (0.122544)	-29.5344 (26.9578)	-5.47457 (17.6911)	-0.082029 (0.116537)	-0.198365* (0.107928)	-26.6480 (130.570)	-41.7435 (67.3468)
Y_96	0.190235 (0.128302)	0.172493 (0.123700)	41.5797 (31.5996)	10.6424 (12.1297)	0.202267* (0.155132)	0.130070 (0.092620)	363.562** (145.884)	175.495 (115.405)
I. Mills R.	2.59068 (2.39072)	-	244.900 (507.978)	-	-35.1647** (15.1432)	-	-24740.9 (17595.4)	-
SIGMA	-	1.46254*** (0.057414)	-	293.305*** (10.6086)	-	1.25997 (~0.00000)	-	1449.00*** (40.4852)
RHO	-	-0.925930*** (0.026287)	-	1.000000 (~0.00000)	-	-1.000000 (~0.00000)	-	1.000000 (~0.00000)

Note: Std. errors in parenthesis. Significance of 2-tail t-test: *** - 1% level; ** - 5% level; * - 10% level.

Table 7 - Statistical Evidence of Selectivity Bias[1] According to the Number of Positive Observations, for the Hispanic Households in the U.S., 1994-96

Food Group	Positive Observations [2]		HP Method		SS Method	
	Number	%	DL	SL	DL	SL
Grains	641	99.7	Yes	Yes	Indeterm.	Indeterm.
Milk	627	97.5	Yes	Yes	Indeterm.	Indeterm.
Vegetables	620	96.4	No	No	Indeterm.	Indeterm.
Beverages	599	93.2	Yes	No	Indeterm.	Indeterm.
Fruits	566	88.0	No	No	Yes.	Indeterm.
Legumes	517	80.4	No	No	Yes	Indeterm.
Sugar	513	79.8	Yes	No	Yes	Indeterm.
Fats	463	72.0	No	No	Indeterm.	Indeterm.
Beef	351	54.6	No	Yes	Yes	Indeterm.
Pork	203	31.6	No	No	Yes	Indeterm.

[1] Selectivity bias: Yes – Statistical evidence; No – Lack of statistical evidence; Indeterm. – Indeterminate; the model suffers from collinearity problems.

[2] Number of households that consumed a positive amount of the corresponding food group during the survey period. The percentage is from a total of 643 households.