ESSAYS ON CHOICE AND DEMAND ANALYSIS OF ORGANIC AND CONVENTIONAL MILK IN THE UNITED STATES

A Dissertation

by

PEDRO AYA-AY ALVIOLA IV

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2009

Major Subject: Agricultural Economics

ESSAYS ON CHOICE AND DEMAND ANALYSIS OF ORGANIC AND CONVENTIONAL MILK IN THE UNITED STATES

A Dissertation

by

PEDRO AYA-AY ALVIOLA IV

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Oral Capps, Jr.
Committee Members,	David A. Bessler Ximing Wu Steven N. Wiggins Denise Mainville
Head of Department,	John P. Nichols

December 2009

Major Subject: Agricultural Economics

ABSTRACT

Essays on Choice and Demand Analysis of Organic and Conventional Milk in the United States. (December 2009)

Pedro Aya-ay Alviola IV, B.S., University of the Philippines; M.A., University of the Philippines

Chair of Advisory Committee: Dr. Oral Capps, Jr.

This dissertation has four interrelated studies, namely (1) the characterization of milk purchase choices which included the purchase of organic milk, both organic and conventional milk and conventional milk only; (2) the estimation of a single-equation household demand function for organic and conventional milk; (3) the assessment of binary choice models for organic milk using the Brier Probability score and Yates partition, and (4) the estimation of demand systems that addresses the censoring issue through the use of econometric techniques.

In the first paper, the study utilized the estimation of both multinomial logit and probit models in examining a set of causal socio-demographic variables in explaining the purchase of three outcome milk choices namely organic milk, organic and conventional milk and conventional milk only. These crucial variables include income, household size, education level and employment of household head, race, ethnicity and region.

Using the 2004 Nielsen Homescan Panel, the second study used the Heckman two-step procedure in calculating the own-price, cross-price, and income elasticities by estimating the demand relationships for both organic and conventional milk. Results indicated that organic and conventional milk are substitutes. Also, an asymmetric pattern existed with regard to the substitution patterns of the respective milk types.

Likewise, the third study showed that predictive outcomes from binary choice models associated with organic milk can be enhanced with the use of the Brier score method. In this case, specifications omitting important socio-demographic variables reduced the variability of predicted probabilities and therefore limited its sorting ability.

The last study estimated both censored Almost Ideal Demand Systems (AIDS) and Quadratic Almost Ideal Demand System (QUAIDS) specifications in modeling non-alcoholic beverages. In this research, five estimation techniques were used which included the usage of Iterated Seemingly Unrelated Regression (ITSUR), two stage methods such as the Heien and Wessells (1990) and the Shonkwiler and Yen (1999) approaches, Generalized Maximum Entropy and the Dong, Gould and Kaiser (2004a) methods. The findings of the study showed that at various censoring techniques, price elasticity estimates were observed to have greater variability in highly censored non-alcoholic beverage items such as tea, coffee and bottled water.

DEDICATION

To my late father, Dr. Pedro L. Alviola III, and mother, Mrs. Nena Alviola

ACKNOWLEDGEMENTS

Many individuals have been instrumental in my pursuit of finishing graduate studies here in Texas A&M University. First and foremost is my major professor, Dr. Oral Capps, Jr., whose integrity, work ethic, research and moral excellence have been my guideposts for which I still continue to strive. He has patiently guided and mentored me, not only on the rigors of applied economic research, but also on the application of a sound and morally principled work ethic that embodies compassion, integrity and honesty. My endless corridor and out-of-the-classroom philosophical, methodological and economic discussions with Dr. David Bessler have been a wonderful haven for recognition of seminal ideas, correction of faulty thinking and the unyielding pursuit of research excellence as a pragmatic way of improving people's lives.

A special thanks to my other committee members, Dr. Ximing Wu, Dr. Steven Wiggins and Dr. Denise Mainville, whose invaluable comments and suggestions have been substantial in improving this dissertation. Also, I would like to thank Dr. Diansheng Dong for assisting me with his estimation technique on censored demand systems. Also, a special note of thanks to Mr. Jim Hill, President of the Southwest Dairy Farmers Association, for his continued support of the project.

I owe an intellectual debt of gratitude to my professors, Dr. George Davis, Dr. Richard Woodward, Dr. Ron Griffin and Dr. Allan Love. I learned through them that faulty thinking is usually a result of bad and idle reasoning. I would also like to thank

Vicki Heard, Dr. David Leatham and Robin Faltesek whose patience I have probably tested a couple of times by my repetitive inquiries.

I thank my friends and classmates in the department especially Ben, Johnny, Ryan, Yongxia, Amy, Xavier, Clarke and Ke-li who made graduate student life more bearable, challenging and interesting.

I would also like to thank the Community and Homebuilders classes at Grace Bible Church for the much needed prayers and encouragement. A special thank you to Mark and Cheryll Mikeska for always being there especially during the trying times. Your mere presence was comfort enough during those circumstances. I'm especially indebted to David and Brenda Roye for their friendship and the loving kindness that they have shown to me and my family.

A special thanks to our Filipino friends in Aggieland and College Station, and to the Tabiens for always welcoming us to their home in Beaumont during semester breaks and for their collective wisdom of which I can always count on during difficult times.

I would like to thank my wife, Novie for her unceasing love, patience and faith. She has been my rock, guide, best friend and mentor and because of her, my pursuit of the PhD degree has been truly inspiring. My special thanks and very proud feelings go to my two sons, Nathan Abraham and John Matthew, whose presence alone ignites joy and propels me to push forward. A special thanks to Mama Nena, Mommy and Daddy and my siblings John, Phillip, Marites and Darel for the constant encouragement and support that you folks have always provided.

And finally, I humbly dedicate this piece of work to my Lord and Savior Jesus Christ. You have always been there for me and my family and because of you, Lord, this academic journey has not only been made possible but was also fruitful. These things I do not own but was only given for stewardship and I’m extremely humbled to be under Your abundant grace and mercy every day.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

TABLE Page

CHAPTER I

INTRODUCTION

The recent shift towards differential diet mechanisms in favor of healthier foods is an indicator that the typical American consumer is now highly conscious of both the different food items being offered in the market and its impact on nutrition and total wellbeing. Similarly, recent trends in supermarkets offering healthier and natural food choices can be seen as a reaction to rising demand for healthier foods. This rapid expansion in the organic food market has in effect triggered in part the increasing growth in the organic milk industry.

Why look at the organic milk and non-alcoholic beverage industry? There are several reasons why these markets deserve research scrutiny. First, the increasing growth of the organic milk market represents the current shift of healthy food items that are increasingly being demanded by the American consumer. Examining this particular market will help define profiles of consumers that are responsive and sensitive to healthy food choices and therefore assist in the fine tuning of policies that addresses significant health concerns in the United States. As well, it is important to focus on the interdependencies of milk with other products such as fruit juice, tea, carbonated soft drinks and bottled water. In this regard, the non-alcoholic beverage complex represents

---

This dissertation follows the style of the *American Journal of Agricultural Economics*.

ideal cases for testing estimation procedures that address the censoring problem in demand systems estimation.

With varying levels of censoring, we are in position to evaluate the performance of several cutting edge estimators. Thus, this dissertation will contribute towards a clearer picture of how choices are made with regards to healthy alternative foods such as organic milk and shed some light on the existing debate on the appropriate estimator to use in estimating censored demand systems.

Organization of the Dissertation

Chapter II begins the study by looking at the possible household choice determinants of three milk choice outcomes, namely: organic milk and conventional milk, organic milk only and conventional milk only. This was achieved through the use of both multinomial logit and probit models. On the other hand, Chapter III estimated a two-stage model, namely the use of a probit model in the first stage to account for selection bias, and then incorporating it in the second stage, where the calculation of price and income elasticity coefficients was done by estimating demand equations for both organic and non-organic milk. In Chapter IV, the extensive use of discrete choice models in the research led naturally to the examination of the quality of predicted probabilities. This chapter assessed the prediction probabilities of fundamental discrete choice models, namely the linear probability model, the logit and probit models by probability scoring techniques such as the Brier Probability Scoring Method and Yates partition. On the other hand, Chapter V estimated a micro-demand system of the non-alcoholic beverages that included conventional milk of which varying levels of data censoring were observed. The central theme of this chapter looked at various methods of estimating censored demand system that have been recently proposed in the literature and made comparative analysis of each estimation technique. Finally, in Chapter VI we summarize the findings of the essays and provide recommendations and key points for future research efforts.

CHAPTER II

EXAMINING THE FACTORS AFFECTING HOUSEHOLD PURCHASE OF ORGANIC AND CONVENTIONAL FLUID MILK IN THE UNITED STATES

In recent years, the fluid milk industry has undergone several notable changes. In the past, fluid milk consumers have looked at fat content levels (Gould, 1996) as an indicator that product choices can be available to support healthy dietary lifestyles. US consumers are now highly conscious of the different food items offered in the marketplace and their impact on nutrition and total well-being. With advances in biotechnology, conventional milk production has increased because growth hormones such as the *recombinant bovine somatotropin* (rBST) have been widely available to the dairy industry. However, despite scientific claims that rBST milk is safe for human consumption, the public perception has been to oppose its introduction and demand other forms of milk variants that are labeled rBST free. Unlike other dairy milk products that arose due to the controversies of the rBST milk dilemma, the organic fluid milk industry has been steadily rising with minimal influence from any coalitional networks. In fact, DuPuis (2000) argued that the industry's increasing market shares were due to the acceptance of mainstream consumers who saw organic milk as a viable alternative in meeting their changing taste and preferences. Also, the industry was characterized as flexible in terms of catering to those changing needs.

According to Dimitri and Venezia (2007), US sales of organic milk have been steadily rising by about 25 percent from the 1990's up until 2004. This growth was largely driven by the increasing market sales of the organic food market. Dimitri and Greene (2002) and Li et al. (2007) opined that consumer acceptance of organic food was largely driven on the grounds that it was a pesticide-free product. Furthermore, the organic milk industry has gained wider distribution from large retail chains such as Costco and Wal-Mart, thus boosting its product exposure in the market (Thompson, 1998; Dimitri and Venezia, 2007). In addition, dairy producers switched from conventional milk production to organic operations in response to opportunities created through the rightward shift in demand for organic milk. Thus, given these developments, organic milk sales have been increasing ever since starting in the early 1990s, while sales of conventional milk have been relatively constant during this time span (Miller and Blayney, 2006).

There have been previous studies concerning the demand interplay between organic and conventional milk. Several studies including Glaser and Thompson (2000), Dhar and Foltz (2005) and Alviola and Capps (2009) revealed that organic and conventional milk are substitutes and that there exist significant differential responses with regard to cross-price effects. These works dealt with the purchased quantities of organic and conventional milk. On the other hand, studies that examine the factors that drive the decision to buy organic and conventional milk have been limited. Dimitri and Venezia (2007) and Alviola and Capps (2009) examined the factors that affect the binary choice decision of buying organic and conventional milk at the household level. However, one can extend the dichotomous choice model to a polychotomous model because households may purchase organic milk only, conventional milk only or both organic and conventional milk conditional on the decision to purchase milk in the first place.

Thus, the objective is to characterize consumer buying behavior with respect to the three aforementioned milk purchase choices. In particular, we wish to identify and assess household characteristics that drive each of these types of milk purchases. In this way, we add to the literature by carrying out an extension of what had been previously undertaken regarding the purchase of organic and conventional milk.

Literature Review

Past studies regarding choice models that deal with organic and conventional milk have been instrumental in understanding the underlying factors that influence the purchase of both milk types. For example Hill and Lynchehaun (2002) cited various socio-demographic factors that affect the buying of organic milk. These factors included personal values, attitudes, age, and ethnicity, presence of children, education, advertising, taste, packaging quality, food scares, prices and income. Similarly Dimitri and Venezia (2007) presented a descriptive analysis of organic milk users based on analysis of Nielsen Homescan data for calendar year 2004. Their findings indicated that the typical organic milk consumer was white, highly educated and less than 50 years old. Also, organic milk users were generally Orientals and Hispanic. However, their analysis was based on descriptive statistics, and no formal statistical analysis was conducted.

Alviola and Capps (2009) utilized a probit model in characterizing the household choice between organic and conventional milk. The source of data was also the 2004 Nielsen Homescan panel. They concluded that households likely to purchase organic milk were single person, affluent, highly educated, located in the west region, black, oriental, Hispanic and have no children. However, the major limitation of their study is that the choice outcome is limited to two (organic versus conventional milk) when in fact these outcomes can extend beyond binary choices.

Using the same data set as that of Dimitri and Venezia (2007) and Alviola and Capps (2009), McKnight (2007) looked at households that purchased organic milk and utilized cluster analysis to differentiate choices between organic milk and conventional milk. The key variable in this analysis was the percentage share of organic milk purchase to total fluid milk purchase. This choice partitioning then was used to construct a multinomial logit model with household socio-demographic variables as choice drivers. The findings indicated that households small in size with well educated household heads were more likely to purchase organic milk. The limitations of the study were twofold: (1) since the choice outcome variable was characterized as percentage of organic milk purchase to total fluid milk purchase, it ignored the interplay of choices between organic milk and conventional milk; (2) choices were assumed to be independent, ignoring the possibility that both organic milk and conventional milk choices might be related.

Methodology

In the literature, the use of multinomial models has been widespread with multinomial logit models dominating over probit models due to the ease of estimation. Starting with the work of McFadden (1978), Dubin and McFadden (1984) and more recently Train (2003), improvements on the multinomial logit model continuously have been refined. The inherent tractability of this model particularly in applied work in agricultural markets and commodities has been well received (Vergara et al., 2004). However the tractability of the multinomial logit model comes with a cost, in that it assumes the independence of irrelevant alternatives (IIA). The fallout of this assumption is the constancy of choice odds even as the number of alternative choices increases. With the use of multinomial probit models, on the other hand, the IIA assumption is relaxed.

With the pioneering work of Hausman and Wise (1978), applications of the multinomial probit model have been employed in various fields such as political science, especially in voter choice of candidates (Dow and Endersby, 2004; Alvarez and Nagler,1994), likelihood of completing high school and college education (Jepsen, 2008), transportation and brand choice (Nobile, Bhat and Pas, 1996 and Hrushka, 2007) and farming adoption decisions resulting in availability of multiple technology (Dorfman, 1996). However, if the number of choice alternatives exceeds four, the practicality of the use of the multinomial choice models diminishes due to mathematical complexity. The current thrust on workable solutions with regard to overcoming this formidable intractability has been the usage and refinement of numerical methods (Train, 2003; Weeks, 1997; Breslaw, 2002 and Bunch, 1991) in achieving solution convergence. Despite the advances in this field, some researchers particularly Maddala (1983) questioned the extra computational burden posed by the multinomial probit model. More recently Greene (2008) noted that while advances in numerical methods are now available for researchers, restrictions on the variance-covariance matrix of the error terms must be in place to achieve convergence.

Random Utility Model

Following Cameron and Trivedi (2005) and Greene (2008), consider a k^{th} choice multinomial model among a class of *m* choices. The utility function of the k^{th} choice can be written as

(1) $$U_k(V_k, \varepsilon_k) = V_k + \varepsilon_k, \qquad k = 1,2,\ldots\ldots,m,$$

where V_k and ε_k are the deterministic and stochastic factors of the *k*th choice. The deterministic component V_k can be expressed as $V_k = W_k\eta_k$ where W_k are the identified drivers of the individual's *k*th choice and η_k are the k-parameters to be estimated. One also can construct an alternative utility function U_r to represent the *r*th choice among the available *m* choices. Therefore to motivate the problem in terms of utility comparisons, an individual chooses the *k*th choice among all other competing choices as indexed by the *j*th choice if and only if $U_k \geq U_r$. This situation implies that an individual chooses choice *k* if and only if it yields the highest level of utility among all choices (McFadden, 1973, 1974a, 1974b and 1978). Following Cameron and Trivedi (2005), if we let *p* be the probability of occurrence, then the probability of occurrence of the *k*th choice (Pr (Y=k)) becomes:

(2) $$\begin{aligned} \Pr(Y = k) &= \Pr(U_k \geq U_r), \\ &= \Pr(V_k + e_k \geq V_r + e_r) \\ &= \Pr(W_k\eta_k + e_k \geq W_r\eta_r + e_r), \\ &= \Pr(e_r - e_k \leq W_k\eta_k - W_r\eta_r), \\ &= \Pr(\mu_{rk} \leq V_{rk}^{'}), \end{aligned}$$

where μ_{rk} and V'_{rk} are defined respectively as $e_r - e_k$ and $W_k\eta_k$ - $W_r\eta_r$. Assumptions can be made about the error terms. If the errors assume an extreme value distribution with mean 0 and variance $\pi^2/6$, then the resulting model is a multinomial logit model (MNL). On the other hand if the errors assume a jointly normal distribution, then a multinomial probit model (MNP) emerges.

The fundamental crux between these models has revolved around the independence of irrelevant alternative (IIA) axiom where the multinomial logit model has the property of its choice odds being invariant to additional alternatives. As additional alternatives are either being added or subtracted, the choice odds remain the same for any pairwise comparison of the relevant alternatives. This invariance property however raises serious concerns on model validity. As noted by Baltagi (2005), when choices are likely to be close substitutes, the MNL and its allied models (conditional logit models) may produce inconsistent estimates if the choices are truly not independent. This assumption maybe appealing in terms of empirical tractability but is very restrictive in terms of characterizing underlying utility preferences (Greene, 2008).

One of the alternative approaches however, is to forego the IIA axiom by assuming an error structure that is multivariate normal, leading to the multinomial probit model. Flexibility is achieved by permitting cross correlations among choices through the specification of a correlated error structure. However, the choice of the multinomial probit model comes with a cost that as the number of alternatives expands, the computational ability to evaluate multiple integrals in finding closed form solutions becomes increasingly difficult (Maddala, 1983). In this exercise, the deployment of the multinomial probit model for modeling organic and non-organic milk choices comes from two major considerations: (1) the IIA axiom may not be a realistic assumption to impose; and (2) the choice variable takes on only three responses, resulting in computationally tractable model from a numerical integration viewpoint (Maddala, 1983).

The choice to buy either organic or conventional milk yields the same odds of either purchasing one or the other milk type. However, if another alternative choice is given such as buying both organic and conventional milk then the IIA axiom presupposes that the odds between purchasing organic or conventional milk will not change. This imposition may not be realistic as one can immediately deduce that purchasing both milk types can affect the odds of purchasing either organic or conventional milk alone. The other reason revolves around numerical ease. While computational burden of estimating the multinomial probit model is exceedingly longer relative to the multinomial logit model, the three choices of either purchasing organic milk, conventional milk or both is still within the purview of the trivariate normal integral limit where standard analytical integration methods can still be applied (Cameron and Trivedi, 2005). However, if multiple responses exceeded four choices then simulation techniques such as frequency simulators, sampling and Bayesian estimation have been in recent years used to make the multinomial probit model tractable (Train, 2003). More recently Greene (2008) opined that caution has to be emphasized that in using multinomial probit models, the requirement of additional restrictions such as zero or equal correlation among the error terms are usually imposed in order to achieve convergence.

Multinomial Probit and Logit Models

A logical extension of the binary choice models is to estimate unordered discrete responses that go beyond two choice outcomes. Thus, for each choice outcome of the dependent variable, the corresponding discrete values range from to *0* to *m-1* where m denotes the maximum number of choice outcome. In this exercise, three choices have been identified wherein a household might purchase both organic and conventional milk (1), organic milk (2) and only conventional milk (3). Thus, these choices are characterized as unordered categorical variables in that the household may arbitrarily choose to purchase organic milk or conventional milk or both without being constrained by any choice-ordering axiom.

In using the multinomial probit model, consider the case where the choice variable takes on three responses and let W_i be a vector of independent variables that are related to the purchase of organic and conventional milk. Following Greene (2008), Cameron and Trivedi (2005), Wooldridge (2002), Gan (2007) and Maddala (1983), the probability of selecting both organic and conventional milk (1st) choice in a multinomial probit model can be represented as:

(3)
$$\Pr(Y=1) = \Pr(U_1 > U_2, U_1 > U_3),$$

$$= \Pr(W_1\eta_1 + e_1 > W_2\eta_2 + e_2, W_1\eta_1 + e_1 > W_3\eta_3 + e_3),$$

$$= \Pr(e_2 - e_1 < W_2\eta_2 - W_1\eta_1, e_3 - e_1 < W_3\eta_3 - W_1\eta_1),$$

$$= \Pr(e_2 < W_2\eta_2 - W_1\eta_1 + e_1, e_3 < W_3\eta_3 - W_1\eta_1 + e_1),$$

$$= \int_{-\infty}^{\infty} f(e_1)\left(\int_{-\infty}^{W_2\eta_2 - W_1\eta_1 + e_1} f(e_2)de_2 \quad \int_{-\infty}^{W_3\eta_3 - W_1\eta_1 + e_1} f(e_3)de_3 \right)de_1,$$

and if the error terms are assumed to be multivariate normal then the last expression becomes;

$$\text{(4)} \qquad \Pr(\text{Y=1}) = \int_{-\infty}^{\infty} f(e_1)(F(W_2\eta_2 - W_1\eta_1 + e_1)F(W_3\eta_3 - W_1\eta_1 + e_1))de_1,$$

$$\Pr(\text{Y=1}) = \int_{-\infty}^{\infty} \varphi(e_1)(\Phi(W_2\eta_2 - W_1\eta_1 + e_1)\Phi(W_3\eta_3 - W_1\eta_1 + e_1))de_1 \ ,$$

where φ(.) and Φ(.) are pdf and cdf respectively. For choice alternatives 2 and 3, the same process can be done in terms of deriving the choice probabilities.

In the multinomial logit model case, the choice variable takes on integer values from $j = 0,\ldots\ldots,m\text{-}1$ and let W_i be a vector of independent variables that are related the purchase of organic and non-organic milk. Following Greene (2008), Cameron and Trivedi (2005) and Wooldridge (2002), the probability of the ith individual selecting the jth choice in a multinomial logit model can be represented as:

$$\text{(5)} \qquad P_{ij} = \frac{\exp(W_i\eta_j)}{\sum_{j=0}^{m-1}\exp(W_i\eta_j)} \qquad j = 0,\ldots.,m-1,$$

where P_{ij} is the probability that the ith choice selected and η_i are the parameters to be estimated. For this exercise, the study evaluated i =3, where the choices are organic and conventional milk (q_i = 1), only organic milk (q_i = 2) and only conventional milk (q_i = 3).

Variance-Covariance Matrix Structure of the Error Terms for Multinomial Probit Model[1]

In order to relax the IIA assumption, the multinomial probit model permits cross correlations between the error terms. In this exercise, STATA's calculation of the variance-covariance matrix requires several restrictions, which translates into constraining one of the variances in the differenced error variance-covariance matrix in order for the matrix to be identified (Note that it does not matter which variance need to be constrained). Following Long and Freese (2006), Kropko (2008) and StataCorp (2005) and assuming that the variance of Choice 1 is fixed, the resulting differenced error variance-covariance matrix can be denoted as;

$$\begin{bmatrix} \sigma_{\alpha_2}^2 & . \\ \sigma_{\alpha_2\alpha_3} & \sigma_{\alpha_3}^2 \end{bmatrix}, \tag{6}$$

where $\sigma_{\alpha_2}^2 = Var(e_2 - e_1)$ and $\sigma_{\alpha_3}^2 = Var(e_3 - e_1)$ and expanding further[2], we have

$$\begin{bmatrix} \sigma_{e_1}^2 + \sigma_{e_2}^2 - \rho_{e_1e_2}\sigma_{e_1e_2} & . \\ \rho_{e_2e_3}\sigma_{e_2}\sigma_{e_3} - \rho_{e_1e_2}\sigma_{e_1}\sigma_{e_2} - \rho_{e_1e_3}\sigma_{e_1}\sigma_{e_3} + \sigma_{e_1}^2 & \sigma_{e_1}^2 + \sigma_{e_3}^2 - \rho_{e_1e_3}\sigma_{e_1e_3} \end{bmatrix}, \tag{7}$$

In order to constrain $\sigma_{\alpha_2}^2$ into a constant, the STATA *asmprobit* routine restricts the variance of both choice 1 and choice 2 equal to 1. Thus in a three choice model, the

---

[1] Kropko (2008) and Long and Freese (2006) provide excellent discussions on how STATA calculates the variance-covariance matrix of the differenced error terms used in its "asmprobit" command. The discussion of the multinomial probit error variance-covariance structure follows their exposition.

[2] See Kropko (2008) for example.

following restrictions are imposed in order for the differenced errors variance-covariance matrix to be identified.

(8) $$\sigma_{e_1}^2 = \sigma_{e_2}^2 = 1,$$

(9) $$\rho_{e_1e_2} = \rho_{e_1e_3} = 0,$$

Thus, it follows from (8) that covariances (ex. $\sigma_{e_1e_2} = \rho_{e\ e_2}\sigma_{e_1}\sigma_{e_2} = 0$) associated with choice 1 are 0. With the restrictions from both equation 8 and 9, the final differenced error variance-covariance matrix as calculated by STATA's "asmprobit" command becomes;

(10) $$\begin{bmatrix} 2 & . \\ \rho_{e_1e_3}\sigma_{e_3} + 1 & 1 + \sigma_{e_3}^2 \end{bmatrix},$$

For this exercise the constrained choice is 3, thus the differenced error variance-covariance matrix is;

	Choice 2	Choice 1
Choice 2	2	.
Choice 1	0.64323	0.25022

As for the other multinomial probit (uncorrelated error) variant, this model was calculated by STATA's "mprobit"command. This type of variant is the normal counterpart of the multinomial logit model and therefore still assumes IIA resulting to error terms that are uncorrelated.

Marginal Effects

For the estimation of the marginal effects, one has to take the partial change of the choice probability with respect to the conditioning variables. Thus, the marginal effects for the multinomial logit model can be written as

$$\text{(14)} \qquad \frac{\partial P_{ij}}{\partial W_i} = P_{ij}\left[n_{ij} - \left(\sum_{i=0}^{k-1} \eta_{ij} P_{ij} \right) \right] \qquad j = 0,....,m-1\,,$$

Equation 14 can be interpreted as the change in probability of the ith choice of the jth household given a change in the independent variables W_i.

As for the multinomial probit model, the derivation of the respective marginal effects is much more complicated (see Dorfman, 1996 for example). The calculation of the marginal effects in both the multinomial logit model and multinomial probit model in STATA is done by numerical approximation.

Empirical Specification

In this empirical exercise, several socio-demographic variables such as household income, household size, employment status and educational level of household head, race, ethnicity, number of children in the household and region are hypothesized factors affecting purchasing choice of organic and conventional milk. The general multinomial model specification is given as follows:

$$\text{(15)} \qquad \begin{aligned} &\Pr(Y_{ij} = j \mid W) = \Phi(\eta_0 + \eta_1 Inc + \eta_2 HHsize + \eta_3 EMP \\ &+ \eta_4 Educ + \eta_5 Race + \eta_6 Ethcy + \eta_7 AgeChild + \eta_8 \operatorname{Re} g) + \varepsilon \end{aligned}$$

where, the ith household has the jth choice (j = 1, 2 and 3) denoting households who purchased both organic and conventional milk, organic milk only and conventional milk only, respectively. Φ is the cdf and W_i is the vector of socio-economic and demographic variables of the household which include *Inc* as household income and *HHsize* is the household size where indicator variables were created for one, two, three, four and five more members representing the number of household members respectively. Other demographic indicator variables include *EMP* as employment status of household head, while *Educ* is the level of education of the household head. The variable *Race* represents the race type and *Ethcy* refers to ethnicity, that is whether the household is Hispanic, or not. *Agechild* represents the presence of children in the household and finally the variable *Reg* represents region. Milk prices are not included in the multinomial logit/probit estimation. Prices were derived as the ratio of expenditure to quantity; but if there was no recorded purchase then no price can be computed[3].

Description of Data

For this empirical exercise, the data pertaining to the choice of purchasing organic and conventional milk, income and household socio-demographic variables are from the AC Nielsen Homescan Panel for calendar year 2004. The AC Nielsen scanner data set is the world's largest, on-going household scanner data survey system, tracking household purchases in the United States. Table 2.1 presents the definition and summary statistics of all the relevant variables partitioned by choice outcome.

For households that purchased both organic and conventional milk (choice 1), the average price paid for both milk types were approximately \$3.15/half gallon and \$2.03/half gallon, respectively. The average purchase quantity was approximately 8.53

---

[3] One may use imputation techniques to derive missing prices, but the empirical results are tied to the use of these procedures.

half gallons of organic milk and 39.39 half gallons of conventional milk. On the other hand, households that only purchased organic milk (choice 2), had an average purchase price and quantity of \$3.25/half gallon and 13.19 half gallons. Finally, the average purchase price and quantity of households that purchased only conventional milk (choice 3) were approximately \$1.75/half gallon and 47.68 half gallons.

From Table 2.1, the variable *Inc* is defined as household income, where for this sample, the average income level for households that purchased both organic and conventional milk was \$55,317, while for those household that purchased only organic milk the average household income is approximately \$49,044.Likewise, the average income for households that purchased only non organic milk is approximately \$49, 356. The study also used indicator variables to describe the number of household members with *hs1*as the base variable with *hs2* pertaining to a household having 2 members. The variables *hs3* and *hs4* denoted 3 and 4 members in a household while the last household size indicator variable *hsp5* describes 5 or more members in the household. The demographic values indicate that more than 70% of the household respondents for choice 1 and choice 3 are households with 1 or 2 members. For those households that purchased only organic milk (choice 2), almost 62 % are single-member households. *Agepcchild* corresponds to a dummy variable with 1 indicating the presence of children and 0 otherwise. Almost 25% of households associated with choices 1 and 3 have children, while only 8% of households associated with choice 2 (organic milk only) have children.

Table 2.1. Descriptive Statistics of Relevant Household Demographic Variables

	(Choice 1 = organic and conventional)				(Choice 2 = organic milk)				(Choice 3 = conventional milk)			
Variables	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
inc (Household Income)	55317	28181	5000	100000	49044	27683	5000	100000	49356	27117	5000	100000
agepcchild (Presence of children in Household)	0.254	0.435	0.000	1.000	0.080	0.271	0.000	1.000	0.254	0.435	0.000	1.000
Household Size												
hs1 (one member)	0.257	0.437	0.000	1.000	0.617	0.487	0.000	1.000	0.260	0.439	0.000	1.000
hs2 (two members)	0.392	0.488	0.000	1.000	0.258	0.438	0.000	1.000	0.392	0.488	0.000	1.000
hs3 (three members)	0.149	0.356	0.000	1.000	0.072	0.259	0.000	1.000	0.142	0.349	0.000	1.000
hs4 (four members)	0.126	0.332	0.000	1.000	0.034	0.182	0.000	1.000	0.128	0.334	0.000	1.000
hsp5 (five members)	0.076	0.264	0.000	1.000	0.019	0.137	0.000	1.000	0.078	0.268	0.000	1.000
Employment Status of Family Head												
emparttime (part time)	0.176	0.380	0.000	1.000	0.167	0.373	0.000	1.000	0.155	0.362	0.000	1.000
empfulltime(fulltime)	0.441	0.497	0.000	1.000	0.542	0.499	0.000	1.000	0.433	0.496	0.000	1.000
unemp(unemployed)	0.383	0.486	0.000	1.000	0.292	0.455	0.000	1.000	0.412	0.492	0.000	1.000
Educational Level of Family Head												
Education less than highschool	0.026	0.159	0.000	1.000	0.004	0.062	0.000	1.000	0.040	0.196	0.000	1.000
eduhighschool (highschool level)	0.188	0.391	0.000	1.000	0.072	0.259	0.000	1.000	0.287	0.453	0.000	1.000
edusomecollege (some college)	0.310	0.462	0.000	1.000	0.273	0.446	0.000	1.000	0.321	0.467	0.000	1.000
educollegeplus (collegeplus)	0.476	0.499	0.000	1.000	0.652	0.477	0.000	1.000	0.351	0.477	0.000	1.000
Race/Ethnicity												
white	0.757	0.429	0.000	1.000	0.674	0.470	0.000	1.000	0.835	0.371	0.000	1.000
black	0.128	0.334	0.000	1.000	0.178	0.383	0.000	1.000	0.092	0.289	0.000	1.000
oriental	0.038	0.192	0.000	1.000	0.053	0.225	0.000	1.000	0.020	0.138	0.000	1.000
other	0.076	0.265	0.000	1.000	0.095	0.293	0.000	1.000	0.054	0.226	0.000	1.000
hispyes(hispanic)	0.092	0.289	0.000	1.000	0.083	0.277	0.000	1.000	0.062	0.242	0.000	1.000
hispno (not hispanic)	0.908	0.289	0.000	1.000	0.917	0.277	0.000	1.000	0.938	0.242	0.000	1.000
Region												
east	0.169	0.375	0.000	1.000	0.178	0.383	0.000	1.000	0.162	0.369	0.000	1.000
central	0.168	0.374	0.000	1.000	0.167	0.373	0.000	1.000	0.244	0.429	0.000	1.000
south	0.392	0.488	0.000	1.000	0.322	0.468	0.000	1.000	0.383	0.486	0.000	1.000
west	0.271	0.445	0.000	1.000	0.333	0.472	0.000	1.000	0.211	0.408	0.000	1.000
obs	4295				264				33633			

The demographic characteristics of the household head also were included in this study. Both the employment status and educational attainment of the household head were represented as dummy or indicator variables. The variables *unemp, empparttime* and *empfulltime* were indicator variables representing whether the household head was unemployed, employed part-time or employed fulltime. The results indicate that for households choosing choice 1, almost 44% are employed fulltime whereas for those households under choice 2 more than 50% were also employed fulltime. For households with choice 3, approximately 43% were employed fulltime. Similarly the variables *edulths, eduhighschool, edusomecollege* and *educolleges* were utilized to describe whether the household head achieved educational attainment below high school, high school, above high school but below college and college and post-college. From the table more than half of the household in all three choices have some college units or have college or higher degrees. For example in choice 1, almost 79% of the households have college education whereas for those household who purchased only organic milk (choice 2) 65% alone comprise those heads which have college and higher degrees. Similar with choice 1, those who purchased only conventional milk (choice 3) have approximately 67 % of their household heads with college units and college plus degrees.

Also included into the model are the race and ethnicity of the household. The indicator variables *white, black, oriental* and *others* represented the racial household distinctions. The majority of the households of all the three choices are white households with choice 1 (76%), choice 2 (67%) and choice 3 (84%). On the other hand, household ethnicity was represented as hispanic (*hispyes*) and nonhispanic (*hispno*) and more than 90% in all of the three choices are non-Hispanic households. Finally, regional dummy variables were also included to describe the regional location of the household. The four major regional dummies that were created were *east, central south* and *west*. The number indicate that 39% of the households for choice 1 are from the south while those household that were under choice 2, approximately 33% were from the west. For choice 3, 38% of the households were from the south.

Empirical Results

Test of Independence of Irrelevant Alternatives (IIA)

A fundamental characteristic of the Multinomial Logit Model (MNL) is its assumption of the Independence of Irrelevant Alternatives (IIA) axiom. However, given that pairwise choice alternatives S and T are close substitutes, then the MNL model may produce inconsistent estimates. Consequently, if choices S and T are truly not independent, the MNL model may not be the optimal model to choose.

The Hausman-McFadden (1984) and Small-Hsiao (1985) tests involve pairwise comparisons of estimated coefficients of the full model vis-a- vis those estimates generated by restricted models where at least one choice alternative has been removed (Long and Freese, 2006). For these tests, the null hypothesis is whether alternatives S and T are independent of other alternatives. If the Chi-square statistic is significant, then the use of the MNL model is deemed inappropriate.

Table 2.2 presents the results of the two tests where the Hausman-McFadden test imply that both two choices failed to reject the null hypotheses and therefore use of the

MNL model is still valid. Notice that choice 2 have negative Chi-squared values. While implausible, these values are to be interpreted as not violating the IIA condition (Hausman-McFadden, 1984). On the other hand, the Small-Hsiao (1985) test results indicate Choice 1 rejecting the null hypotheses while Choice 2 failing to reject it. This situation implies that for Choice 1, IIA is not valid while for Choice 2, the independence axiom holds. Notice that both the Hausman-McFadden and Small-Hsiao test produced contradictory results. Apparently, these conflicting results from the Hausman-McFadden and Small-Hsiao tests were investigated by Cheng and Long (2007) by running Monte Carlo simulations on the size properties of these two tests. The study concluded that the Hausman-McFadden test results in poor estimates even if the sample size is larger than 1000 while the Small-Hsiao test performance were ambiguous with different data structures.

The study further concludes that these tests are inadequate in evaluating IIA validity or violations and note McFadden’s (1973) recommendation that care and valid judgment must be taken into account in using the MNL models especially if the partitioned choice outcomes are really distinct from each other. On the other hand, this exercise explicitly assumes *a priori* that the choices might not be distinct and therefore prompts us to use other models that would explicitly assume choice correlations (i.e. multinomial probit model).

Table 2.2. Hausman and Small Hsiao Tests for Independence of Irrelevant Alternatives (IIA) for a Multinomial Logit Model (MNL)

Omitted Choice[a]	Chi-Squared Statistic	df	P-value	Evidence
Hausman Test (n=38192)				
Choice1	4.5470	19	1.0000	Accept Ho
Choice2	-0.0370[b]	19	1.0000	Accept Ho
Small Hsiao Test (n=38192)				
Choice 1	44.809	19	0.001	Reject Ho
Choice 2	15.34	19	0.701	Accept Ho

Ho: Difference in the coefficients are not systematic

Note if Chi2 < 0 then the model does satisfy the asymptotic assumptions of the test.

[a] Since there are 3 alternatives in this model, 2 test variations are expected where omission of choice 1 results in the first restricted model and omission of choice 2 produces the second restricted model. Both the Hausman and Small-Hsiao tests compare the restricted models' coefficients with the full model where all choices are included.

[b], Hausman and McFadden (1984) opined that a possible negative result is evidence that IIA is not violated.

Multinomial Model Parameter Estimates

Table 2.3 presents three multinomial models namely the multinomial logit model and variants of the multinomial probit models. For the multinomial probit models, the variation comes from the various assumptions made about the error variance-covariance matrix. These variations include uncorrelated and unequal correlation of error terms.

The findings of the three models indicate that as the number of household size increases, the less likely that these households will purchase the combination of both organic and conventional milk (choice 1) and organic milk (choice 2) and this finding is readily apparent in choice 2 relative to choice 1. This situation implies that a single household is more likely to purchase both organic and a combination of organic and conventional milk relative to households with two, three, four and five or more household members. Household income although insignificant is positive throughout all models suggesting increasing likelihood of buying both organic and combination of organic and conventional milk. On the other hand households with children are less likely to buy both organic and combination of organic and conventional milk relative to households without children.

As for employment status, household heads that are employed fulltime are less likely to buy milk relative to those whose employment status is part time or not employed. The estimates for the household head’s level of education suggest a pattern indicating increasing likelihood of purchasing organic and combination of both milk types as educational level increases. As for race, the results show white households are

Table 2.3. Multinomial Logit and Probit[a] Estimated Coefficients and P-values of Fluid Milk Purchase

	Multinomial Logit		Multinomial Probit (Uncorrelated Error Terms)		Multinomial Probit (Unequal Correlation of Error Terms)	
Variables	Response 1[b] Coefficients	Response 2[c] Coefficients	Response 1 Coefficients	Response 2 Coefficients	Response 1 Coefficients	Response 2 Coefficients
inc	0.1712	0.0368	0.1289	0.0433	0.0448	0.0664
	(0.000)[d]	(0.647)	(0.000)	(0.272)	(0.060)	(0.053)
agepcchild	-0.0622	-0.5136	-0.0498	-0.2475	-0.0212	-0.2047
	(0.306)	(0.141)	(0.266)	(0.100)	(0.275)	(0.087)
Household Size						
hs2	-0.0566	-1.1466	-0.0532	-0.5863	-0.0368	-0.4941
	(0.206)	(0.000)	(0.106)	(0.000)	(0.100)	(0.000)
hs3	-0.0862	-1.4279	-0.0750	-0.7173	-0.4652	-0.5924
	(0.181)	(0.000)	(0.116)	(0.000)	(0.115)	(0.000)
hs4	-0.1876	-2.0069	-0.1565	-1.0130	-0.0776	-0.8593
	(0.016)	(0.000)	(0.006)	(0.000)	(0.087)	(0.000)
hsp5	-0.2069	-2.0398	-0.1694	-0.9729	-0.0819	-0.8049
	(0.023)	(0.000)	(0.011)	(0.000)	(0.094)	(0.000)
Employment of Family Head						
empparttime	0.1074	0.4930	0.0815	0.2307	0.0327	0.1601
	(0.028)	(0.011)	(0.025)	(0.016)	(0.130)	(0.037)
empfulltime	-0.1732	0.1917	-0.1292	0.0545	-0.0427	-0.0056
	(0.000)	(0.226)	(0.000)	(0.4780)	(0.076)	(0.932)
Education of Family Head						
eduhighschool	0.0375	0.9948	0.0252	0.3910	0.0119	0.2916
	(0.723)	(0.333)	(0.733)	(0.334)	(0.667)	(0.257)
edusomecollege	0.3435	2.0012	0.2462	0.8717	0.0948	0.6989
	(0.001)	(0.047)	(0.001)	(0.028)	(0.111)	(0.001)
educollegeplus	0.6197	2.6918	0.4607	1.2509	0.1777	1.0293
	(0.000)	(0.007)	(0.000)	(0.0020)	(0.078)	(0.000)
Race/Ethnicity						
white	-0.2025	-0.8757	-0.1564	-0.4300	-0.0634	-0.3535
	(0.017)	(0.008)	(0.015)	(0.009)	(0.093)	(0.017)
black	0.2350	-0.0054	0.1784	0.0561	0.0619	0.0789
	(0.014)	(0.988)	(0.014)	(0.752)	(0.155)	(0.600)
oriental	0.2843	0.2163	0.2283	0.2023	0.0822	0.1980
	(0.018)	(0.605)	(0.02)	(0.34)	(0.145)	(0.246)
hispyes	0.3057	0.2548	0.2376	0.1970	0.0840	0.2066
	(0.000)	(0.467)	(0.000)	(0.246)	(0.099)	(1.540)

Table 2.3 Continued

	Multinomial Logit		Multinomial Probit (Uncorrelated Error Terms)		Multinomial Probit (Unequal Correlation of Error Terms)	
Variables	Response 1[b] Coefficients	Response 2[c] Coefficients	Response 1 Coefficients	Response 2 Coefficients	Response 1 Coefficients	Response 2 Coefficients
Region						
central	-0.3774 (0.000)	-0.3885 (0.066)	-0.2730 (0.000)	-0.1994 (0.054)	-0.0970 (0.063)	-0.1615 (0.057)
south	-0.0327 (0.494)	-0.2528 (0.171)	-0.0250 (0.480)	-0.1226 (0.180)	-0.0109 (0.416)	-0.0934 (0.057)
west	0.1375 (0.008)	0.2908 (0.120)	0.1059 (0.006)	0.1909 (0.041)	0.0405 (0.113)	0.1825 (0.014)
Constant	-2.1369 (0.000)	-5.5852 (0.000)	-1.7395 (0.000)	-3.4493 (0.000)	-0.5923 (0.064)	-3.1550 (0.000)
Wald chi2(36)	1006.8700		983.0900		501.8000	
	0.0000		0.0000		0.0000	
log Pseudoliklihood	-1463.2930		-1460.7740			
log simulated Pseudoliklihood					-1458.9130	
obs	38192		38192		38192	

[a] Base outcome is response 3 (only conventional milk)

[b] Response 1 is purchase of both organic and conventional milk

[c] Response 2 is purchase of organic milk only

[d] values in parentheses are p-values

less likely to buy organic and conventional milk and organic milk only compared to black and oriental households while Hispanics are more likely to buy both organic and combination milk relative to non-Hispanics. The findings for regions are relatively the same for all models as households located in the west are more likely to buy strictly organic milk and combination of organic and combination milk. However, those households located in the South and Midwest are less likely to buy organic milk and a combination of organic and conventional milk.

Marginal Effects Analysis

Multinomial Logit Analysis

Looking at the multinomial logit model, as household incomes increase the purchase probability increases by 0.0162 and 0.0001 and decreases by 0.0163 if the milk purchase is combination, strictly organic milk and strictly conventional milk (Table 2.4). For the marginal effects of household size equal or greater than 5 members, the probability of purchasing declines by 0.0180 and 0.0035 and increase by 0.0215 respectively, in purchasing a combination of organic and conventional milk and strictly organic milk and strictly conventional milk. Also we find a similar trend with respect to presence of children in that, the probability of purchase declines by 0.0057 and 0.0016 and an increase of 0.0072 if the choice is to buy the combination of organic and conventional milk, strictly organic milk and strictly conventional milk.

Table 2.4. Marginal Effects of Multinomial Logit and Probit[1] Models of Fluid Milk Purchase by Organic, Conventional or Both

	Multinomial Logit (Uncorrelated Error Terms)			Multinomial Probit (Uncorrelated Error Terms)			Multinomial Probit (Unequal Correlation of Error Terms)		
Variables	Response 1	Response 2	Response 3	Response 1	Response 2	Response 3	Response 1	Response 2	Response 3
inc	0.0162	0.0001	-0.0163	0.0169	0.0001	-0.0170	0.0169	0.0002	-0.0171
	(0.000)	(0.817)	(0.000)	(0.000)	(0.761)	(0.000)	(0.000)	(0.447)	(0.000)
agepcchild	-0.0057	-0.0016	0.0072	-0.0060	-0.0017	0.0078	-0.0063	-0.0018	0.0080
	(0.314)	(0.104)	(0.204)	(0.293)	(0.079)	(0.182)	(0.880)	(0.733)	(0.852)
Household Size									
hs2	-0.0050	-0.0037	0.0086	-0.0059	-0.0044	0.0102	-0.0096	-0.0045	0.0141
	(0.239)	(0.000)	(0.042)	(0.169)	(0.000)	(0.018)	(0.700)	(0.720)	(0.646)
hs3	-0.0076	-0.0032	0.0109	-0.0087	-0.0037	0.0124	-0.0135	-0.0037	0.0172
	(0.189)	(0.000)	(0.063)	(0.145)	(0.000)	(0.039)	(0.734)	(0.748)	(0.691)
hs4	-0.0165	-0.0039	0.0203	-0.0185	-0.0043	0.0228	-0.0233	-0.0044	0.0277
	(0.013)	(0.000)	(0.002)	(0.007)	(0.000)	(0.001)	(0.632)	(0.754)	(0.587)
hsp5	-0.0180	-0.0035	0.0215	-0.0199	-0.0038	0.0238	-0.0250	-0.0038	0.0288
	(0.017)	(0.000)	(0.004)	(0.010)	(0.000)	(0.002)	(0.659)	(0.760)	(0.613)
Employment of Family Head									
empparttime	0.0103	0.0020	-0.0122	0.0104	0.0020	-0.0125	0.0113	0.0016	-0.0129
	(0.036)	(0.032)	(0.014)	(0.036)	(0.051)	(0.014)	(0.743)	(0.718)	(0.689)
empfulltime	-0.0164	0.0007	0.0156	-0.0170	0.0007	0.0163	-0.0166	0.0004	0.0162
	(0.000)	(0.188)	(0.000)	(0.000)	(0.263)	(0.000)	(0.645)	(0.757)	(0.656)
Education of Family Head									
eduhighschool	0.0031	0.0045	-0.0076	0.0024	0.0039	-0.0062	0.0011	0.0036	-0.0047
	(0.761)	(0.451)	(0.508)	(0.808)	(0.428)	(0.552)	(0.992)	(0.787)	(0.963)

Table 2.4 Continued

	Multinomial Logit (Uncorrelated Error Terms)			Multinomial Probit (Uncorrelated Error Terms)			Multinomial Probit (Unequal Correlation of Error Terms)		
Variables	Response 1	Response 2	Response 3	Response 1	Response 2	Response 3	Response 1	Response 2	Response 3
edusomecollege	0.0328	0.0112	-0.0441	0.0311	0.0098	-0.0409	0.0287	0.0094	-0.0381
	(0.003)	(0.220)	(0.001)	(0.003)	(0.147)	(0.000)	(0.846)	(0.735)	(0.755)
educollegeplus	0.0607	0.0166	-0.0773	0.0598	0.0150	-0.0748	0.0577	0.0145	-0.0722
	(0.000)	(0.139)	(0.000)	(0.000)	(0.061)	(0.000)	(0.754)	(0.708)	(0.623)
Race/Ethnicity									
white	-0.0197	-0.0040	0.0238	-0.0203	-0.0043	0.0246	-0.0216	-0.0040	0.0256
	(0.027)	(0.053)	(0.009)	(0.025)	(0.061)	(0.009)	(0.671)	(0.673)	(0.609)
black	0.0240	-0.0001	-0.0239	0.0249	0.0001	-0.0250	0.0249	0.0002	-0.0251
	(0.022)	(0.925)	(0.024)	(0.020)	(0.951)	(0.021)	(0.763)	(0.906)	(0.756)
oriental	0.0299	0.0007	-0.0306	0.0325	0.0014	-0.0338	0.0330	0.0015	-0.0345
	(0.033)	(0.690)	(0.030)	(0.026)	(0.554)	(0.022)	(0.740)	(0.771)	(0.716)
hispyes	0.0321	0.0008	-0.0329	0.0337	0.0013	-0.0349	0.0334	0.0016	-0.0350
	(0.000)	(0.572)	(0.000)	(0.000)	(0.474)	(0.000)	(0.694)	(0.757)	(0.664)
Region									
central	-0.0330	-0.0011	0.0342	-0.0333	-0.0010	0.0344	-0.0341	-0.0007	0.0348
	(0.000)	(0.065)	(0.000)	(0.000)	(0.151)	(0.000)	(0.564)	(0.730)	(0.543)
south	-0.0030	-0.0008	0.0039	-0.0030	-0.0009	0.0040	-0.0033	-0.0009	0.0042
	(0.505)	(0.166)	(0.397)	(0.510)	(0.188)	(0.394)	(0.906)	(0.721)	(0.880)
west	0.0133	0.0010	-0.0144	0.0139	0.0015	-0.0154	0.0142	0.0017	-0.0160
	(0.010)	(0.181)	(0.006)	(0.009)	(0.112)	(0.004)	(0.700)	(0.722)	(0.637)
Prob(Outcome)	0.1062	0.0035	0.8904	0.1075	0.0036	0.8889	0.1091	0.0035	0.8874

values in parentheses are p-values

As for the household head’s employment status, the findings indicate that for a household head that is employed full time, the probability of purchasing decreases by 0.0164 and increases by 0.0007 and 0.0156 respectively, if the purchase choice is combination, strictly organic milk and strictly conventional milk. If on the other hand the household head is employed part time, then the probability of purchasing a combination of organic and conventional milk and organic milk only increases by 0.0103 and 0.0020. However, the purchase probability decreases by 0.0122 if the milk purchase is conventional. On the other hand, if the household head education is college level we find that the purchase probability increases by 0.0607 and 0.0166 and decreases by 0.0773 respectively, if the choice purchase is combination, strictly organic and strictly conventional. The same purchase probability trends are observed if the household head is either a high school graduate or has some college level units.

As for race, if the household is white then the probability of purchase declines by 0.0197 and 0.0040 and increases by 0.0238 if the milk purchase is a combination, organic milk only or conventional milk. With regards to black and oriental households, the probability of purchase increases by 0.0240 and 0.0299 if the milk purchase is a combination of organic and conventional milk. If on the other the purchase is organic milk only, then the probability declines by 0.0001 if the household is black and increases by 0.0007 for an oriental household. Both purchase probabilities of black and oriental household decline by 0.0239 and 0.0306 if the milk purchase is conventional milk. For the ethnicity variable, the findings indicate that for hispanic households, the probability of purchasing organic milk and combination increases by 0.0321 and 0.0008 whereas the probability of purchasing conventional milk declines by 0.0329. As for regions, households in the west purchase probability increases by 0.0133 and 0.0052 and declines by 0.0144 if the purchase choice is combination milk, strictly organic milk and strictly conventional. On the other hand, for those households in the south, the probability of purchase declines by 0.0030 and 0.0008 and increases by 0.0039 if the purchase is combination of organic and conventional milk, strictly organic and strictly conventional. Relative to the south households, those located in the central region a have similar purchase probability trends.

Multinomial Probit Analysis

The marginal effects for the two multinomial probit variants seem to be close in both magnitude and signs relative to the multinomial logit model marginal effects. For the multinomial probit model with uncorrelated error terms, the closeness and same sign magnitudes relative to the multinomial logit may be attributed to the fact that the error terms are assumed to be independent standard normal random variables. The difference however of the said multinomial probit model is the relatively longer computation time to achieve convergence due to solving standard numerical integration as required by an error structure that is standard normal. Thus, this type of multinomial probit still assumes IIA.

There is little difference in estimated marginal effects generated by the multinomial logit model and the other multinomial probit variants. Thus, differences in the marginal effects can only occur if there is indeed a significant departure of both probability distributions. As argued by Dow and Endersby (2004), the relatively narrower confidence interval of the marginal effects estimates found in the multinomial logit model relative to its probit analog seem to justify the use of the multinomial logit model over its probit counterpart in terms of the confidence it generates. It should noted that for this exercise, the 95 percent confidence bands of the marginal effects estimates are narrower in the multinomial logit model relative to the multinomial probit estimates. Similarly, the work of Kropko (2008) strongly suggests that even when the independence of irrelevant alternative (IIA) axiom is severely violated, the multinomial logit model estimates provide more accurate results vis-à-vis those generated by the multinomial probit model.

Numerical Stability and Precision of Multinomial Logit (MNL) and Probit (MNP) Marginal Effects Estimates

When the respective multinomial model variants are compared, we find that little differences exist in the magnitudes of the marginal effects. However, the estimated marginal effects for the Multinomial Probit (unequal error correlation) are mostly insignificant. The standard errors generated by maximum simulated likelihood are larger relative to the other two cases. Following Dow and Endersby (2004), Greene (2008) and Judd (1998), we calculate the condition numbers of the three models respectively. The condition number (CI) is defined as the square root of the ratio between the largest and smallest eigenvalues (Greene, 2008). Likewise, Judd (1998) suggests a measure that can indicate numerical stability and accuracy. By taking the $\log_{10}$ (CI), indices that are less than or equal than 3 or 4 indicated numerical optimization stability while those greater than 10 imply instability.

Table 2.5 presents the conditional numbers and $\log_{10}$ (CI) values for the multinomial logit, multinomial probit (uncorrelated error terms) and multinomial probit (correlated error terms) models. Results show that the $\log_{10}$ (CI) value for the MN Probit error correlated variant is 8.36 while for the MN Logit and MN Probit uncorrelated error variant, the $\log_{10}$ (CI) values are approximately 2.32 and 2.045. This finding implies that the MN Logit and MN Probit uncorrelated error variant likelihood estimation procedure is numerically more stable and accurate than the MN Probit error correlated variant. This finding lends support to the notion that because of the inherent instability of likelihood estimation in the MN Probit error correlated variant, its estimated coefficients and/or standard errors are suspect, yielding greater likelihood of statistical insignificance for the estimated marginal effects.

Assessment of Predictive Capacity of the Organic Milk Multinomial Choice Model (Case of the Multinomial Logit and Probit Model)

We also examined the predictive capacity of both the multinomial logit and probit for organic milk (uncorrelated error term variant). Several studies including Park and Capps (1997) and Capps et al. (1999) have utilized prediction success tables in evaluating the predictive ability of multinomial/polychotomous choice models. In this approach, a successful prediction refers to a situation where both actual and predicted outcomes match in each of the outcome choices. To illustrate, suppose that the associated predicted probabilities of the ith household are as follows: choice 1 (0.2), choice 2 (0.3) and choice 3 (0.5). From the predicted values, the ith household should

Table 2.5. Conditional Indices and $\log_{10}$ (CI) of Multinomial Model Variants

	Multinomial Logit (Uncorrelated Error Terms)	Multinomial Probit (Uncorrelated Error Terms)	Multinomial Probit (Unequal Correlation of Error Terms)
Max Eigenvalue	4.0457	0.6275	0.46766937
Min Eigenvalue	0.00009396	0.00005104	0.0000000000000000087
Condition Number(CI)[a]	207.5031	110.8788	232131789.3914
$\log_{10}$(CI)[b]	2.3170	2.0448	8.3657

[a] The condition number is defined as the square root of the ratio between highest and lowest eigen values (Greene, 2008)

[b] The $\log_{10}$ (CI) provides a measure of numerical precision with numbers ≤ 3 indicating numerical stability and those > 10 showing potential instability (Judd, 1998 cited in Dow and Endersby, 2004)

choose outcome 3 because it has the highest probability and if the actual choice is indeed choice 3, then the model has made a correct prediction. Now if we sum all the correct predictions in all the choice outcomes and divide it by the total number of actual choices, then we get a measure of how successful the multinomial model is in making right predictions. Likewise, the ratio of a choice outcome's right predictions and its corresponding number of actual choices determines the model's ability to predict that particular outcome.

In this exercise however, an attempt was made to generate the usual prediction success table but was unsuccessful due to the dominant frequency of choice 3 (conventional milk). Almost all of the generated predicted probabilities pointed to choice 3 as the choice that should be chosen. This outcome however reduces the likelihood of having right predictions for choice 1 (organic and conventional milk) and choice 2 (organic milk only) and therefore constrains the ability of the model to correctly predict both choices 1 and 2. In order to circumvent this problem, we utilize the percentage of the observed frequencies of each choice as cutoff points in constructing the various conditions that will likely lead to the predicted choice of a particular outcome. The cutoff values are 0.112458 (Choice 1), 0.006912 (Choice 2) and 0.8806295 (Choice 3).

Denoting P(*xb1*), P(*xb2*) and P(*xb3*) as the predicted probabilities for choice 1 , 2 and 3, the following are conditions by which each of the 3 choices can be predicted :

Choice 1

$$P(xb1) \geq 0.112458 \,\&\, P(xb2) \prec 0.0069`12 \,\&\, P(xb3) \prec 0.8806295$$

$$P(xb1) \geq 0.112458 \,\&\, P(xb2) \geq 0.006912 \,\&\, P(xb3) \prec 0.8806295$$

$$P(xb1) \geq 0.112458 \,\&\, P(xb2) \prec 0.006912 \,\&\, P(xb3) \geq 0.8806295$$

Choice 2

$$P(xb1) \prec 0.112458 \,\&\, P(xb2) \geq 0.006912 \,\&\, P(xb3) \prec 0.8806295$$
$$P(xb1) \geq 0.112458 \,\&\, P(xb2) \geq 0.006912 \,\&\, P(xb3) \prec 0.8806295$$
$$P(xb1) \prec 0.112458 \,\&\, P(xb2) \geq 0.006912 \,\&\, P(xb3) \geq 0.8806295$$

Choice 3

$$P(xb1) \prec 0.112458 \,\&\, P(xb2) \prec 0.006912 \,\&\, P(xb3) \geq 0.8806295$$
$$P(xb1) \geq 0.112458 \,\&\, P(xb2) \prec 0.006912 \,\&\, P(xb3) \geq 0.8806295$$
$$P(xb1) \prec 0.112458 \,\&\, P(xb2) \geq 0.006912 \,\&\, P(xb3) \geq 0.8806295$$

Results indicate that for the multinomial logit case, the model predicts that approximately 19.16 percent of the time that choice 1 (organic and conventional milk) will be chosen. On the other hand, choice 2's (organic milk only) prediction is 8.4 percent. As for the last choice, the model predicts that 72.37 percent of the time, choice 3 (conventional milk) will be selected. Similar findings were also observed for the multinomial probit case where choice 1 is 19.27 percent, while for choice 2 is 8.64 percent and 72.08 percent for choice 3. Also, the results tend to favor the multinomial probit over the multinomial logit in having a higher prediction rate in choices 1 and 2. However, for choice 3, the multinomial logit model has a higher prediction probability relative to its counterpart multinomial probit model.

Conclusions and Implications

The findings of both models indicate that as the number of household member increases, the less likely that these households will purchase organic milk and combination of both organic and conventional milk. This result implies that a single household is more likely to purchase both organic and conventional milk relative to households with two, three, four and five or more household members. Household income is positive suggesting increasing likelihood of buying both organic and combination milk. On the other hand households with children are less likely to buy both organic and combination of organic and conventional milk relative to households without children. The estimates for the level of education of the household head indicate increasing likelihood of purchasing organic and combination of both as educational level increases. As for race, the results show that white households are less likely to buy compared to black and oriental household while Hispanics are more likely to buy both organic and combination milk relative to non-Hispanics. For regions, households located in the west are more likely to buy strictly organic milk and a combination of organic and conventional milk. As for employment status, household heads that are employed fulltime are less likely to buy milk relative to those whose employment status is part time or not employed.

This work provides input in designing marketing strategies that can target particular demographic groups such as single person, college educate household heads, oriental, Hispanic and western located households. We note that these findings represent the 2004 conditions and that a more current data set may further update recent behavioral changes with regards to the interplay between factors that affect organic and conventional milk purchase.

CHAPTER III

HOUSEHOLD DEMAND ANALYSIS OF ORGANIC AND CONVENTIONAL FLUID MILK IN THE UNITED STATES BASED ON THE 2004 NIELSEN HOMESCAN PANEL[*]

In recent years, consumer concerns have moved beyond issues of fat content (Gould, 1996) to issues related to the environment, genetically modified organisms (GMOs), health risks, and pesticide use. Recent trends in supermarkets offering healthier and natural food choices can been seen as a reaction to consumer concerns. The rapid expansion in the organic food market (Thompson, 1998) in particular has, in effect, triggered growth in the organic milk industry. Dairy products, along with fresh produce, were among the first organic products experienced by consumers (Demeritt, 2004). As reported by Dimitri and Venezia (2007), beginning in the early 1990s, the distribution of organic milk was mainly done through specialty shops and other small-scale operators. Currently, organic milk is available in nearly all food retail venues, including conventional supermarkets and mass merchandisers (e.g. Costco and Wal-Mart), implying wide distribution of the product within the last decade. Glaser and Thompson (2000) also observed that organic milk sold in gallons and pints barely registered any sales, but organic milk sold in half-gallon containers recorded impressive sales. Because

---

[*] Reprinted with permission from “Household Demand Analysis of Organic and Conventional Fluid Milk in the United States Based on the 2004 Nielsen Homescan Panel” by Pedro A. Alviola IV, and Oral Capps, Jr., in press. *Agribusiness: an International Journal*, Copyright[2009] by Wiley Periodicals, Inc., A Wiley Company

retail sales of organic milk have been growing since the mid 1990s, while overall sales of conventional milk have remained relatively constant over the same time period (Miller and Blayney, 2006), market shares for organic milk are on the rise. Organic milk currently constitutes about six percent of retail milk sales (Dimitri and Venezia, 2007).

Organic dairy is a rapidly growing market sector, offering opportunities for farmers to boost their incomes through conversion from commercial to organic production. Organic milk retails at premiums as high as 80 percent over conventional milk (Glaser and Thompson, 2000), while producers can accrue premiums of more than 40 percent over conventional prices (Organic Valley, 2005). For producers who are facing the decision of whether or not to invest in the conversion to organic production methods, it is crucial to have information on the prospects for the market, in particular, issues concerning consumer demand.

In this light, the objective of this research is to analyze household demand for organic milk and for conventional milk in the United States, addressing most of the limitations indigenous to previous research efforts. We wish to better understand the drivers of the demand for organic milk and for conventional milk, particularly own-price effects, cross-price effects, and income effects, as well as the effects of socio-demographic characteristics of households. Similar to the descriptive work done by Dimitri and Venezia (2007), we employ the Nielsen Homescan Panel in our analysis. Initially, we center attention on the factors affecting the decision to purchase organic milk and conventional milk at the household level.

Once the decision to purchase organic milk and conventional milk is made, we subsequently focus on factors affecting the amount purchased. Consequently, we identify the impacts of socio-demographic variables such as household size, the presence of children, employment status, education level, race, and ethnicity of the household head and region associated with the quantities of organic fluid milk and conventional milk purchased, and we estimate own-price, cross-price, and income elasticities for organic milk and conventional milk at the household level. In this way, we add to the store of knowledge in dealing with a formal econometric analysis of the demand for organic milk and conventional milk, by offering a micro-perspective at the household level across the United States.

Literature Review

Previous research on consumer demand for organic milk has made important contributions to the understanding of the market. For example, Bernard and Mathios (2005) find that consumers are willing to pay substantially more for organic milk and rBST-free milk than for conventional milk. Glaser and Thompson (2000), through the use of scanner data, find that purchases of organic milk are very sensitive to changes in prices. Dhar and Foltz (2005) considered demand interrelationships for rBST free milk, organic milk, and unlabeled (conventional) milk through the estimation of a Quadratic Almost Ideal Demand System (QUAIDS). The data indigenous to this analysis are weekly milk prices and sales for twelve U.S. cities over the period of March 9, 1997, to February 24, 2002. Findings revealed that rBST free milk and organic milk were complements, conventional milk and rBST free milk were substitutes, and conventional milk and organic milk were substitutes. Additionally, own-price elasticities for rBST free milk, organic milk, and conventional milk were estimated to be -4.40, -1.37, and -1.04, respectively. The limitations of this research were threefold: (1) because the analysis only covered 12 U.S. cities, it may not be representative of national demand patterns; (2) the period of the analysis may not reflect current market trends; and (3) the analysis did not deal with socio-demographic characteristics of individual consumers or households.

Dimitri and Venezia (2007) relied on the use of Nielsen Homescan data from 2004, with coverage of 38,375 households that purchase milk. The Nielsen Homescan data are a nationwide panel of households who scan their food purchases for home use from all retail outlets. Data include detailed product characteristics, quantities, and expenditures for each food item purchased by each household. The data are unique in that purchase information and demographic information about the households is available. In conducting descriptive analysis of the 2004 Nielsen data, they concluded that the typical consumer of organic milk is white, well-educated, and living in a household headed by someone younger than 50 years old. Further, households of all income levels purchase organic milk. Across ethnic groups, a higher share of Oriental, Hispanic, and “other” households purchase organic milk rather than conventional milk. The limitations of this research were twofold: (1) no formal statistical analysis of these data was conducted; and (2) no own-price, cross-price, or income elasticities were estimated.

Methodology

We plan to address most of the limitations of previous research efforts in analyzing household demand for organic milk and for conventional milk. Through the use of the 2004 Nielsen Homescan Panel, we employ the Heckman two-step procedure in this analysis.

Random Utility Model

The decision of whether or not to purchase organic milk can be modeled as a binary choice, wherein the outcome variable Y_i takes on two values: 1 with the occurrence of the event (purchase organic milk) or 0 (purchase conventional milk), $i = 1, 2, \ldots, n$, with n referring to the number of households in our sample[4]. With this specification, we can assume a utility function given as:

(1) $U(W_i, \varepsilon_i)$,

where utility is function of the covariates W_i involved in the decision process. Assuming that the utility function U exists, this choice problem can be represented as

(2) $U_1 = W_1\eta_1 + e_1$,

(3) $U_0 = W_0\eta_0 + e_0$,

where U_1 and U_0 are the utility levels associated with purchasing organic milk (U_1) and conventional milk (U_0); the disturbance terms e_1 and e_0 are random error components.

---

[4] Other choice possibilities also included households which purchase no milk at all and households which purchase both organic and conventional milk. The number of households which purchase no milk during calendar year 2004 was extremely small. The number of households purchasing only organic milk was 264, the number of households purchasing both organic milk and conventional milk was 4,295, and the number of households purchasing only conventional milk was 33,633. We were concerned with the decision to buy organic (conventional) milk or not over the entire year. Work is underway, in a separate analysis, in estimating a polychotomous choice model dealing with the aforementioned three choices.

For this exercise, we assume the *i*th household chooses to purchase organic milk (Y_i=1) because more utility is derived relative to the purchase of conventional milk (Y_i=0). Thus, if the *i*th household chooses to purchase organic milk, then $U_1 > U_0$ and consequently:

(4) $\Pr(Y_i = 1) = \Pr(U_1 > U_0)$,

(5) $\Pr(Y_i = 1) = \Pr(W_1\eta_1 + e_1 > W_0\eta_0 + e_0)$,

(6) $\Pr(Y_i = 1) = \Pr(e_0 - e_1 < W_1\eta_1 - W_0\eta_0)$, and

(7) $\Pr(Y_i = 1) = \Pr(\mu < W_1\eta_1 - W_0\eta_0)$.

Subsequently, if we assume that e_1 and e_0 are normally distributed, then the random variable $\mu = (e_1 - e_0)$ also is normally distributed. Consequently, $\Pr(Y_i = 1) = \Phi(W_1\eta_1 - W_0\eta_0)$, where Φ represents the cumulative distribution function (cdf). This relationship holds across all households, i = 1,…, n. Through standardization of μ, Φ then represents the standard normal cumulative distribution function. In this way, we justify the use of the probit model in investigating the decision to purchase organic fluid milk. Given the binary nature of the choice problem, we also justify the use of the probit model in investigating the decision to purchase conventional milk.

Hill and Lynchehaun (2002) identified various factors that influence consumer preferences in purchasing organic milk (Figure 3.1). Factors considered are grouped according to: (1) personal factors such as values and lifestyles; (2) intrinsic factors such as price and packaging; (3) cultural and social factors including age, ethnicity, and income; (4) knowledge factors; (5) extrinsic factors; and (6) uncontrollable factors. As

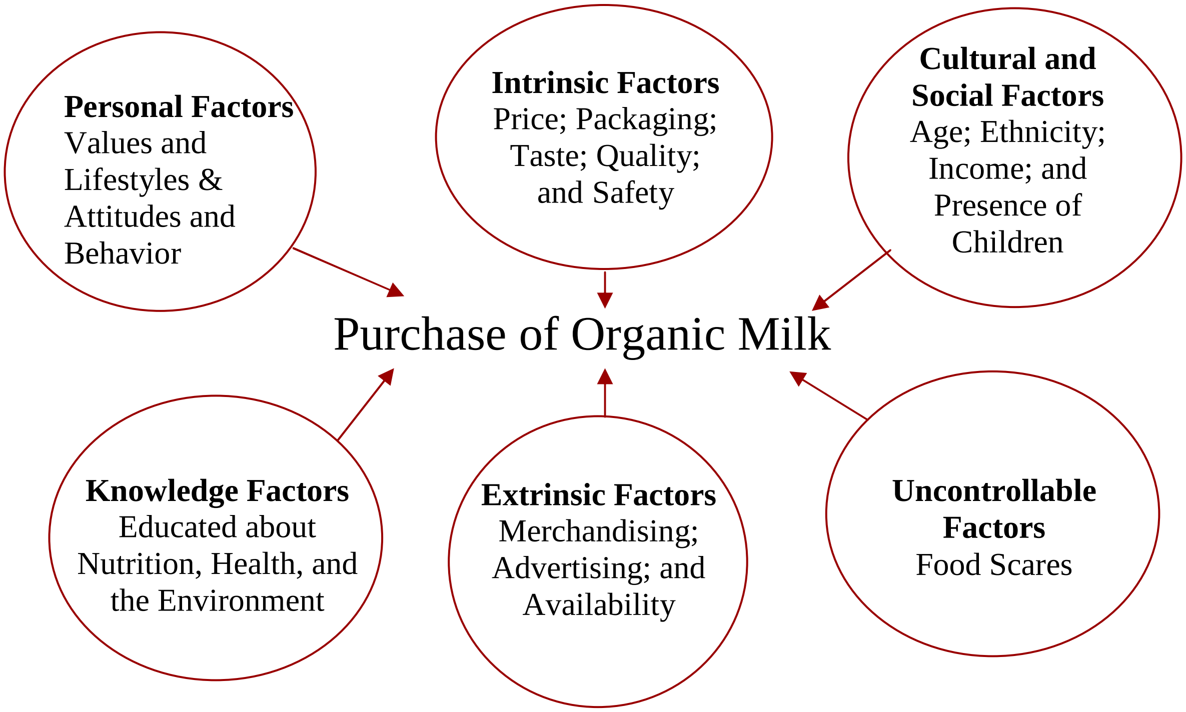


Figure 3.1 Factors that influence the consumer preference towards organic milk (Hill and Lynchehaun, 2002).

well, as stated previously, Dimitri and Venezia (2007) provided hypotheses regarding typical consumers purchasing organic milk.

Heckman Sample Selection Approach

Following Heckman (1976, 1979), the issue of sample selection bias may arise if we limit our sample to those households who purchase organic milk or to those households who purchase conventional milk. The remedy, as proposed by Heckman, is to use a two-step approach where the first-stage involves the usage of a binary choice specification (i.e. probit model) to account for the selection bias. In the second stage, we estimate the model using least squares, with the inclusion of the omitted variable, representing the selection bias, as an additional covariate or regressor.

There have been previous studies that have looked at censoring and sample selection issues in regard to estimating the demand for conventional fluid milk. Schmit et al (2002) utilized a two-step sample selection model based on a Nielson Homescan Panel of U.S. households from January 1996 through December 1999 in order to estimate at-home demand for fluid milk and cheese. Likewise, Dong et al., (2004b) examined milk purchasing behavior using a double-hurdle model, accounting for not only the censored nature of commodity purchases, but also for the dynamics of the purchase process. This work involved data from a panel of upstate New York households over the period 1996 to 1999. In our analysis, we consider only purchase patterns of organic and conventional milk over calendar year 2004.

First Stage of the Heckman Procedure

Using the probit model, we denote q_i as an indicator variable that takes on value of 1 if the *i*th household purchases organic milk and 0 if the *i*th household purchases conventional milk. Mathematically the probit model can be represented as:

(8) $\Pr(q_i = 1) = \Phi(W_i \eta)$, and

(9) $\Pr(q_i = 0) = 1 - \Phi(W_i \eta)$,

where Φ is the standard normal cumulative distribution function (cdf) and W_i is vector of variables that are related to the decision to purchase organic milk, similar to those described by Hill and Lynchehaun (2002) and by Dimitri and Venezia (2007). The corresponding vector of first-stage parameter estimates is represented by η_i. Thus, with $\phi(W_i \hat{\eta})$ as the calculated probability density function from this first-stage estimation, the Inverse Mill’s Ratio (IMR) can be calculated as

(10) $$IMR = \frac{\phi(W_i^T \hat{\eta})}{\Phi(W_i^T \hat{\eta})}$$

The IMR captures all the effects of the omitted variable regressor; hence the IMR is added to the set explanatory variables in the model in the second stage.

Second- Stage of the Heckman Procedure

In the second stage estimation, the demand equation for organic milk becomes:

(11) $$E(z_i^o \mid q_i = 1) = X_i \beta + \alpha \left[\frac{\phi(W_i^T \hat{\eta}_i)}{\Phi(W_i^T \hat{\eta}_i)} \right] + \mathrm{v}_{\mathrm{i}}, \text{ or}$$

(12) $$E(z_i^o \mid q_i = 1) = X_i \beta + \alpha IMR_i + \mathrm{v}_{\mathrm{i}},$$

where z_i^o is the quantity of organic milk purchased by the *i*th household, *W* represents the vector of variables related to the decision to purchase organic milk, and *X* constitutes the vector of explanatory variables related to the amount of organic milk purchased. Importantly, observations for which $q_i = 1$, i =1, 2, …, n_1 are used in the second-stage estimation—n_1 corresponds to the number of households who purchase organic milk. Following Saha, Capps, and Byrne (1997) and Greene (2008), let X_{ij} denote the *j*th regressor common to both W_i and X_i. The estimated marginal effect (ME) of a change in this regressor is given by:

$$(13) \quad \hat{M}E_{ij} = \beta_j + \alpha \frac{\partial IMR_i}{\partial X_{ij}}.$$

Thus, the ME is composed of two parts: a direct effect on the expected quantity of organic milk purchased, reflected by β_j, and a change in the IMR with respect to a unit change in X_{ij}. After some simplification, equation (13) can be rewritten as

$$(14) \quad \hat{M}E_{ij} = \hat{\beta}_j - \hat{\alpha} \cdot \eta_j (\hat{W}_i^T \eta \cdot I\hat{M}R_i + I\hat{M}R_i^2),$$

where:

$\hat{M}E_{ij}$ = marginal effect of the *j*th explanatory variable for the *i*th household,

$\hat{\beta}_j$ = parameter estimate associated with the *j*th explanatory in the second-stage of the model,

$\hat{\alpha}$ = parameter estimate associated with the IMR variable in the second stage of the model,

$\hat{\eta}_j$ = parameter estimate of the *j*th explanatory variable associated with the first-stage probit analysis,

$W_i^T\hat{\eta}$ = the prediction from the probit analysis for the *i*th household, and

$I\hat{M}R_i$ = the Inverse Mills Ratio for the *i*th household purchasing organic milk.

Equation (14) represents the appropriate expression in calculating the marginal effects associated with the Heckman two-step procedure. In general, $\hat{M}E_{ij} \neq \hat{\beta}_j$; the only cases where $\hat{M}E_{ij} = \hat{\beta}_j$ are as follows: (1) either $\hat{\alpha}$ is not statistically different from zero or (2) the *j*th explanatory variable in the second stage of estimation does not appear in the first-stage. Finally, since the estimated ME is observation-dependent, we propose to evaluate the marginal effects at the sample means.

Of note, the demand equation for conventional milk is quite similar to the specifications given in equations (11) and (12). In those equations, we replace $E(z_i^o \mid q_i = 1)$ with $E(z_i^c \mid q_i = 0)$, where z_i^c is the quantity of conventional milk purchased by the *i*th household. We replace *W* with *W** to represent the vector of variables related to the purchase of conventional milk. Further, we replace *X* with *X** to represent the vector of explanatory variables related to the amount of conventional milk purchased. Finally, we replace *IMR* with *IMR** to represent the inverse Mill's ratio in the demand equation for conventional milk. The number of households who purchase conventional milk is n_2, that is $i = 1, 2, \ldots, n_2$.

Empirical Specification

In this empirical exercise, the first-stage probit model specification is hypothesized to be a function of household income; employment status and education of the household head; race and ethnicity of the household head; region in which the household is located; and the presence of children (less than 18 years of age) in the household. The basis of this specification comes from the work of Hill and Lynchehaun (2002) as well as the work of Dimitri and Venezia (2007).

Mathematically, we write the probit specification for the decision to purchase organic milk as follows:

(15)

$$P(q_i = 1|W_i) = \eta_0 + \eta_1 Income_i + \eta_2 Hs2_i + \eta_3 Hs3_i + \eta_4 Hs4_i + \eta_5 Hs5_i + \eta_6 Agepchild_i +$$

$$\eta_7 Empparttime_i + \eta_8 Empfulltime_i + \eta_9 Eduhighschool_i + \eta_{10} Edusomecollege_i + \eta_{11} Educollegeplus_i +$$

$$\eta_{12} White_i + \eta_{13} Black_i + \eta_{14} Oriental_i + \eta_{15} Hisyes_i + \eta_{16} Central_i + \eta_{17} South_i + \eta_{18} West_i + \in_i$$ [5]

A description of the variable names in this specification is given in Table 1, along with their associated descriptive statistics. The majority of the explanatory variables are dummy or indicator variables. The reference categories to avoid the dummy variable trap are: (1) household size of 1, (2) no children under 18 years of age in the household, (3) the household head is unemployed, (4) the household head did not complete high school, (5) the household head is not white, black, or Oriental, (6) the

---

[5] We may also write mathematically the probit specification for the decision to purchase conventional milk as $P(q_i = 0|W_i*)$. The explanatory variables in this specification W* are the same as those in equation (15). Further, the parameter estimates in the specification are opposite in sign but are of the same magnitude.

household head is not Hispanic, and (7) the household is located in the East. By definition, the household head is the female head or the male head if no female head exists in the household.

Note that our specification does not include price as a potential explanatory variable influencing the decision to purchase organic milk. Prices are imputed as the ratio of expenditure to quantity in the Nielsen data; in essence, prices are unit values[6]. However, if organic milk is not purchased, it is not possible to derive the corresponding unit value. One can use other mechanisms in order to impute the missing prices, but we do not use additional imputation procedures in the probit analysis.

On the other hand, the second-stage specification deals with the amount of organic milk purchased, given that the decision to purchase was made.

Mathematically, we write the second-stage specification in the Heckman routine as:

$$(16)\, Q_{ji} = \beta_0 + \beta_1 Porg_{ji} + \beta_2 Pnonorg_{ji} + \beta_3 Income_{ji} + \beta_4 Hs2_{ji} + \beta_5 Hs3_{ji} + \beta_6 Hs4_{ji} + \beta_7 Hs5_{ji} + \beta_8 Agepchild_{ji} + \beta_9 Empparttime_{ji} + \beta_{10} Empfulltime_{ji} + \beta_{11} Eduhighschool_{ji} + \beta_{12} Edusomecollege_{ji} + \beta_{13} Educollegeplus_{ji} + \beta_{14} White_{ji} + \beta_{15} Black_{ji} + \beta_{16} Oriental_{ji} + \beta_{17} Hisyes_{ji} + \beta_{18} Central_{ji} + \beta_{19} South_{ji} + \beta_{20} West_{ji} + \beta_{21} IMR_{ji} + v_{ji}$$

, where

Q_{ji}, corresponds to the quantities of organic milk purchased (j =1) and conventional milk purchased (j = 2) respectively for the ith household; $Porg_{ji}$ and $Pnonorg_{ji}$ are the prices

---

[6] These calculated unit values may also reflect quality differences, and, consequently, the estimated income and price elasticities may be biased. However, we believe that the commodities involved are sufficiently disaggregated and homogeneous so as to minimize the degree of bias (Cox and Wohlgenant, 1986).

for unit values of organic milk and conventional milk, respectively, faced by the *i*th household. The rest of the variables are the same as those in the probit specification given by equation (12). Data pertaining to prices of soymilk or rBST-free milk were not available in the Nielsen Homescan panel. Consequently, these prices are excluded from our analysis[7].

Unlike the situation in the estimation of the probit model given by equation (15), equation (16) requires the use of price variables for both organic milk and for conventional milk. In the estimation of the second-stage demand equation for organic milk (conventional milk), we use only those observations for which purchases of organic milk (conventional milk) were made. Consequently, no imputation of own-price variables in the respective demand equations is necessary. However, for the cross-price variables in the respective demand equations, we need to impute these values. In cases when purchases of organic milk were made, households may not have purchased conventional milk and vice versa. Our imputation process in this analysis rests on the use of regional dummy variables: (1) when $Porg_{ji} = 0$, then

$Porg_{ji} = \exp[1.20705 - 0.09014 * Central_{ji} - 0.12081 * South_{ji} - 0.03836 * West_{ji}]$ and

(2) when $Pnonorg_{ji} = 0$, then

---

[7] The exclusion of rBST-free and soy milk prices may bias the parameter estimates and therefore affect the values of price elasticities of organic and conventional milk. The direction of the bias is difficult to ascertain. Based on the current literature on milk demand analysis, organic milk and rBST-free milk are complements while conventional milk and rBST free milk are substitutes (Dhar and Foltz, 2005). Likewise conventional milk and soymilk are complements (Dhar and Foltz, 2004).

$$Pnonorg_{ji} = \exp[0.56082 - 0.12828 * Central_{ji} + 0.02107 * South_{ji} - 0.00543 * West_{ji}].^{8}$$

Issues of Price Endogeneity

Because the prices in the analysis are unit values derived from the ratio of total expenditures to quantities purchased, there exists the possibility of price endogeneity (Dong, Shonkwiler and Capps, 1998). To determine the existence or nonexistence of price endogeneity in both the organic and conventional milk demand models, we conducted Hausman tests.

In conducting these tests, we identified socio-demographic variables such as household income, race, region and poverty status as instrumental variables (IV) for prices of organic and conventional milk. However, our data set corresponds to a cross-section of U.S. households, and, as such, the availability of valid instruments was severely limited if not lacking. Lewbel (1997), Nakamura and Nakamura (1998), and Park and Davis (2001), contended that if the chosen instruments were not highly correlated with the endogenous variable under investigation (prices in our case), then the IV estimator is biased and inefficient. Furthermore inference results generated from Hausman tests become suspect because the likelihood increases of accepting the null hypothesis of exogeneity as the instruments become less relevant (Nakamura and Nakamura, 1998, and Park and Davis, 2001). Thus, with severely limited instruments inherent in any data set, Ordinary Least Squares (OLS) estimates may be more appropriate to use relative to those generated by IV estimation.

---

[8] No problems of collinearity with the regional indicator variables and the respective price variables were evident. In addition to capturing price variation, region also may be capturing the effects of non-economic factors, such as environmental issues.

In performing the Hausman test, the first stage involved both regressions of organic and conventional milk prices as a function of income, race, region, and poverty levels. Further, the demand specification included the residuals of the first-stage estimation and F-tests were conducted to determine whether the coefficients corresponded to the residuals from the augmented regressions were statistically different from zero. Our findings indicated that endogeneity was not present in the organic milk demand relationship (p-value = 0.8647). However, for the conventional milk demand relationship, the hypothesis of price exogeneity (p-value=0.000) was rejected, which prompted the use of two-stage least squares (TSLS). Results from TSLS estimation for the conventional milk equation, however, indicated degrading collinearity patterns and non-significance of most of the estimated parameters, thus prompting the choice of OLS generated parameters. In keeping with Nakamura and Nakamura (1998), as well as Park and Davis (2001), given the limited instruments inherent in this cross-sectional data set, OLS estimates were deemed more appropriate than those generated by IV methods.

Data Description

For this empirical exercise, the data pertaining to the choice of purchasing organic milk, price and quantity of organic milk and conventional milk, income, and household socio-demographic variables are from the 2004 Nielsen Homescan Panel. Table 3.1 presents the definition and summary statistics of all the relevant variables considered in the analysis. For each household, we aggregate their purchases of organic milk and conventional milk over the entire calendar year.

The variable *Yesorg* (*Noorg*) is the dependent variable for the probit model and is defined as 1 to represent the purchase of organic (conventional) milk and 0 otherwise. Roughly 12 percent of the sample of households purchased organic milk sometime during the calendar year of 2004, and thus 88 percent of the sample of households purchased conventional milk during the 2004 calendar year.

The price and quantity variables of organic milk (*Porg, Qorg)* and conventional milk *(Pnonorg and Qnorg*) are standardized for a half gallon milk container. Most organic milk is sold by the half gallon (Glaser and Thompson, 2000), so we use the half gallon as the standard volume metric for this analysis. Conditional on making purchases, the average amounts of organic and conventional milk bought for calendar year 2004 were 9 and 47 half gallons, respectively. The average price paid for organic milk was $3.16 per half gallon and the average paid for conventional milk was $1.78 per half gallon. Consequently, there is a substantial premium paid for organic milk on the order of $1.38 per half gallon.

The average household income level of the sample is slightly above $50,000. Concerning household size, 26 percent of the sample consists of single-person households, while nearly 40 percent consists of two-person households. The proportions of households with three, four, and five or more members are 14 percent, 13 percent, and 8 percent, respectively. Additionally, households with children less than 18 years old (*Agechild*) are roughly 25 percent of the sample.

Table 3.1. Summary Statistics of Variables Used in the Analysis

Variable	Description	Observation	Mean	Std. Dev.	Min	Max
Yesorg (q_i = 1)	Household purchased organic milk	38,192	0.119	0.324	0	1
Noorg (q_i = 0)	Household did not purchase organic milk	38,192	0.881	0.324	0	1
Qorg	Quantity of organic milk purchased (half gallons)	4,559	8.798	14.798	0.5	293
Qnorg	Quantity of conventional milk purchased (half gallons)	37,928	46.739	42.641	0.25	1011
Porg	Price of organic milk (half gallons)	4,559	3.155	0.541	2.12	4.58
Pnonorg	Price of conventional milk (half gallons)	38,192	1.780	0.541	0.99	4.36
Income	HH income	38,192	50,024	27,306	5,000	100,000
Hs1	HH size of 1[a]	38,182	0.262	0.440	0	1
Hs2	HH size of 2	38,192	0.391	0.488	0	1
Hs3	HH size of 3	38,192	0.143	0.350	0	1
Hs4	HH size of 4	38,192	0.127	0.333	0	1
Hs5	HH size > 4	38,192	0.077	0.267	0	1
Agepcchild	HH has at least 1 child less than 18 yrs of age	38,192	0.253	0.435	0	1
No children	HH has no children less than 18 years of age	38,192	0.747	0.435	0	1
Unemployed	Head of HH is unemployed	38,192	0.408	0.491	0	1
Empparttime	Head of HH is employed part-time	38,192	0.157	0.364	0	1
Empfulltime	Head of HH is employed full-time	38,192	0.435	0.496	0	1
Edulths	HH head completed less than 12 years of schooling[a]	38,192	0.038	0.192	0	1
Eduhighschool	HH head is high school graduate	38,192	0.275	0.446	0	1
Edusomecollege	HH head has completed some college	38,192	0.320	0.446	0	1
Educollegeplus	HH head has at least a college education	38,192	0.367	0.482	0	1
White	HH head is white	38,192	0.825	0.380	0	1
Black	HH head is black	38,192	0.096	0.295	0	1
Oriental	HH head is Oriental	38,192	0.022	0.146	0	1
Other	HH head is classified as other[a]	38,192	0.057	0.232	0	1
Hispyes	HH head is Hispanic	38,192	0.066	0.248	0	1

Table 3.1 Continued

Variable	Description	Observation	Mean	Std. Dev.	Min	Max
Hispno	HH is not hispanic[a]	38,192	0.934	0.248	0	1
East	HH is located in the East[a]	38,192	0.163	0.370	0	1
Central	HH is located in the Midwest	38,192	0.235	0.424	0	1
South	HH is located in the South	38,192	0.384	0.486	0	1
West	HH is located in the West	38,192	0.219	0.413	0	1

Source: Nielsen Home Scan Panel for Calendar Year 2004

HH denotes household; the HH head is defined as the female head. If a female head of household does not exist, then the HH head is the male head.

[a] Reference category so as to avoid the dummy variable trap.

Demographic characteristics of the household head also are included in this analysis. Both the employment status and educational attainment of the household head are represented as dummy or indicator variables. The variables *Unemp, Empparttime,* and *Empfulltime* are indicator variables representing whether the household head is unemployed, employed part-time, or employed full-time. Roughly 60 percent of household heads are employed either part-time or full-time. Similarly the variables *Edulths, Eduhighschool, Edusomecollege,* and *Educollegeplus* are utilized to describe whether the household head completed less than a high school education, was a high school graduate, completed some college, or obtained at least an undergraduate degree. Nearly 70 percent of the sample had at least some college, while slightly more than 25 percent completed high school but not attended college.

Also included into the model are race and ethnicity of the household. The indicator variables *White, Black, Oriental,* and *Other* represent the major racial household distinctions. About 83 percent of the sample is classified as white, 10 percent is classified as black, and slightly more than 2 percent is classified as Oriental. Household ethnicity is represented as either Hispanic (*Hispyes*) or non-hispanic (*Hispno*). About 7 percent of our sample is classified as Hispanic. Finally, dummy variables labeled as *East*, *Midwest, South*, and *West* are included to describe the regional location of the household. The majority of the households are located in the South (38.4 percent), followed by the Midwest (23.5 percent), West (21.9 percent), and East (16.3 percent).

Empirical Results

First-Stage Analysis: Probit Model

The maximum likelihood estimates of the parameters and the accompanying estimates of the marginal estimates of the first-stage probit model analysis are provided in Table 3.2. From the Wald chi-squared statistic, at least one of the coefficients associated with the set of explanatory variables is statistically significant despite the magnitude of pseudo R^2 (McFadden R^2 statistic) of 0.029. This magnitude of the measure of goodness-of-fit is not atypical in probit models.

From Table 3.2, as the number of household member increases, it is less likely that households will purchase organic milk. Hence a single-person household is more likely to purchase organic milk relative to households with two, three, four, and five or more members. Looking at the marginal effects, we find that for household size equal to or greater than 5 members, the probability of purchasing organic milk is less by 0.0293, relative to a single household. For other household size categories, the probability of purchasing organic milk is less by 0.0283 for *Hs4*, 0.0178 for *Hs3*, and 0.0146 for *Hs2*. On the other hand, as household income increases, the likelihood of purchasing organic milk is greater. The presence of children in the household is not a statistically significant factor affecting the likelihood of purchasing organic milk. Household heads employed part-time are more likely to purchase organic milk relative to unemployed heads. This probability is higher by 0.0130. On the other hand, household heads employed full-time are less likely to purchase organic milk relative to unemployed household heads. This probability is lower by 0.0159 relative to those who are unemployed.

Table 3.2. Parameter and Marginal Effects Estimates of Probit Analysis of Organic Milk Choice[a]

Variable	Estimates	(P>|z|)	Marginal Effects	(P>|z|)
Hs2	-0.0768	0.0010	-0.0146	0.0010
Hs3	-0.0968	0.0040	-0.0178	0.0020
Hs4	-0.1589	0.0000	-0.0283	0.0000
Hs5	-0.1673	0.0000	-0.0293	0.0000
Income	3.27E-06	0.0000	6.26E-07	0.0000
Agepcchild	-0.0429	0.1740	-0.0081	0.1680
Empparttime	0.0659	0.0090	0.0130	0.0110
Empfulltime	-0.0837	0.0000	-0.0159	0.0000
Eduhighschool	0.0245	0.6380	0.0047	0.6410
Edusomecollege	0.1908	0.0000	0.0381	0.0000
Educollegeplus	0.3555	0.0000	0.0721	0.0000
White	-0.1292	0.0040	-0.0260	0.0060
Black	0.1215	0.0170	0.0246	0.0240
Oriental	0.1619	0.0130	0.0339	0.0230
Hispyes	0.1673	0.0000	0.0349	0.0000
Central	-0.1933	0.0000	-0.0348	0.0000
South	-0.0222	0.3710	-0.0042	0.3690
West	0.0807	0.0030	0.0159	0.0040
Constant	-1.3431	0.0000		
McFadden R^2	0.029			
Number of Observations	38,192			
Wald Statistic (18)	800			
p-value	0.000			
Wald Tests Joint tests of hypotheses associated with the indicator variables	Chi-squared statistic	p-value		
(1) Hs2=Hs3=Hs4=Hs5=0	20.42	0.0004		
(2) Empparttime= Empfulltime=0	40.09	0.0000		
(3) Eduhighschool=Edusomecollege= Educollegeplus=0	208.42	0.0000		
(4) White=Black=Oriental=0	114.35	0.0000		
(5) Central=South=West=0	113.95	0.0000		

---

[a] The exact same magnitudes of parameter estimates are obtained in the probit analysis of conventional milk choice. However, the signs of the respective coefficients are reversed.

The level of education of the household head plays an important role in the purchase of organic milk. From Table 3.2, as the educational level of the household head increases, the probability of purchasing organic milk increases. For household heads with at least a college level education, the probability of purchasing organic milk increases by 0.0721 relative to household heads with less than a high school education. For those households with educational levels corresponding to some college, the likelihood of buying organic milk increases by 0.0381 relative to household heads with less than a high school education.

Hispanic households are more likely to purchase organic milk relative to non-Hispanic households. The likelihood of purchase of Hispanic households (*Hisyes*) increases by 0.0349 relative to non-hispanic households. Black and Oriental households are more likely to purchase organic milk relative to other race types. For the black and Oriental households, the probability of purchasing organic milk is higher by 0.0246 and 0.0339, respectively, relative to other race types. For white households, the probability of purchasing organic milk decreases by 0.0260 relative to other types. Consequently, white households are the least likely to purchase organic milk, controlling for other socio-economic and demographic factors.

Finally, for the regional indicator variables, the findings indicate that households located in the West are more likely to purchase organic milk, while those located in the Midwest are least likely to purchase organic milk. For households located in the West, the probability of purchasing organic milk increases by 0.0159 relative to households located in the East. For households located in the Midwest, the probability of purchasing organic milk decreases by 0.0348 relative to those in the East.

These results are consistent with the findings of the sparse literature. For example, according to Pittman (2004) and Dong et al (2004b), blacks are less likely to consume conventional milk than other races, while Hispanic households are more likely to consume conventional milk. Dong et al (2004b) also found that household size was positively correlated with the likelihood of purchasing conventional milk.

Assessment of Predictive Capacity of the Probit Choice Model

A prediction success table is used to assess the usefulness of the probit model. Several studies (Park and Capps, 1997; Capps et al., 1999) use this approach in evaluating qualitative choice models. In generating the appropriate classification values, we use a cut-off value equal to 0.119 instead of the default 0.500[9]. This value corresponds to the ratio of the total number of households purchasing organic milk to the total number of households in the sample, that is, the market penetration. From Table 3.3, the percentage of correct predictions is approximately 0.58.

In short, using our decision rule or cut-off probability of 0.119, the model is correct 58 percent of the time in predicting choices for both organic milk and conventional milk, respectively. In terms of sensitivity or the ability to correctly predict the decision to purchase organic milk, the model is correct approximately 61% of the time. On the other hand, in terms of specificity or the ability of the model to correctly

---

[9] If the 0.5 default value is used instead of the market penetration of organic milk, then the model is not be able to correctly classify any households that purchased organic milk.

Table 3.3. Prediction-Success Table: Choices of Organic Milk and Conventional Milk

	Actual Choice		
Predictions	Organic Milk	Conventional Milk	Total
Organic Milk	2,772	14,266	17,038
Conventional Milk	1,787	19,367	21,154
Total	4,559	33,633	38,192
Percentage of Right Predictions (%)	57.97		
Sensitivity(%)[a]	60.80		
Specificity(%)[b]	57.58		
Cutoff value	0.119		

[a] The percentage of correctly predicting the choice of choosing organic milk (2,772/4,559)

[b] The percentage of correctly predicting the choice of choosing conventional milk (19,367/33,633)

classify the decision to purchase conventional milk, the model is correct approximately 58% of the time.

Second Stage Analysis: Estimation of Demand Equations[10]

For the second stage, estimation of the two demand equations for organic and conventional milk is performed using least squares. For the organic milk demand equation, the goodness-of-fit statistic is 0.074, while for the conventional milk demand equation, the goodness-of-fit ratio is 0.226. The parameter estimates and the associated p-values are exhibited in Table 3.4. Note, however, that in the organic milk demand equation the variable *Invmills* (inverse mills ratio) is statistically significant at the 0.05 level (p-value=0.0300), indicating evidence of sample selection bias. Thus, for explanatory variables common to both the probit (first-stage equation) and the second-stage equation, the parameter estimates are not the appropriate marginal effects. However, for the conventional milk demand model, the inverse mills ratio (p-value=0.4540) is statistically insignificant; hence sample selection bias is not evident. Thus, the estimated coefficients in the second-stage equation for conventional milk correspond to the appropriate marginal effects.

Second-Stage Results for Organic Milk

For the second-stage estimation, once the decision to purchase organic milk has been made, from Table 3.4, holding other things constant, for every unit increase in the price of organic milk the quantity purchased of organic milk declines by 5.6 half gallons.

---

[10] Attempts were made to estimate the first and second-stage equations simultaneously. However, the estimation routine failed to converge. Consequently, the estimation of the Heckman two-step procedure is done sequentially. The software package used in this analysis was STATA 9.2.

Table 3.4. Second Stage Parameter Estimates of Demand Analysis of Organic and Conventional Milk

Variable	Organic Milk	P>|t|	Marginal Effects	Conventional Milk	P>|t|
Porg	-5.5893	0.0000		2.6621	0.0100
Pnonorg	3.4728	0.0000		-22.9181	0.0000
Invmills	-128.3534	0.0300		16.2653	0.4540
Income	-0.0003	0.0430	0.00005	-0.00001	0.6160
Hs2	9.3840	0.0110	0.9698	13.3728	0.0000
Hs3	12.2548	0.0080	1.6492	18.8843	0.0000
Hs4	19.4247	0.0110	2.0185	26.1841	0.0000
Hs5	17.5322	0.0290	-0.7872	32.1777	0.0000
Agepcchild	5.8490	0.0080	1.1541	5.2759	0.0000
Empparttime	-7.5255	0.0180	-0.3099	-1.8289	0.0210
Empfulltime	8.2172	0.0410	-0.9537	-5.5228	0.0000
Eduhighschool	-0.7233	0.6600	1.9627	-0.9707	0.4110
Edusomecollege	-19.1697	0.0440	1.7218	-3.6604	0.0250
Educollegeplus	-33.9874	0.0490	4.9483	-3.2492	0.2260
White	16.3271	0.0090	2.1750	5.0256	0.0000
Black	-12.8049	0.0290	0.4969	-15.3822	0.0000
Oriental	-15.4636	0.0370	2.2629	-7.1158	0.0010
Hispyes	-15.8197	0.0470	2.4986	-3.1379	0.0550
Central	21.5736	0.0260	0.4058	0.5385	0.7060
South	2.2439	0.0820	-0.1881	2.3717	0.0010
West	-7.0075	0.0660	1.8313	-3.0126	0.0010
Constant	243.8373	0.0230		63.7563	0.0000
R-squared	0.074			0.226	
Number of Observations	4,559			37,928	
F(21, 4537)	19.23			617.63	
Prob > F	0.000			0.000	

Table 3.4 Continued

F-tests	Organic Milk		Conventional milk	
Joint tests of hypotheses associated with the indicator variables	F-value	Prob > F	F-value	Prob > F
Hs2=Hs3=Hs4=Hs5=0	3.98	0.0032	108.86	0.0000
Empparttime= Empfulltime=0	3.11	0.0466	36.00	0.0000
Eduhighschool=Edusomecollege= Educollegeplus=0	4.94	0.0020	6.41	0.0002
White=Black=Oriental=0	3.23	0.0214	73.19	0.0000
Central=South=West=0	3.81	0.0097	25.16	0.0000

On the other hand, a unit increase in the price of conventional milk translates to increases in the purchase of organic milk by almost 3.5 half gallons.

Relative to single-person households, two-person households purchase almost 1 more half gallon of organic milk annually; three-person households purchase 1.65 more half gallons of organic milk annually; and four-person households purchase 2 more half gallons of organic milk annually. However, for households with five or more persons, annual purchases of organic milk are lower by 0.8 half gallons relative to single-person households. Households with children less than 18 years of age purchase almost 1.2 more half gallons of organic milk annually relative to households with no children less than 18 years of age. However, household heads who are employed either part-time or full-time annually purchase less organic milk, on the order of 0.3 to 1 half gallons, relative to households with heads that are unemployed.

However, the reverse is true regarding educational levels of household heads. Relative to household heads who have less than a high school education, those with a high school education purchase almost two more half gallons of organic milk annually; those with some college education purchase 1.7 more half gallons of organic milk annually. Additionally, those with at least an undergraduate education purchase nearly 5 more half gallons of organic milk annually relative to those household heads with less than a high school education. As for race, whites and Orientals purchase roughly 2.2 more half gallons of organic milk annually relative to other races. Hispanics buy more than 2.5 half gallons of organic milk annually relative to non-Hispanics. Regionally, marked differences exist in the volumes of organic milk purchased. Relative to households located in the East, those located in the West buy more than 1.8 half gallons of organic milk annually; households located in the Midwest buy 0.4 more half gallons of organic milk annually than households located in the East; but households located in the South buy almost 0.2 half gallons less annually than households located in the East.

Second Stage Results for Conventional Milk

A unit increase in the price of conventional milk translates to a decline of approximately 23 half gallons of conventional milk, while an increase in the unit price of organic milk leads to an increase in purchase of conventional milk by almost 2.7 half gallons. Also, the presence of children in the household translates to increased purchases of conventional milk by roughly 5.3 half gallons annually. For household size, the purchases of conventional milk increase as number of household members increase. To illustrate, two-person households buy 13.4 more half gallons of conventional milk annually relative to single-person households; three-person households purchase almost 19 more half gallons of conventional milk annually, relative to single-person households. For four-person households, the gap is 26 more half gallons annually, and for five or more person households the gap is 32 half gallons annually.

Similar to the findings for organic milk, employed household heads purchase less conventional milk than unemployed household heads. The difference is between 1.8 and 5.5 half gallons annually, depending if household heads are employed part-time or full-time. In contrast with the findings for organic milk, purchases of conventional milk decline as the level of education increases. Concerning race, whites purchase more conventional milk relative to other races; blacks and Orientals purchase less conventional milk relative to other races. Hispanic households purchase less conventional milk than non-Hispanic households. Households located in the South buy more conventional milk relative to households located in other regions, while households located in the West buy less conventional milk relative to households located in other regions.

Elasticity Estimates[11]

We now present estimates of own-price, cross-price and income elasticities of both organic and conventional milk (Table 3.5). The standard definition of price elasticity is the percent change in the quantity demanded brought about by a one-percent change in price. Using this definition, we find that the own-price and cross-price elasticities of organic milk are -2.00 and 0.70 respectively. These numbers imply that a one-percent increase (decrease) in the price of organic milk translates to a 2.00 percent decline (rise) in the quantity demanded for organic milk. On the other hand, if the price of conventional milk increases (decreases) by one-percent, the quantity demanded for organic milk increases (decreases) by 0.70 percent. The income elasticity for organic milk is approximately equal to 0.27, which implies that a one-percent increase in household income leads to nearly a 0.30 percent increase in quantity demanded for organic milk.

---

[11] Tomek and Robinson's (2003) formula of total elasticity is $T_i = E_{ii} + E_{ij} * S_{ji}$ where E_{ii} and E_{ij} are the own price and cross price elasticties and S_{ji} represents the elasticities of "price transmission". The concept behind the formula denotes that a change in price of say commodity *i* will result in changes in prices of other commodities as well (*mutatis mutandum*). We assume that changes in the price of organic milk do not affect the price of conventional milk and vice versa. Also, in calculating the income elasticities we abstract from price rationing and assume perfect competition in supply.

Table 3.5. Price and Income Elasticity Estimates for Organic and Conventional Milk[a]

Variable	Organic Milk	Conventional Milk
Own-Price Elasticity	-2.0046	-0.8729
Cross-Price Elasticity	0.7027	0.1797
Income Elasticity	0.2672	-0.0135

[a] elasticities are computed at the sample means.

On the other hand, the own-price, cross-price, and income elasticities for conventional milk are -0.87, 0.18, and -0.01, respectively. The interpretations are the same as that of the organic case. The demand for organic milk is elastic, but the demand for conventional milk is inelastic. The sensitivity to own-price changes of organic milk is at odds with the findings of Hammarlund (2001), who finds that, on average, consumers are willing to pay up to five times the price of conventional milk to buyorganic milk. Owing to the positive cross-price elasticities, evidence indicates that organic and conventional milk are substitutes. Evidence also seems to indicate that organic milk is a necessity (income elasticity estimated to be 0.27), but conventional milk is an inferior good (income elasticity estimated to be -0.01). Dhar and Foltz (2005) estimated own-price elasticities as follows: rBST-free milk (-4.40), organic (-1.37), and conventional milk (-1.04). Dhar and Foltz (2005) also found that both rBST-free and organic milk were substitutes for conventional milk. Our own-price elasticity estimates for organic milk (-2.00) and for conventional milk (-0.87) differ significantly from those by Dhar and Foltz (2005)[12].

Our cross-price elasticity estimates indicate that a one-percent change in the price of conventional milk leads to a 0.70 percent change in the quantity demanded for organic milk, whereas a one-percent change in the price of organic milk results in a 0.18 percent change in the quantity demanded for conventional milk. This asymmetric pattern in the respective cross-price elasticities as suggested by Dhar and Foltz (2005) may be

---

[12] Statistical tests were performed to consider whether our elasticity estimates were different from those elasticities generated by Dhar and Foltz (2005). In looking at comparisons of own-price and cross-price elasticities, we reject in all cases the equivalence of our estimates with those of Dhar and Foltz (2005).

attributed to the difficulty or unwillingness of consumers to switch back from a high-quality product to a relatively lower-quality product, even if there are notable price changes. The cross-price elasticity of organic milk with respect to conventional milk was estimated to be 3.15 by Dhar and Foltz (2005). The cross-price elasticity of conventional milk with respect to organic milk was estimated to be 0.02 by Dhar and Foltz (2005). Our estimates of the cross-price elasticities are significantly different from those of Dhar and Foltz (2005).

To highlight the importance of generating elasticities, we also calculated the effect of a one-percent increase in the price of organic milk and the price of conventional milk on *total* milk sales. Using Dimitri and Venezia's (2007) calculated organic milk expenditure share of 0.32, our results show that a one-percent increase in the price of organic milk translates to a 0.20 percent decrease in total milk sales. Likewise, a one-percent increase in the price of conventional milk translates to a 0.31 percent increase in total milk sales.[13] The effects in each case, however, are modest.

Implications, Conclusions and Limitations

The findings from the probit analysis indicate that single-person households are more likely to purchase organic milk relative to other households with more family

---

[13] The basis of this calculation is as follows. Letting TR (total revenue) = P_1Q_1 (total revenue from organic milk) + P_2Q_2 (total revenue from conventional milk),

$$\frac{\partial TR}{\partial P_1} = P_1\frac{\partial Q_1}{\partial P_1} + Q_1 + P_2\frac{\partial Q_2}{\partial P_1}$$

, which implies that the percentage change in total revenue due to a one-percent change in P_1 may be expressed as $V_1G_{11} + V_1 + V_2G_{21}$, where V_1 is the expenditure share of organic milk, V_2 is the expenditure share of conventional milk, G_{11} is the own-price elasticity of organic milk, and G_{21} is the cross-price elasticity of conventional milk with respect to organic milk. The percentage change in total revenue due to a one-percent change in P_2 similarly may be expressed as $V_2G_{22} + V_2 + V_1G_{12}$, where G_{22} is the own-price elasticity of conventional milk and G_{12} is the cross-price elasticity of organic milk with respect to conventional milk.

members. Likewise, affluent households are more likely to purchase organic milk, and household heads with some college are more likely to purchase organic milk than heads of households with lower levels of education. In terms of region, households located in the West are the most likely to purchase organic milk, and those in the Midwest are the least likely to purchase organic milk. The presence of children in the household may reduce the likelihood of purchasing organic milk relative to those with no children. As for race, black and Oriental households are most likely to purchase organic milk and white households are least likely to purchase organic milk. Finally, Hispanic households are more likely to buy organic milk than households that are non-hispanic. Thus, from these demographic profiles, we find that variables such as household size, number of children, employment status and education of household head, race, ethnicity, and region have a significant effect on the likelihood of purchasing organic milk.

However, once the decision to purchase either organic milk or conventional milk has been made, our findings indicate that as household size increases, purchases of both organic and conventional milk increase. The presence of children in the household also leads to increases in the purchase of both milk types. However, as the level of education increases, purchases of organic milk rise but purchases of conventional milk fall. Whites and Orientals purchase more organic milk than other races; whites also buy more conventional milk, but blacks and Orientals buy less conventional milk. Hispanic households purchase more organic milk but less conventional milk than non-Hispanic households. Finally, households located in the West purchase the most organic milk relative to other regions, whereas households located in the South purchase more conventional milk than households located in other regions. Our second-stage results concerning the impacts of socio-demographic factors on purchases of organic milk largely are in agreement with those found by Dimitri and Venezia (2007).

From the estimated elasticities, we find that organic and conventional milk are substitutes, although an asymmetric pattern exists in this relationship. The demand for organic milk is more sensitive to changes in the price of conventional milk, but the demand for conventional milk is not very sensitive to changes in the price of organic milk. Additionally, the demand for organic milk is elastic but the demand for conventional milk is inelastic. Finally, organic milk technically is a necessary good but conventional milk is an inferior good.

The results from our work will enhance marketing efforts of organic milk in targeting particular demographic groups, particularly college-educated households, households located in the West, and Hispanic households. Also, owing to our findings concerning own-price elasticities, retailers should lower the prices of organic milk but raise prices of conventional milk in order to increase sales revenue, holding all other factors constant. As well, increases in the prices of conventional milk, all other things equal, will lead to increases in purchases of organic milk.

The major limitation of our analysis is that we provide only a snapshot of the organic and conventional milk market in 2004. Whether this demand picture continues to hold in the future is a function of the interplay among retailers, the supply of milk from organic and conventional dairies, and the socio-demographic characteristics of the population. A replication of our analysis with more recent data certainly is worthwhile to monitor demand patterns for organic milk and for conventional milk. Further, in future work, attention should be centered on household choices of buying organic milk only, conventional milk only, or buying both organic and conventional milk. Finally, in lieu of centering attention on purchase patterns over a calendar year, future work should also consider transactions throughout the year in order to ascertain seasonal patterns, as well as dynamic aspects of milk purchasing behavior.

CHAPTER IV

THE IMPORTANCE OF SOCIO-DEMOGRAPHIC VARIABLES ON THE QUALITY OF PREDICTED PROBABILITIES FROM BINARY CHOICE MODELS: AN APPLICATION OF THE BRIER PROBABILITY SCORE METHOD CONCERNING THE CHOICE OF ORGANIC MILK

Introduction

The use of binary choice models has been standard in explaining behavioral choice between two alternatives or events. Because of the pervasiveness of these models in terms of looking at the underlying drivers associated with dichotomous choice, the task of evaluating these models in terms of their ability to predict correct predictions becomes paramount. One popular measure of fit is the use of the prediction-success/expectation-prediction contingency tables. This approach classifies correct predictions from the following rule: if the predicted probability is greater than 0.5 and the first choice is selected, then the decision of choosing the first choice is correctly predicted. Likewise, if the probability is less than 0.5 and the second alternative is chosen, then the model has made a correct classification of the alternative choice. Accordingly, summing the correctly classified cases over the total number of observations gives the percentage of correct predictions. The higher the percentage of right predictions, the better predictive power the model possesses. Another alternative rule is to forego the 0.5 cut-off and use the mean frequency of observations of the choice variable as the cut-off (Capps and Kramer, 1985). There is flexibility in this approach because if the mean frequency is lower than 0.5, then the model will not be able to predict correct classifications.

The advantage of the approach is its simplicity and ease in calculations and if a symmetric loss function is assumed then 0.5 cutoff rule is justified (Cameron and Trivedi, 2008). However Stock and Watson (2007) argued that the equal odds cutoff does not take into account the quality of the predicted probabilities as the approach does not discriminate whether the predicted probabilities are 51 percent or 99 percent. Likewise Wooldridge (2002) opined that the percent of correctly predicted can be misleading because there is relative ease in predicting one of the outcome and while the opposite is true in predicting the other alternative. Thus, Wooldridge suggested that the more appropriate values to look at are the sensitivity and specificity where the former is the ability to predict outcome Y=1 while the latter is ability to correctly classify outcome Y=0. Several studies including Alviola and Capps (2009) argued that the appropriate cutoff should be based on the frequency of the observations corresponding to the binary choice. This cutoff reflects the actual probability because the equal odds rule does not take into account the number of observations that chose a certain event. Also Cameron and Trivedi (2005, 2008) suggested the comparison of the average value of the binary outcome variable (Y=1) and the mean of the predicted probabilities.

The Stock and Watson (2007) and Wooldridge (2002) critiques and the Cameron and Trivedi (2005, 2008) approach represent the standard textbook orthodoxy in measuring goodness of fit of binary choice models with the use of prediction-success contingency tables. Although most of these studies opine that the approach is suboptimal, they do not offer any superior alternative. We attempted to address this gap by assessing the predictive capacity of binary choice models through the use of probability scores.

We examined the prediction probabilities of fundamental discrete choice models, namely the logit and probit models as well as the linear probability model (LPM), through the Brier Probability Scoring Method. The Brier score is a type of incentive compatible probability forecast method that is used to assess subjective probability forecasts. We also applied the Yates Brier Sore Partition in order to determine the effect of differing model specifications on the ability to sort events that occurred and those that did not occur. Finally, in our analysis, we utilized the 2004 Nielsen Homescan panel in constructing three choice models associated with the purchase/nonpurchase of organic milk.

Methodology

Random Utility Model

The choice of whether to purchase organic milk can be modeled as a binary choice wherein the outcome variable Y_i takes on two values where 1 can be thought of an occurrence of an event or 0 otherwise. In this alternative specification, an agent can assume a utility function where utility comparisons can be made. Given the utility function;

$$U(x_i, \varepsilon_i) \tag{1}$$

where *U* is function of the covariates vector *x*, the agent can assign 1 to a choice where he/she derives higher level of utility and 0 if the alternative choice produced a lower utility level. Assuming that the utility function can be approximated as linear, this choice problem can represented as

(2) $$U_1 = x^T \beta_1 + e_1$$

(3) $$U_0 = x^T \beta_0 + e_0$$

where U_1 and U_0 are the corresponding deterministic utility choices and errors terms e_1 and e_0 are random error components. So for this exercise the household chooses to purchase organic milk (Y_i=1) because higher utility is derived relative to conventional milk. If the household chooses organic milk i.e. $U_1 > U_0$ and if we let *p* be the probability of occurrence, then the probability of occurrence Pr (Y_i=1) becomes:

(4) $$\Pr(Y_i = 1) = \Pr(U_1 > U_0)$$

(5) $$\Pr(Y_i = 1) = \Pr(x^T \beta_1 + e_1 > x^T \beta_0 + e_0)$$

(6) $$\Pr(Y_i = 1) = \Pr(e_0 - e_1 < x^T \beta_1 - x^T \beta_0)$$

(7) $$\Pr(Y_i = 1) = \Pr(\mu < x^T \beta_1 - x^T \beta_0)$$

(8) $$\Pr(Y_i = 1) = F(x^T \beta)$$

where F(.) can be designated as the cumulative density function (cdf). If we assume that e_1 and e_0 are normally distributed, then the difference $\mu = e_1 - e_0$, also is normally distributed. If F(.) is assumed to be the standard normal cdf, then the probit model emerges. If, on the other hand, the error terms e_1 and e_0 follow an extreme value distribution, then the difference follows a logistic distribution. Also, since the Linear

Probability Model (LPM) does not rely on any distribution function, the probability of occurrence is equal to $\Pr(Y_i = 1) = x^T \beta$.[14]

Binary Choice Models and Brier Probability Score

Following the determination of event probabilities from the probit, logit and LPM models, the derivation of the predicted probabilities can be calculated by replacing the β's in equation (8) with their corresponding estimated coefficients ($\hat{\beta}$'s). Thus for this exercise, the respective predicted probabilities can be denoted as $p_{ij}^{m} = F(x^T \hat{\beta})$ where p_{ij}^{m}, represents the predicted probabilities of individual *i* on choice *j* (j = 0, 1) in model m. In this case, *m* = probit (*P*), logit (*L*) or LPM. The respective predicted probabilities of the three models are as follows:

(9) $$p_{ij}^{P} = \Phi(x^T \hat{\beta}_P)$$

(10) $$p_{ij}^{L} = \varphi(x^T \hat{\beta}_L)$$

(11) $$p_{ij}^{LPM} = x^T \hat{\beta}_{LPM}$$

where Φ and φ are standard normal and logistic cdfs for the probit and logit specifications.

With extensive use of binary choice models in modeling dichotomous product choices, assessing both forecast accuracy and sorting capability become paramount.

---

[14] Of course, the problem with the LPM is the possibility that probabilities may fall outside the unit interval (0 to 1). That is, probabilities may either be less than zero, between 0 and 1, or greater than 1. The use of the probit model or logit model eliminates any possibility that probabilities are outside the unit interval.

Following the approach of Bessler and Ruffley (2004) and Olvera and Bessler (2006), let the probability of occurrence of individual *i* on the j^{th} event be p_{ij} and denote d_{ij} as a binary index number that takes on the values of one if the j^{th} event occurred and zero otherwise. Thus, the individual level quadratic probability score (PS) can be written as:

(12) $$PS(p,d) = (p_{ij} - d_{ij})^2$$

where, the values of PS can range from zero to one. This equation can be generalized with a mean probability score (Brier score) indexed over N observations (households in our example) at *i = 1,…,N.* Therefore, the Brier score can be written as:

(13) $$\bar{PS}(p,d) = \left(\frac{1}{N}\right)\sum_{i=1}^{N}(p_{ij} - d_{ij})^2$$

Given equation (13), a Brier Score of 0 means perfect forecast accuracy while a score of 1 denotes complete forecast inaccuracy. In this exercise, estimation of the mean probability score was calculated in order to assess the quality of probability forecasts from binary choice models and to determine the importance of socio-demographic variables in terms of the ability to discriminate events that occurred and those that did not occur.

Yates Decomposition of the Brier Score

Furthermore, the Yates covariance partition (1982, 1988) of the Brier score was utilized to address the issue of relationship between reported and actual forecasts. The Yates partition discussed in Bessler and Ruffley (2004) and Olvera and Bessler (2006), separates the Brier score into decomposable factors such as bias, scatter, minimum variance probability score, variance of outcome index (*d*) and covariance between *p* and *d*. In notation form, this decomposition can be written as:

(14) $$\overline{PS}(p,d) = Var(d) + MinVar(p) + Scatter(p) + Bias^2 - 2 * Cov(p,d)$$

Starting with the term *Var(d)*, defined as outcome index variance, the notational representation can be written as:

(15) $$Var(d) = \bar{d}_{ij}(1 - \bar{d}_{ij})$$

with $\bar{d}_{ij} = \frac{1}{N}\sum_{i=1}^{N}\bar{d}_{ij}$ as the mean of the outcome index *d*. This term reflects the factors that are exogenous to the forecaster (Yates 1982, 1988).

Scatter (*p)* is defined as:

(16) $$Scatter(p) = \frac{1}{n}\left[n_1 Var(p_{1j}) + n_0 Var(p_{0j})\right]$$

where $Var(p_1) = \frac{1}{n_1}\sum_{i=1}^{n_1}(p_{1j} - \bar{p}_1)^2$ and $Var(p_0) = \frac{1}{n_0}\sum_{i=1}^{n_1}(p_{0j} - \bar{p}_0)^2$ denote conditional variances of the predicted probabilities for events that occurred (p_1) and for those events that did not occur (p_0). Thus, scatter is the weighted average value of the two conditional variances and is defined as an indicator of the total noise contained in the predicted probabilities of the two events. Note that $n_0 + n_1 = N$.

MinVar(p) represents the total variance and is defined as:

(17) $$MinVar(p) = Var(p) - Scatter(p)$$

where $Var(p) = \frac{1}{N}\sum_{i=1}^{N}(p_{ij} - \bar{p}_{ij})^2$ with $\bar{p}_{ij}$ as the mean probability of occurrence $\frac{1}{N}\sum_{i=1}^{N} p_{ij}$. Likewise, the component *Bias* is denoted as:

(18) $$Bias = \bar{p}_{ij} - \bar{d}_{ij}$$

This term measures the difference of the mean predicted probability and the mean outcome index. Thus, *Bias* measures, on average, the deviation associated with the forecasted probabilities to their true outcomes. The deviation also is the rate of miscalibration because the bias term measures how probability forecasts are overpredicted or underpredicted (Yates 1982, 1988).

The term *Cov(p,d)* reflects ability to filter relevant information that enables a proper assignment of probabilities for events that occurred and for those that did not occur. This term is given as:

(19) $$Cov(p,d) = \bar{p}_1 - \bar{p}_0 (Var(d))$$

where $\bar{p}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} p_{i1}$ and $\bar{p}_0 = \frac{1}{n_0}\sum_{i=1}^{n_1} p_{i0}$ are mean probability of occurrence for events that occurred and those that did not occur.

Empirical Specification

In this exercise, two model specifications were estimated for each binary choice model. The respective model specifications were modeled as:

(20)

$$P(q_i = 1|W_i) = \beta_0 + \beta_1 Income_i + \beta_2 Hs2_i + \beta_3 Hs3_i + \beta_4 Hs4_i + \beta_5 Hs5_i + \beta_6 Agepchild_i$$
$$+ \beta_7 Empparttime_i + \beta_8 Empfulltime_i + \beta_9 Eduhighschool_i + \beta_{10} Edusomecollege_i +$$

$$\beta_{11}EduCollegePlus + \beta_{12}White_i + \beta_{13}Black_i + \beta_{14}Oriental_i + \beta_{15}Hisyes_i + \beta_{16}Central_i + \beta_{17}South_i + \beta_{18}West + \varepsilon_i$$

(21) $\Pr(q_i = 1 | X_i) = F(\beta_0 + \beta_1 Income),$

In each specification as given by equation (20) or equation (21), q_i represents household i’s choice to purchase organic milk and 0 otherwise. Also, F(.) is the cdf, either a standard normal distribution to represent a probit specification or a logistic distribution to represent a logit specification. With the LPM model, the cdf is omitted in its specification. The set of explanatory variables include household socio-demographic variables associated with the household head such as household income (*Inc)*, type of employment, level of education, race, and ethnicity of the household, the presence or absence of children and region (*Reg)*.

Equation (21) omits everything except for the income covariate. We use this specification to determine the impact of censoring potentially important socio-demographic variables on the forecasting ability of binary choice models. Thus, two sets of predicted probabilities for each choice model (probit, logit and LPM) were estimated. These in turn were used to derive two sets of Brier Scores, prediction success tables, and Yates Brier Score partition (decomposition) factors.

Data

For this empirical exercise, the data pertaining to the choice of purchasing organic milk, income and household socio- demographic variables are from the 200 Nielsen Homescan Panel. Table 4.1 presents the definition and summary statistics of all

Table 4.1. Summary Statistics of Variables Used in the Analysis

Variable	Description	Observation	Mean	Std. Dev.	Min	Max
Yesorg ($q_i = 1$)	Household purchased organic milk	38,192	0.119	0.324	0	1
Noorg ($q_i = 0$)	Household did not purchase organic milk	38,192	0.881	0.324	0	1
Income	HH income	38,192	50,024	27,306	5,000	100,000
Hs1	HH size of 1[a]	38,182	0.262	0.440	0	1
Hs2	HH size of 2	38,192	0.391	0.488	0	1
Hs3	HH size of 3	38,192	0.143	0.350	0	1
Hs4	HH size of 4	38,192	0.127	0.333	0	1
Hs5	HH size > 4	38,192	0.077	0.267	0	1
Agepcchild	HH has at least 1 child less than 18 yrs of age	38,192	0.253	0.435	0	1
No children	HH has no children less than 18 years of age	38,192	0.747	0.435	0	1
Unemployed	Head of HH is unemployed	38,192	0.408	0.491	0	1
Empparttime	Head of HH is employed part-time	38,192	0.157	0.364	0	1
Empfulltime	Head of HH is employed full-time	38,192	0.435	0.496	0	1
Edulths	HH head completed less than 12 years of schooling[a]	38,192	0.038	0.192	0	1
Eduhighschool	HH head is high school graduate	38,192	0.275	0.446	0	1
Edusomecollege	HH head has completed some college	38,192	0.320	0.446	0	1
Educollegeplus	HH head has at least a college education	38,192	0.367	0.482	0	1
White	HH head is white	38,192	0.825	0.380	0	1
Black	HH head is black	38,192	0.096	0.295	0	1
Oriental	HH head is Oriental	38,192	0.022	0.146	0	1
Other	HH head is classified as other[a]	38,192	0.057	0.232	0	1
Hispyes	HH head is Hispanic	38,192	0.066	0.248	0	1

Table 4.1 Continued

Variable	Description	Observation	Mean	Std. Dev.	Min	Max
Hispno	HH is not hispanic[a]	38,192	0.934	0.248	0	1
East	HH is located in the East[a]	38,192	0.163	0.370	0	1
Central	HH is located in the Midwest	38,192	0.235	0.424	0	1
South	HH is located in the South	38,192	0.384	0.486	0	1
West	HH is located in the West	38,192	0.219	0.413	0	1

Source: Nielsen Home Scan Panel for Calendar Year 2004

HH denotes household; the HH head is defined as the female head. If a female head of household does not exist, then the HH head is the male head.

[a] Reference category so as to avoid the dummy variable trap.

the relevant variables that were used in the study. The Nielsen scanner data set is the world's largest, on-going household scanner data survey system wherein it tracks household purchases in the United States.

The variable *Yesorg* is the dependent choice variable and is indexed as 1 to represent purchase of organic milk and 0 otherwise. *Income* is defined as household income and the average income level of the sample was $50,025/household.As for the household size, the study used indicator variables to describe the number of household members where *Hs1 (26%)* and *Hs2 (40%)* pertain to households having one and two members while *hs3* has 3 household members with a mean proportion of 14 percent. The two last household size indicator variables *hs4 and hs5* describes 4 and 5 or more members in the household. The respective mean proportion are 13 and 8 percent respectively Also, households with children less than 18 years old (*agepcchild*) were 25 percent of the sample.

The demographic characteristics of the household head were also included in this study. Both the employment status and educational attainment of the household head were represented as dummy or indicator variables. The variables *Unemp, Empparttime* and *Empfulltime* are indicator variables representing whether the household head was unemployed, employed part-time or employed fulltime. Their respective mean proportions are 41 percent, 16 percent and 43 percent. Similarly the variables *Edulths, Eduhighschool, Edusomecollege* and *Educolleges* denote household head educational attainment whether it is below high school, high school, above high school but below college and college and beyond. The respective mean proportions are 4 percent, 28 percent, 32 percent and 37 percent.

Also included into the model were race and ethnicity of the household. The indicator variables *White, Black, Oriental* and *Others* represented the major racial household distinction. Approximately 83 percent are white households. On the other hand household ethnicity was represented as either Hispanic (*Hispyes-7 percent*) or non-hispanic (*Hispno-93 percent*). Finally, regional dummy variables such as *East*, *Central, South* and *West* were included to describe the regional location of the household. The respective mean proportions are 16 percent, 24 percent, 38 percent and 22 percent respectively.

Results

Inter-Binary Choice Model Comparisons

For this exercise, three models were used, namely the probit, logit and linear probability models to represent the binary choice between organic and conventional milk. Tables 4.2 and 4.3 report the logit, probit and LPM estimated parameters of both the full model and income only model .The Brier Score and Yates partition components are exhibited in Table 4.4. The calculated Brier Scores (BS) for the three respective models are given as follows: Probit (BS=0.1028960), Logit (BS=0.1029092) and LPM (BS=0.1028963). Furthermore, the Probit model has the highest forecast covariance value compared to the other two models. These results imply that the probit model predicts better than the logit and LPM models by having both the lowest Brier scores and highest forecast covariance values (Table 4.4).

Table 4.2. Full Model Parameter Estimates of Logit, Probit and LPM Analysis of Organic Milk Choice

Variable	Logit Model		Probit Model		Linear Prob. Model	
	Estimates	(P>|z|)	Estimates	(P>|z|)	Estimates	(P>|z|)
Hs2	-0.1420	0.0010	-0.0768	0.0010	-0.0148	0.0010
Hs3	-0.1818	0.0040	-0.0968	0.0040	-0.0191	0.0040
Hs4	-0.2921	0.0000	-0.1589	0.0000	-0.0304	0.0000
Hs5	-0.3105	0.0010	-0.1673	0.0000	-0.0329	0.0000
Income	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Agepcchild	-0.0790	0.1880	-0.0429	0.1740	-0.0082	0.1740
Empparttime	0.1272	0.0080	0.0659	0.0090	0.0138	0.0070
Empfulltime	-0.1532	0.0000	-0.0837	0.0000	-0.0160	0.0000
Eduhighschool	0.0529	0.6150	0.0245	0.6380	0.0045	0.5490
Edusomecollege	0.3808	0.0000	0.1908	0.0000	0.0309	0.0000
Educollegeplus	0.6830	0.0000	0.3555	0.0000	0.0663	0.0000
White	-0.2429	0.0040	-0.1292	0.0040	-0.0273	0.0090
Black	0.2212	0.0180	0.1215	0.0170	0.0258	0.0320
Oriental	0.2789	0.0170	0.1619	0.0130	0.0461	0.0080
Hispyes	0.2997	0.0000	0.1673	0.0000	0.0355	0.0000
Centrak	-0.3779	0.0000	-0.1933	0.0000	-0.0339	0.0000
South	-0.0431	0.3560	-0.0222	0.3710	-0.0044	0.3740
West	0.1470	0.0030	0.0807	0.0030	0.0175	0.0020
Constant	-2.3285	0.0000	-1.3431	0.0000	0.0958	0.0000
Pseudo R^2	0.0287		0.029			
Obs	38192		38192		38192	
Wald chi2(18)	804.39		800			
Prob>chi2	0.000		0.000			
R2					0.0212	
F(18, 38173)					43.5	
Prob > F					0.000	

Table 4.3. Income-Only Model Parameter Estimates of Logit, Probit and LPM Analysis of Organic Milk Choice

Variable	Logit Model		Probit Model		Linear Prob. Model	
	Estimates	(P>|z|)	Estimates	(P>|z|)	Estimates	(P>|z|)
Income	7.34E-06	0.0000	3.88E-06	0.0000	7.89E-07	0.0000
Constant	-2.38081	0.0000	-1.3788	0.0000	0.079893	0.0000
Pseudo R^2	0.0059		0.0059			
Obs	38192		38192		38192	
Wald chi2(1)	165.54		164.12			
Prob>chi2	0.0000		0.0000			
R2					0.0044	
F(1, 38190)					156.94	
Prob > F					0.0000	

Table 4.4. Brier Score and Decompositions of Probit, Logit and Linear Probability Model (LPM) and Model Variants for Organic Milk Choice

PROBIT MODEL	Probit (Full Model)	Probit (Income Only)[a]	% Change
Brier Score (BS)	0.1028960	0.1046501	1.705
Variance of *d* (Var(d))	0.1051212	0.1051212	0.000
Minimum variance of *p* *(Min Var(p))*	**0.0000487**	**0.0000020**	**-95.873**
Scatter (Scatter(p))	**0.0022488**	**0.0004615**	**-79.478**
Bias^2	1.1E-10	8.1E-13	-99.264
Forecast covariance (2Cov(p,d))	0.0045228	0.0009346	**-79.336**
Slope	0.0215121	0.0044453	**-79.336**
Intercept	0.1167921	0.1188407	1.754

LOGIT MODEL	Logit (Full Model)	Logit (Income Only)	% Change
Brier Score (BS)	0.1029092	0.1046490	1.691
Variance of *d* (Var(d))	0.1051212	0.1051212	0.000
Minimum variance of *p* *(Min Var(p))*	**0.0000484**	**0.0000015**	**-96.921**
Scatter (Scatter(p))	**0.0022520**	**0.0004645**	**-79.374**
Bias^2	0.0000000	0.0000000	0.000
Forecast covariance (2Cov(p,d))	0.0045124	0.0009388	**-79.195**
Slope	0.0214629	0.0044655	**-79.194**
Intercept	0.1168085	0.1188375	1.737

LINEAR PROBABILITY MODEL	LPM (Full Model)	LPM (Income Only)	% Change
Brier Score (BS)	0.1028963	0.1046569	1.711
Variance of *d* (Var(d))	0.1051212	0.1051212	0.000
Minimum variance of *p* *(Min Var(p))*	**0.0000471**	**0.0000021**	**-95.520**
Scatter (Scatter(p))	**0.0021779**	**0.0004623**	**-78.773**
Bias^2	0.0000000	0.0000000	0.000
Forecast covariance (2Cov(p,d))	0.0044500	0.0009288	**-79.128**
Slope	0.0211657	0.0044175	**-79.129**
Intercept	0.1168440	0.1188432	1.711

[a] Model variant has income as the only explanatory variable for all the three choice models.

Prediction success tables also were utilized to assess the ability of the “complete” model to classify outcomes (Table 4.5). Instead of the default 0.5 cut-off value, the appropriate critical values were calculated based on the purchase frequency of organic milk relative to the whole sample size. The choice of cut-off value was made to reflect the actual probability of choosing organic milk and not the usual application of the equal odds approach in both choices. For all three choice models utilized, the cutoff value was equal to 0.119. Results indicate that the logit model garnered the highest percentage of right predictions (58.41 percent) relative to the probit (57.97 percent) and the LPM (54.64 percent). The implication is that the logit model results in 58 percent correct predictions, the probit just fewer than 58 percent correct predictions, and the LPM slightly more than 54 percent correct predictions. Thus, among the three models, the logit model performs best in correctly classifying those households that chose organic and/or conventional milk. Although both methods resulted in different outcome in terms of model superiority, the observed values are very close that inference suggests that there is no significant difference. The observed values are in agreement with Capps and Kramer (1985) where they analyze food stamp participation using probit and logit model specification. Their conclusions include that both models empirical performance were indeed minimal.

Table 4.5. Prediction-Success Evaluation for Probit, Logit and Linear Probability Models (LPM) in Both Full Model and Income-only Specifications

PROBIT	Actual Choice			
	Complete		Income Only	
Predictions	Organic Milk	Conventional	Organic Milk	Conventional
Organic Milk	2772	14266	2340	14336
Conventional	1787	19367	2219	19297
Total	4559	33633	4559	33633
	Full Model	Income Only		
% Right Predictions[a]	**57.97**	**56.65**		
Sensitivity (%)[b]	**60.80**	**51.33**		
Specificity (%)[c]	**57.58**	**57.38**		
Cut-off value	0.12	0.12		

LOGIT[d]	Actual Choice			
	Complete		Income Only	
Predictions	Organic Milk	Conventional	Organic Milk	Conventional
Organic Milk	2747	14073	2340	14336
Conventional	1812	19560	2219	19297
Total	4559	33633	4559	33633
	Full Model	Income Only		
% Right Predictions	**58.41**	**56.65**		
Sensitivity (%)	**60.25**	**51.33**		
Specificity (%)	**58.16**	**57.38**		
Cut-off value	0.12	0.12		

Table 4.5 Continued

LPM[e]	Actual Choice			
	Complete		Income Only	
Predictions	Organic Milk	Conventional	Organic Milk	Conventional
Organic Milk	2962	15727	2340	14336
Conventional	1597	17906	2219	19297
Total	4559	33633	4559	33633
	Full Model	Income Only		
% Right Predictions	**54.64**	**56.65**		
Sensitivity (%)	**64.97**	**51.33**		
Specificity (%)	**53.24**	**57.38**		
Cut-off value	0.12	0.12		

[a] For full model ((2772+19367)/38192)*100 and for income only ((2340+19297)/38192)*100

[b] This is the percentage of correctly predicting the choice of choosing organic milk. For full model (2772/4559)*100 and for income only (2340/4559)*100

[c] This is the percentage of correctly predicting the choice of choosing conventional milk. For full model (19367/33633)*100 and for income only (19297/33633)*100

[d, e] Same calculations as with the probit example

Inter-Model Probabilistic Graphs

Following Yates (1982, 1988) and Olvera and Bessler (2006), illustrative constructs called probabilistic or covariance graphs were utilized to demonstrate the ability to differentiate binary choice events that had occurred or did not occur. The graphs illustrate the ability to discriminate between the choice of purchasing organic and conventional milk across three binary choice models, namely probit, logit and linear probability models (LPM). Results indicate that the slope and intercept of the three probabilistic graphs (Figures 4.1a, 4.2a and 4.3a) have values that are close to one another

Intra-Binary Choice Model Comparisons

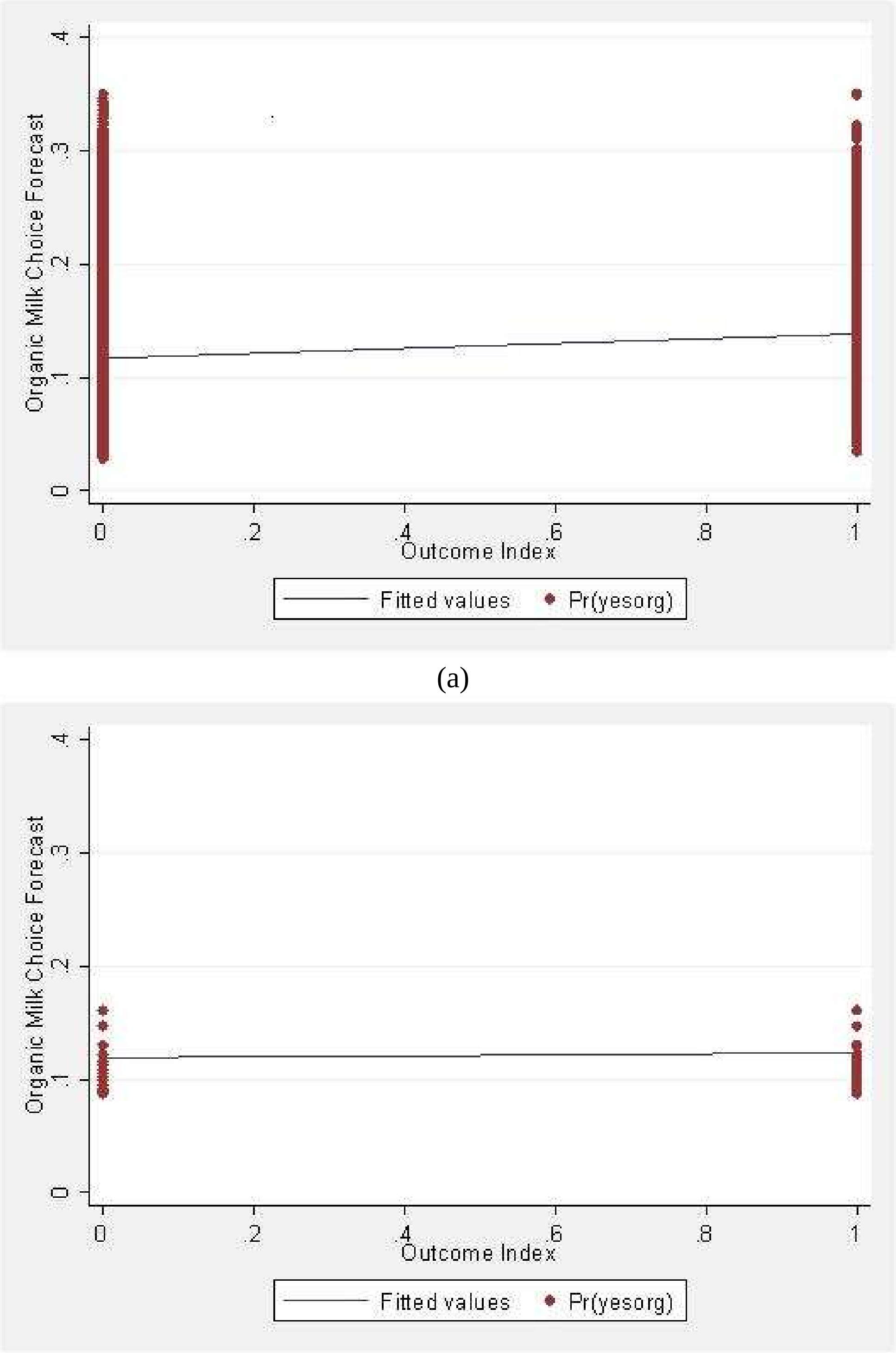
In this section of the paper, the analysis shifts from comparing different binary choice models to looking at one choice model and its respective model variant. More specifically, we compare a choice model containing covariates such as income and various socio-demographic variables with a model variant which contains income as its only explanatory variable.

Results from Table 4.4 indicate that for all three models, Brier scores had increased between complete models and their variants with income as the only explanatory variable. More specifically, the increase in terms of percent change for the probit versus probit variant (income only) model was approximately 1.71 percent. For the logit model and its respective logit variant, the percent change increased by 1.69 percent. As for the LPM and model variant, the approximate increase in percentage change was 1.711 percent. The increase in the Brier scores implies diminishing forecasting ability of all three models with respect to predicting both choices (Table 4.4). This difference in Brier score was brought about by the declining variability of the predicted probabilities due to the omission of critical socio-demographic variables in a binary choice model specification (MinVar(p)). Thus, the results imply that when important socio-demographic determinants are removed, the variability of predicted probabilities is reduced and therefore forecasting ability is diminished.

Results from the prediction success-tables exhibited in Table 4.5 indicate that for both probit and logit models, the percent of right predictions declined by approximately 2.27 percent and 3 percent. As for the LPM model, percentage of right predictions increased by 3.69 percent. Also for both the probit and logit models, we find that in terms of sensitivity or the ability to classify correctly the choice of organic milk, the sensitivity declined by 15.58 percent and 14.82 percent. Likewise, the specificity, or the ability to correctly predict the choice of conventional milk, declined by 0.36 percent and 1.34 percent among model variants. The sensitivity of the LPM decreased by 21 percent while its specificity increased by 7.77 percent. Again based on the results of the prediction-success or contingency tables, censure of critical important socio-demographic variables reduces in most cases the ability of choice models to make right predictions.

Intra-Model Probabilistic Graphs

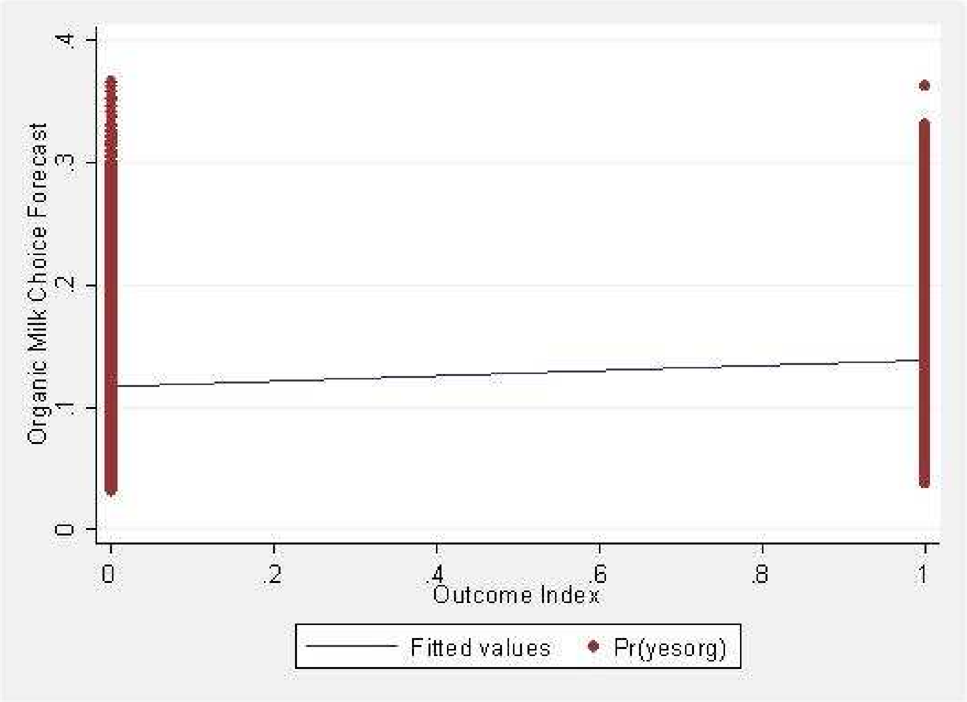
Figures 4.1a, 4.1b, 4.2a, 4.2b, 4.3a 4.3b illustrate pairwise covariance graphs for probit, logit, LPM specifications and their respective model variants. Results show that



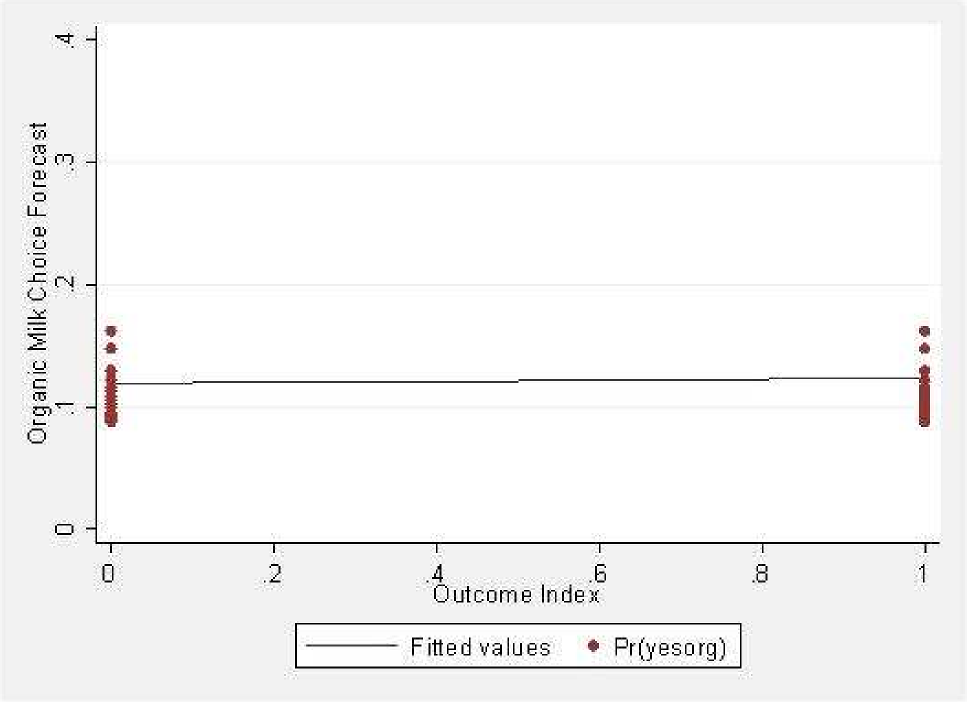
(a)

(b)

Figure 4.1. Probit (a) and probit-income variant (b) model probabilistic graphs

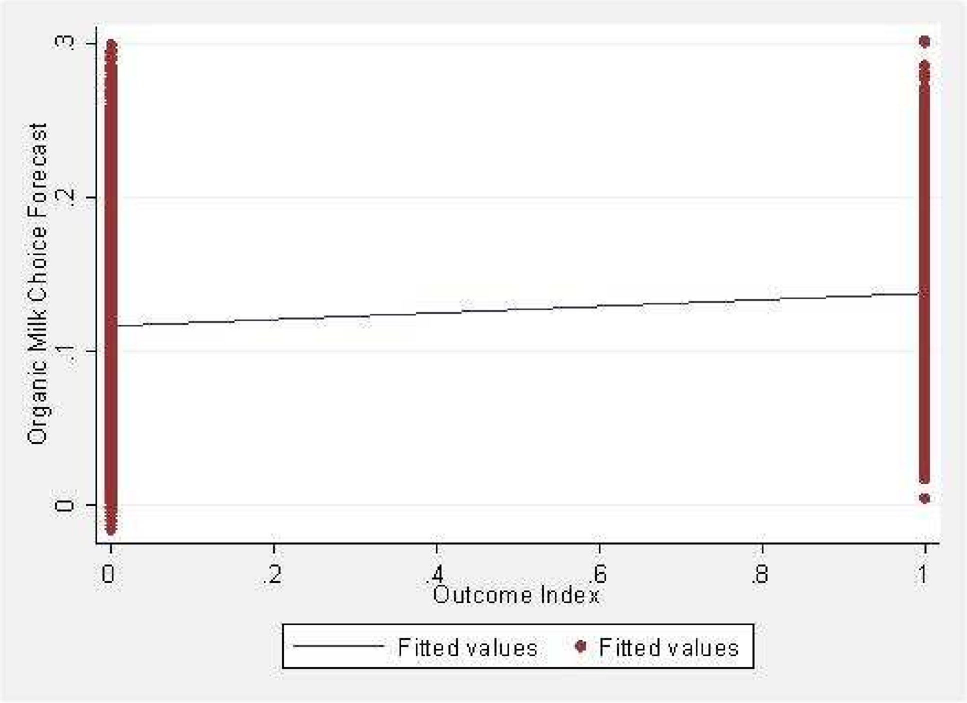


(a)

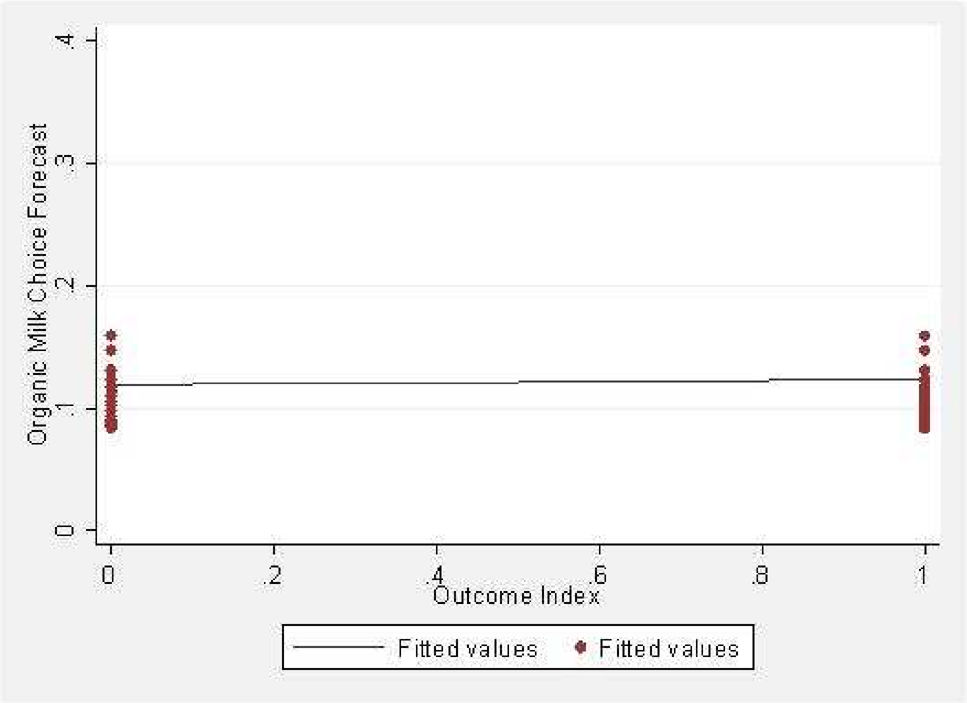


(b)

Figure 4.2. Logit (a) and logit-income variant (b) model probabilistic graphs



(a)



(b)

Figure 4.3. Linear Probability Model (a) and LPM-income variant (b) model probabilistic graphs

the slopes of the probit, logit and LPM covariance graphs declined significantly when socio-demographic variables were removed from the original binary choice specification. For example, percentage changes in the slope for the probit and its income-only variant declined by approximately 79 percent. For the logit and LPM models, the percentage change in slope also decreased by 79 percent. These numbers are confirmed by the flatter probabilistic graphs that characterize choice models that are income-only variants.

Intra-Model Analysis of the Yates Partition

The Yates partition decomposes the Brier score into factors such as bias, scatter, minimum forecast variance, variance of outcome index (d) and covariance between p and d. In this section we center attention to the effect on scatter and minimum variance components. Results from Table 4.4 show that across the three models, the values of both factors declined noticeably when the number of explanatory variables were reduced to only the income variable. For example, the declining percent change for the probit model and its income only variant in both minimum forecast variance and scatter were 95.87 percent and 79.48 percent. Likewise, for the logit model and its income-only model variant, the decline in percentage change were approximately 96.92 (minimum forecast variance) and 79.37 percent (scatter). As for the LPM model, similar changes also were observed in both direction of change and magnitude relative to the probit and logit models.

The effect of omitting important socio-demographic variables resulted then in reducing the variability of predicted probabilities. This reduction however also can mean limited information flow which can constrain the ability of choice models to discriminate between events that occurred and those that did not occur. With limited information flow, we find that there is increased filtering of irrelevant information, and therefore the value of the scatter component decreases. As with the minimum variance, the limited information reduced the overall variance of the respective probabilities. Finally, with reduced information flow, the gap between probabilities assigned to binary events diminishes, thus we find that the forecast covariance decreases. In summary, model specifications that limit information flow in binary choice models can bring about increased noise filtering (declining scatter), lessening of overall forecast variance (decreased minimum forecast variance) and weakening of the ability to filter relevant information that enables the proper assignment of probabilities for events that occur and did not occur (reduced forecast covariance).

Conclusions

There were two levels of analysis done in this study; considering comparisons across choice models and considering comparisons of alternative specifications within choice models. Utilizing probit, logit and linear probability choice models to represent the choice of organic milk or conventional milk, both Brier scores and prediction-success tables were evaluated to determine their usefulness in making accurate predictions. Results indicate that the probit model predicts better among the three models by having the lowest Brier Score and highest forecast covariance values. However, when the prediction-success was used, the logit model performed best in terms of correct classifications. One notable observation was that across the three models, the values of the Brier score, Yates partition factors and prediction-success tables were very close in magnitude. The study also utilized probabilistic graphs in order to illustrate the ability of all models to differentiate between events that occurred (choosing organic milk) and those that did not occur (choosing conventional milk).

When important socio-demographic variables are omitted in a binary choice model, the variability level of the predicted probabilities becomes significantly reduced. Consequently, this diminishes the ability of the model to sort binary events or choices. Estimates from the Brier scores indicate that for each of the choice models vis-à-vis their respective income-only variant, the values increased indicating diminished forecasting ability. Likewise, results from the prediction-success table point to declining percentages of correct classifications. The declining slope change of the covariance graphs between “complete” models and their income-only variants is indicative of diminished binary event discriminatory ability.

With regards to the effect on the factors from the Yates partition, the study focused on the scatter and minimum variance. Results show that when important socio-demographic variables are omitted, scatter and minimum variance values are significantly reduced. An intuitive explanation for this change lies in the reduction of the variability of predicted probabilities. Also, the removal of important socio-demographic variables resulted in a weakened ability to sort between events that occurred and did not occur. And as such, points to the tradeoff between sorting and variability. As to the use of prediction success tables, one must also utilize other methods such as probability scoring as this paper showed that logit was the superior model in using the prediction success tables, whereas the probit performed best under the probability score criteria.

CHAPTER V

MICRO-DEMAND SYSTEMS ANALYSIS OF NON-ALCOHOLIC BEVERAGES IN THE US: AN APPLICATION OF ECONOMETRIC TECHNIQUES DEALING WITH CENSORING

The move towards different diet mechanisms that favor nutritious foods has in recent years led to the emergence of healthier and natural food choices. In particular, manufacturers and retailers have been responsive in introducing new products to the non-alcoholic beverage industry, especially juices, energy drinks and others. This chapter focuses on the interdependencies of milk, and demand for certain non-alcoholic beverages, namely: fruit juices, tea, coffee, carbonated soft drinks, and bottled water. In the case of the non-alcoholic beverage complex, these products have different levels of market penetration. Consequently, the dependent variables associated with these non-alcoholic beverages are censored at zero. That is, certain households have zero expenditures, but the corresponding information on household characteristics, which forms the basis of the explanatory variables are often readily observed. Thus, several competing estimation methods have been developed in order to address the censoring issue in the estimation of micro-demand systems. Importantly, no prior research has been done in terms of utilizing these respective approaches with regard to a particular data set.

In this study, the estimation of the demand system made use of Quadratic Almost Ideal Demand System (QUAIDS) model (Banks, Blundell and Lewbel, 1997) and Almost Ideal Demand System (AIDS) (Deaton and Mulbauer, 1980). The advantages of the QUAIDS model are its flexibility in incorporating nonlinear effects and interactions of price and expenditures in the demand relationships. Since the data used are at the household level, censoring is typically observed as some households report expenditures of a beverage product say coffee and none on say bottled water. Thus, in order to model the censoring problem in demand systems, the research utilized estimation procedures that range from the use of two-step estimators (Heien and Wessells, 1990; Shonkwiler and Yen, 1999), maximum entropy and maximum simulated likelihood estimation (Dong, Gould and Kaiser, 2004a). The use of the iterated seemingly unrelated regression (ITSUR) estimation without adjustments for censoring serves as a basis of comparison for the aforementioned estimation techniques. Finally, the source of data is the 1999 Nielsen Homescan Panel due to its vast array of household demographic information.

Literature Review

The use of the Quadratic Almost Ideal Demand System (QUAIDS) model in applied work has been well documented. For example, Dhar and Foltz (2005) utilized a quadratic AIDS model to estimate values and benefits derived from rBST, organic milk and unlabelled milk. Their study used scanner time-series data of milk consumption of 12 key cities in the United States. Their findings indicate that rBST and organic milk are complements, while conventional milk and rBST milk, as well as conventional milk and organic milk are substitutes. Their own-price elasticity estimates were -4.40 (rBST free milk), -1.37 (organic milk) and -1.04 (conventional milk).

Likewise, a study done by Mutuc, Pan and Rejesus (2007) investigated household demand for vegetables in the Philippines through the use of QUAIDS. Their findings indicated significant differences in expenditure elasticities in both rural and urban areas whereas for the respective own and cross-price elasticities, no significant variations across rural and urban areas were evident. Dhar and Foltz (2005) used Full Information Maximum Likelihood (FIML) as an estimation procedure, the Mutuc et al.'s (2007) study had a censoring problem because of the presence zero expenditures on some vegetable commodities being consumed by the household, hence their usage of the Shonkwiler and Yen (1999) two step procedure. As microdata become increasingly available and more detailed, the estimation of micro-demand systems at the household level becomes problematic due to censoring.

The work of Heien and Wessells (1990) was one of the pioneering studies to address the censoring problem in demand systems estimation. Their approach mimics the Heckman two-stage method by estimating probit models to compute the inverse Mill's ratios for each commodity. Subsequently, these measures are incorporated into the second step SUR estimation of the budget shares. On the other hand, Shonkwiler and Yen (1999) proposed a consistent estimation procedure that utilizes a probit estimator in the first step and then using the cdf to multiply the covariates in the demand shares and including the pdf as an independent variable in the second step. Both methods fall under the purview of utilizing two-step estimators.

While the Shonkwiler and Yen approach worked well with the problem of zero expenditure, Arndt, Liu and Preckel (1999) claimed that it had limitations with respect to dealing with corner solutions. Several studies including Arndt (1999) and Golan, Perloff and Shen (2001) propose an alternative approach called maximum entropy to estimate censored demand systems. This approach allows for consistent and efficient estimation of demand systems without putting any restrictions on the error terms. Other researchers such as Meyerhoefer, Ranney and Sahn (2005) use the general method of moments (GMM) estimator to address censoring problems in demand systems estimation. The GMM method was not used in this study.

Several studies have criticized the two step method stating that it has ignores the “adding up” restriction in estimating share equations in the censored demand systems (Dong, Gould and Kaiser, 2004a; Yen, Lin and Smallwood, 2003). Together with Golan, Perloff and Shen (2001), these classes of estimators fall under the Amemiya-Tobin framework where the former does not employ maximum likelihood estimation in evaluating multivariate probability integrals while Dong, Gould, Kaiser (2004a) and Yen, Lin Smallwood (2003) utilize numerical methods such as maximum and quasi-maximum simulated likelihood estimation in approximating the likelihood function. The literature regarding the use of alternative estimation techniques such as Bayesian and non-parametric approaches on micro-demand system estimation have been limited (Tiffin and Aquiar, 1995).

Methodology

Almost Ideal Demand System (AIDS) Model

This research utilizes the AIDS (Deaton and Mulbauer, 1980) model in the demand system estimation of six non-alcoholic beverages, namely: fruit juices, tea, coffee, carbonated soft drinks, bottled water and milk. Equation (1) describes the general specification of the AIDS model where p_j and w_i are the price and budget share of the ith beverage commodity with the price index ln *α(p)* being specified further as a function of own and cross prices. The average budget share w_i is computed as p_iq_i/M *where* $M = \sum p_iq_i$ is the total expenditure. The parameters of this system are α_i, γ_i and β_i, respectively. One can also incorporate household demographic characteristics into the demand system thru the intercept parameter α_i. These variables include *HHsize* for household size, *Inc* as household income and *Race* is race type. Also, the variable *Season* represents the seasonality component and *Rg* is the Region.

(1) $$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{M}{a(p)}\right] + \varepsilon_i \text{, i-1,2,..n}$$

where;

$$\ln a(p) = \alpha_o + \sum_{i=1}^{n} \alpha_i \ln p_i + 0.5 \sum_i \sum_j \gamma_{ij} \ln p_i \ln p_j$$

$$\alpha_i = \alpha_i^* + \alpha_{i1} HHsize + \alpha_{i2} Inc + \alpha_{i3} Race + \alpha_{i4} Season + \alpha_{i5} Rg$$

On the other hand, the classical theoretical restrictions of adding up, homogeneity and symmetry imposed in the AIDS demand system estimation have the following notational representation;

Adding up: $\sum_{i=1}^{n} \alpha_i = 1$, $\sum_{i=1}^{n} \beta_i = 0$, $\sum_{j=1}^{n} \gamma_{ij} = 0$

Homogeneity: $\sum_{j=1}^{n} \gamma_{ij} = 0$

Symmetry: $\gamma_{ij} = \gamma_{ji}$

By imposing these restrictions, the model satisfies the Engel Aggregation thru the adding up condition and from the parameter γ_{ij}, homogeneity and symmetry are imposed.

Quadratic Almost Ideal Demand System (QUAIDS) Model

The Quadratic Almost Ideal Demand System (QUAIDS) model (Banks, Blundell and Lewbel, 1997) also is utilized in this demand analysis. The advantages of using this model over competing flexible demand systems is its unparalleled capability of incorporating non-linear effects and interactions of price and expenditures on the demand specifications. The mathematical representation of the QUAIDS demand system is as follows:

$$(2) \qquad w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{M}{a(p)}\right] + \frac{\lambda_i}{b(p)}\left[\ln\left(\frac{M}{a(p)}\right)\right]^2 + \varepsilon_i \text{, i-1,2,..n}$$

where:

$$\ln a(p) = \alpha_o + \sum_{i=1}^{n} \alpha_i \ln p_i + 0.5\sum_{i} \sum_{j} \gamma_{ij} \ln p_i \ln p_j$$

$$b(p) = \prod_{i=1}^{n} p_i^{\beta_i}$$

$$\alpha_i = \alpha_i^* + \alpha_{i1} HHsize + \alpha_{i2} Inc + \alpha_{i3} Race + \alpha_{i4} Season + \alpha_{i5} Rg;$$

The QUAIDS model is a more generalized version of the AIDS model. Also, if the joint significance test of the parameter $\lambda_i = 0$ is rejected then the QUAIDS model is a superior model at least statistically relative to the AIDS model system. In this research, the intercept parameter α_i incorporates the household demographic characteristics just as with the AIDS model. Since the QUAIDS model has a quadratic term, then another parameter restriction associated with adding up is imposed in addition to the classical theoretical restrictions that were applied.

Adding up: $\sum_{i=1}^{n}\alpha_i = 1$, $\sum_{i=1}^{n}\beta_i = 0$, $\sum_{j=1}^{n}\gamma_{ij} = 0$, $\sum_{i=1}^{n}\lambda_i = 0$

Homogeneity: $\sum_{j=1}^{n}\gamma_{ij} = 0$

Symmetry: $\gamma_{ij} = \gamma_{ji}$

Again the imposition of these restrictions satisfies the Engel Aggregation and the homogeneity and symmetry conditions are subsumed thru the parameter γ_{ij}.

Elasticity Estimation in AIDS and QUAIDS Demand Systems

When the needed parameters are already estimated, the elasticity estimates can now be calculated for the AIDS and QUAIDS demand systems. Following Green and Alston (1990) and Bank, Blundell and Lewbel (1997) formulas, the expenditure, uncompensated and compensated price elasticities are given by the following formulas;

(3) $$\eta_i = \frac{\beta_i}{w_i} + 1$$, for the AIDS model

(4) $$\eta_i = \frac{\beta_i + \frac{2\lambda_i}{b(p)}\left[\ln\left(\frac{m}{a(p)}\right)\right]}{w_i} + 1$$, for the QUAIDS model.

On the other hand the Marshallian or uncompensated price elasticities are given by

(5) $$\varepsilon_{ij}^{u} = \frac{\gamma_{ij} - \beta_i(\alpha_j + \sum_k \gamma_{ik}\ln P_k)}{w_i} - \delta_{ik}$$, for the AIDS model

(6) $$\varepsilon_{ij}^{u} = \frac{\gamma_{ij} - \mu_i(\alpha_j + \sum_k \gamma_{ik} \ln P_k) - \frac{\lambda_i \beta_j}{b(p)}\left[\ln\left(\frac{m}{a(p)}\right)\right]^2}{w_i} - \delta_{ik}$$, for

QUAIDS

where $\mu_i = \beta_i + \frac{2\lambda_i}{b(p)}\left(\ln\left[\frac{m}{a(p)}\right]\right)$ and δ_{ik} = Kronecker delta

Finally, from the Slutsky's Equation, the Hicksian or compensated elasticties are calculated via the formula; $\varepsilon_{ij}^{c} = \varepsilon_{ij}^{u} + \eta_i w_j$, where ε_{ij}^{u} is the uncompensated price elasticity of i with respect to j and η_i is the budget elasticity of good i. The term w_j is the mean budget share of good j.

Estimation Techniques That Address Censoring in a Demand System

Two-Step Estimators

A class of estimation techniques that deal with censored systems equation is the two-step estimation procedure. In this paper we highlight the two approaches namely the Heien and Wessells (1990) approach and the Shonkwiler and Yen (1999) method. These techniques usually consist of estimating a binary choice model in the first step, whose purpose is to account for those households that purchased and did not purchase the said commodity. In this exercise a probit model was estimated where the outcome variable takes on two values namely those households that purchased (1) and those that did not purchase (0). Two important derivatives of the probit estimation include the calculation of the cumulative distribution function (cdf) and probability density function (pdf) from the choice model.

In the case of the Heien and Wessells (1990) approach, the calculation of the inverse Mill's ratio (ratio of the pdf and cdf) from the first step probit estimation is now included as an added regressor into the estimation of the demand system. We note however that for those households that consumed and did not consume the beverage item, the inverse mills ratio had the following formula:

(7) $$IMR = \frac{\phi(W_i^T \hat{\eta})}{\Phi(W_i^T \hat{\eta})}, \text{ for those that consume}$$

(8) $$IMR = \frac{-\phi(W_i^T \hat{\eta})}{1-\Phi(W_i^T \hat{\eta})}, \text{ for those that did not consume}$$

where $IMR, \phi(W_i\hat{\eta})$, $\Phi(W_i\hat{\eta})$ and W_i are the inverse mills ratio, pdf , cdf and vector of socio-demographic variables including income, race and region. Thus, the Heien and Wessells (1990) two step approach of estimating a demand system can be represented as:

(9) $$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{M}{a(p)}\right] + \upsilon_i IMR + \varepsilon, \text{ for AIDS}$$

(10) $$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{M}{a(p)}\right] + \frac{\lambda_i}{b(p)}\left[\ln\left(\frac{M}{a(p)}\right)\right]^2 + \upsilon_i IMR + \varepsilon, \text{ for}$$

QUAIDS

On the other hand, the Shonkwiler and Yen (1999) consistent two step approach utilizes the calculated cdf to multiply the whole RHS variables of the share equation and include the pdf as an additional regressor in the system of budget shares. In notational form this can be represented as:

(11) $$w_i = \Phi(W_i\eta)\left[\alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{M}{a(p)}\right)\right] + \phi(W_i\eta) + \varepsilon, \text{ for AIDS}$$

(12)

$$w_i = \Phi(W_i\eta)\left[\alpha_i + \sum_{j=1}^{n}\gamma_{ij}\ln p_j + \beta_i \ln\left(\frac{M}{a(p)}\right) + \frac{\lambda_i}{b(p)}\left(\ln\left(\frac{M}{a(p)}\right)\right)^2\right] + \phi(W_i\eta) + \varepsilon$$, for

QUAIDS

Dong, Gould and Kaiser Approach (2004)

In this paper, we use the Dong, Gould and Kaiser (2004a) approach which is a variant of the Amemiya-Tobin model in estimating a censored AIDS model. In this approach the AIDS demand model can be written as:

(13) $$w_i^* = \alpha_i + \sum_{j=1}^{n}\gamma_{ij}\ln p_j + \beta_i \ln\left[\frac{M}{a(p)}\right] + \varepsilon,$$

where $w_i^* = p_i q_i$ represents the latent budget share with p_i and q_i correspond to the price and quantity of ith beverage. As pointed out by Stockton, Capps and Dong (2007), the censored system will take into account the latent budget share if the vector mapping of the latent shares to its corresponding actual shares addresses the following conditions concerning the latent share; w_i^*. These conditions are i) $0 \le w_i \le 1$ and ii) $\sum_i w_i = 1$. Thus, Dong, Gould and Kaiser (2004a) proposed an approach that addresses both restrictions by applying the following mapping condition;

(14) $$w_i = \frac{w_i^*}{\sum_{j\varepsilon\Omega} w_j^*}$$, if $w_i^* > 0$ and Ω corresponds to the positive latent share space.

$$w_i = 0, \qquad \text{if } w_i^* \le 0$$

In this mapping rule, we find that not only is the adding up condition for latent and observed shares satisfied but because the rule addressed the two constraints imposed on the latent share, non-negative expenditure shares are expected. As for the estimation procedure, the error structure of the respective share equation assumes a multivariate normal distribution, thus the method of maximum simulated likelihood was used to evaluate the integrals inherent in a multivariate normal distribution.

Generalized Maximum Entropy Procedure

Following the SAS ETS 9.2 ENTROPY Procedure guide (SAS ETS 9.2 User Guide, 2008), the procedure selects the parameter estimates consistent with the maximization of the entropy distribution. Thus, the entropy metric for a given distribution is given as;

$$\text{(15)} \qquad \max\ -\sum_{i=1}^{n} p_i \ln(p_i) \ \text{ s.t. } \sum_{i=1}^{n} p_i = 1,$$

where p_i is the probability of the ith support point.

In a regression framework, since this method assumes no parametric assumptions about the error terms and coefficients, a transformation known as reparameterization is necessary in order to identify the said parameters. For a two point support case, a reparametrized error term can be written as $\varepsilon = r_{z1}e_{z1} + r_{z2}e_{z2}$ where r_1 and r_2 are associated weights of the error term's upper and lower bound values of e_1 and e_2. As for the reparameterized coefficients, this can be written as $\beta = p_{h1}s_{h1} + p_{h2}s_{h2}$ where p_1 and p_2 represent the probabilities of β and s_1 and s_2 are the upper and lower bounds values based on prior information involving β. From this reparameterization, the GME maximization problem can be notationally written as:

(16)
$$
\begin{aligned}
\max\ & G(p,r) = -p'\ln(p) - r'\ln(r) \\
\text{s.t.}\ & q = X\,S\,p + E\,r \\
& 1_H = (I_H \Theta 1_L^1)\, p \\
& 1_Z = (I_Z \Theta 1_L^1)\, r
\end{aligned}
$$

where q is the vector of response variable, X is the matrix of independent covariate observations. *S* and *p* denote the vectors of weight and their associated probabilities with respect to β, while r is the weight associated with the boundary point contained in E. And finally I_H and I_Z are identity matrices. The symbol Θ is the Kronecker product.

However for this exercise, we deal with censored shares in a demand system such that we make modifications in solving the primal problem of the entropy procedure found in equation 16. For example, given that q = w_i is the share in the AIDS model, $w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{M}{a(p)}\right] + \varepsilon$, we apply the following conditions:

$$
w_i^* = \begin{matrix} \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{M}{a(p)}\right) + \varepsilon : w_i > 0 \\ lowerbound : w_i \le 0 \end{matrix}
$$

Thus for this case, the primal optimization problem can be written as

$$
\begin{aligned}
\max\ & G(p,r) = -p'\ln(p) - r'\ln(r) \\
\text{s.t}\ & w_i = \left[\alpha_i + \sum_{i=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{M}{a(p)}\right)\right] Sp + Er
\end{aligned}
$$

$$w_i^{LB} \le \left[\alpha_i + \sum_{i=1}^{n}\gamma_{ij}\ln p_j + \beta_i \ln\left(\frac{M}{a(p)}\right)\right]^{LB} S^{LB} p + E^{LB} r^{LB}$$

$$1_H = (\mathrm{I}_{\mathrm{H}} \Theta 1_L^1)\, p$$

$$1_Z = (\mathrm{I}_{\mathrm{Z}} \Theta 1_L^1)\, r$$

Following this example in the AIDS model, a similar construction can be done in the QUAIDS model.

Estimation Issues

This research also attempted to use the Dirichlet distribution to model the censored shares of the non-alcoholic beverage demand system. The Dirichlet distribution is a multivariate generalization of the beta distribution and imposes following properties in terms of modeling the shares in the demand system; i) $0 < w_i < 1$ and ii) $\sum_i w_i = 1$. However, in the Dirichlet distribution, densities do not exist in the distribution's boundaries (0 and 1) and therefore only those observations that are in the interior are valid. Thus, modeling censored demand systems with Dirichlet distribution is not possible.

The estimation of the AIDS and QUAIDS specification using the maximum entropy technique was done using the experimental SAS procedure called PROC ENTROPY. However, this experimental procedure at present is only limited to estimation of systems of linear regressions. Thus, attempts were made to linearize the demand system by using the starting values generated from the ITSUR specification and simplifying through the use of mean values of the non linear components such as the nonlinear price index *ln(a(p)*) and Cobb–Douglas price aggregator *b(p)* into constants in both the AIDS and QUAIDS model. Thus, in this case, the linearized AIDS and QUAIDS model can be represented as:

(18) $$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \Delta_i + \varepsilon \text{, for AIDS}$$

Where $\Delta_i = \ln\left(\frac{M}{C}\right)$ and ln C is a calculated constant of ln *a(p)*

(19) $$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \Delta + \lambda_i \Gamma^2 + \varepsilon \text{, for QUAIDS}$$

where $\Gamma^2 = \frac{\left[\ln\left(\frac{M}{C}\right)\right]^2}{D}$ with lnC as the calculated constant of ln *a(p*) and D is the constant representing the Cobb-Douglas price aggregator *b(p)*.

Another simplification that the study did was to forego the imposition of classical restriction of adding up, symmetry and homogeneity in the maximum entropy estimation of the demand system. This is because of the difficulty of identifying the values of support points of those coefficients being restricted. And with so many restrictions being imposed, the identification of problematic constraints becomes a major problem. Thus, the estimation of the AIDS and QUAIDS models were done without the usual imposition of the classical theoretical constraints.

The usage of the Dong, Gould and Kaiser (2004) technique was only performed in the AIDS model. This study did not attempt to use it in a QUAIDS model specification. This was primarily due to the highly non linear nature of the QUAIDS model and because the estimation approach requires the evaluation of multivariate integrals in a highly non linear model, convergence maybe difficult to achieve.

Data

The data used in the study is the 1999 AC Nielsen HomeScan Panel where the data set is a compilation of household purchase transactions of the said year. In this data set, the household's transaction records with respect to total expenditures and quantities of commodities are purchased primarily in retail groceries which include the usage of either discounts or coupons. The household transactions are performed by the use of scanner equipment. The number of household used is 7, 195 and because it was further disaggregated by quarter the total sample size numbered to 28,780. This sample size can be thought of a national representative sample of the huge amounts of item purchases of U.S. households for the year 1999.

In this study, the various specific socio-demographic variables used in the study were household income, household size, race, region and seasonal indicator for quarter. From Table 5.1, we find the mean household income is $51,740 and dominant household size for the sample is those with two members (38%). As for race, approximately 94 percent are white and black households. As for region, 33 percent come from South while rest has the following shares: east (20%), Central (25%) and West (20%).

Another feature of the data set is that commodity prices are not readily available, instead one uses the derivation of total expenditures over total quantity of the purchased item and it is called unit values and this is used as a proxy for the item price. If both the

Table 5.1. Descriptive Statistics of Relevant Household Demographic Variables

Variables	Mean	Std. Deviation	Min	Max
Household Income($)	51,740	26,254	5,000	100,000
Household Size (%)				
One member	22	41	0	1
Two members	38	48	0	1
Three members	16	37	0	1
Four members	15	36	0	1
Five members	10	29	0	1
Race (%)				
White	84	37	0	1
Black	10	30	0	1
Oriental	1	11	0	1
Other	5	22	0	1
Region (%)				
East	20	40	0	1
Central	25	43	0	1
South	34	47	0	1
West	20	40	0	1
Quarter (%)				
Q1	25	43	0	1
Q2	25	43	0	1
Q3	25	43	0	1
Q4	25	43	0	1
Observations	28,780			

expenditures and quantities were zero, then this study utilized a simple price imputation procedure where the process rested on the use of income, race and regional dummy variables. If $P_i = 0$, then

Pfruitjuice = 4.53912 + (*hinc**0.00000345) + (*white**-0.0885) + (*black**-0.24972) + (*oriental**0.01158) + (*central**-0.07377) + (*south**-0.02857) + (*west**0.60825);

Ptea = 2.07429 + (*hinc**0.00000716) + (*white**-0.39710) + (*black**-0.08642) + (*oriental**-0.13340) + (*central**0.03567) + (*south**-0.29073) + (*west**0.24558);

Pcoffee = 1.26359 + (*hinc**0.00000539) + (*white**-0.26017) + (*black**-0.18400) + (*oriental**0.86170)+ (*central**0.10697) + (*south**0.00532) + (*west**0.33853);

Pcsd = 2.29327 + (*hinc**0.0000006510327) + (*white**0.02942) + (*black**0.03566) + (*oriental**0.14496) + (*central**0.07624) + (*south**0.16520)+ (*west**0.21459);

Pwater = 1.98661 + (*hinc**0.00000218) + (*white**0.04082) + (*black**-0.06763) + (*oriental**0.01389) + (*central**-0.00548) + (*south**-0.06986) + (*west**-0.20992);

Pmilk = 3.21833 + (*hinc**-0.000000112181) + (*white**-0.13875) + (*black**0.28677) + (*oriental**0.22932) + (*central**-0.24758) + (*south**-0.05396) + (*west**0.17670);

Tables 5.2, 5.3 and 5.4 present the mean total expenditures, quantity purchased and prices for the 6 non-alcoholic beverages used in the study. In this case we find that the top household purchases with respect to non-alcoholic beverages were carbonated soft drinks, fruit juices, milk and coffee. The mean price are as follows fruit juices ($4.71/gal), tea ($2.06/gal), coffee ($1.41/gal), carbonated soft drinks ($2.48/gal), bottled water ($2.06/gal) and milk ($3.08/gal). On the other hand, Table 5.5 presents the mean budget shares of the beverage items. For the period 1999, approximately 81 percent of total expenditures for non alcoholic beverages are captured by carbonated soft

Table 5.2. Descriptive Statistics for Total Expenditure for Each Non-Alcoholic Beverage Item (n=28,780)

	Mean ($)	Std. Deviation ($)	Min ($)	Max ($)
Fruit Juices	14.19	19.15	0	268.82
Tea	3.42	7.36	0	177.26
Coffee	8.45	13.21	0	230.59
Carbonated Soft Drinks	31.14	41.24	0	1814.93
Bottled Water	3.02	8.34	0	206.96
Milk	22.86	23.87	0	304.05

Table 5.3. Descriptive Statistics for Quantities for Each Non Alcoholic Beverage Item (n=28,780)

	Mean (gallons)	Std. Deviation (gallons)	Min (gallons)	Max (gallons)
Fruit Juices	3.17	4.25	0	63.31
Tea	2.76	6.03	0	137.50
Coffee	8.27	13.73	0	305.51
Carbonated Soft Drinks	13.27	16.83	0	681.75
Bottled Water	2.44	7.51	0	151.45
Milk	8.30	9.22	0	98.00

Table 5.4. Descriptive Statistics for Prices[1] for Each Non-Alcoholic Beverage Item (n=28,780)

	Mean ($/gallon)	Std. Deviation ($/gallon)	Min ($/gallon)	Max ($/gallon)
Fruit Juices	4.71	1.31	0.99	15.09
Tea	2.06	1.24	0.08	16.08
Coffee	1.41	1.32	0.13	16.03
Carbonated Soft Drinks	2.48	0.85	0.30	11.44
Bottled Water	2.06	1.04	0.05	12.83
Milk	3.08	0.89	0.88	15.56

[1] When expenditure and quantities are equal to zero, price imputation was used where if qty=0 then P_i=f(income,race and region).

Table 5.5. Mean Budget Shares for Each Beverage Item for Calendar Year 1999

Beverage Product	Average Budget Share	Std. Deviation	Min	Max
Fruit Juices	0.175	0.188	0.000	1.000
Tea	0.047	0.096	0.000	1.000
Coffee	0.109	0.153	0.000	1.000
Carbonated Soft Drinks	0.343	0.247	0.000	1.000
Bottled Water	0.038	0.094	0.000	1.000
Milk	0.288	0.210	0.000	1.000

drinks, fruit juices and milk. The 19 percent are devoted to tea (4.7 %), coffee (11%) and bottled water (3.8 %).

Table 5.6 describes the degree of censoring associated with each type of non-alcoholic beverages. From the table items with minimal to medium censoring are milk (6.77%), carbonated soft drinks (8.84 %) and fruit juices (23.09 %). On the other hand the remaining highly censored non-alcoholic beverage items are tea (54.88 %), coffee (42.77 %) and bottled water (60.65 %).

Empirical Results

Estimated Demand Parameters

Both the censored AIDS and QUAIDS specifications and their unrestricted analogs were estimated using the various techniques addressing the censoring issue. These included the Iterated Seemingly Unrelated regression (ITSUR), the two step procedure approaches; Heien & Wessells (1990) and Shonkwiler & Yen (1999), the Generalized Maximum Entropy and the Simulated Maximum Likelihood estimation (Dong et al., 2004a). Tables 5.7 to 5.10 provide the estimated parameters of AIDS and QUAIDS plus their unrestricted specifications.

Almost all of the socio-demographic parameters in both specifications and across all estimation techniques are statistically significant. Also, almost all of the parameters in both AIDS and QUAIDS and across estimation techniques are relatively close to one another and the same can be said for the AIDS and QUAIDS unrestricted cases. Thus it can be postulated that because of a relatively large sample size, the various estimation

Table 5.6. Number of Censored Responses for Each Beverage Item

	Number of Observations	Percentage
Fruit Juices	6,646	23.09
Tea	15,795	54.88
Coffee	12,310	42.77
Carbonated Soft Drinks	2,544	8.84
Bottled Water	17,454	60.65
Milk	1,949	6.77

Table 5.7. Parameter Estimates of Different Censoring Techniques for AIDS Estimation

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value	Dong et al Actual	T-value
I. Fruit Juice										
constant (af0)	0.214472	<.0001	0.23759	<.0001	0.219606	<.0001	0.128333	<.0001	-0.1565	-6.3447
hinc (afa)	7.98E-07	<.0001	8.59E-07	<.0001	9.42E-07	<.0001	6.93E-07	<.0001	0.044	20.2886
hs2(afb)	-0.01708	<.0001	-0.01751	<.0001	-0.02443	<.0001	-0.01304	<.0001	-0.0222	-5.4082
hs3(afc)	-0.0299	<.0001	-0.02968	<.0001	-0.0407	<.0001	-0.02392	<.0001	-0.0365	-6.6273
hs4(afd)	-0.04384	<.0001	-0.04243	<.0001	-0.05824	<.0001	-0.03647	<.0001	-0.0535	-9.041
hs5(afe)	-0.04321	<.0001	-0.04035	<.0001	-0.05902	<.0001	-0.03435	<.0001	-0.0523	-7.6816
white(aff)	-0.03569	<.0001	-0.03543	<.0001	-0.04462	<.0001	-0.03073	<.0001	-0.053	-7.6536
black(afg)	0.056458	<.0001	0.053506	<.0001	0.062805	<.0001	0.058266	<.0001	0.0698	8.673
oriental(afh)	0.058398	<.0001	0.054969	<.0001	0.064485	<.0001	0.054533	<.0001	0.0814	6.1896
central(afi)	-0.05463	<.0001	-0.05483	<.0001	-0.06587	<.0001	-0.05309	<.0001	-0.0701	-14.5763
south(afj)	-0.0382	<.0001	-0.03899	<.0001	-0.04285	<.0001	-0.03852	<.0001	-0.0504	-11.4912
west(afk)	-0.05601	<.0001	-0.05792	<.0001	-0.06122	<.0001	-0.06379	<.0001	-0.0794	-16.005
Q1(afl)	-0.00255	0.402	-0.00141	0.6164	-0.00383	0.3215	-0.00251	0.4083	-0.0029	-0.6597
Q2(afm)	-0.01287	<.0001	-0.01064	0.0002	-0.01737	<.0001	-0.01109	0.0003	-0.0172	-3.8355
Q3(afn)	-0.01244	<.0001	-0.00997	0.0004	-0.01645	<.0001	-0.01074	0.0004	-0.0193	-4.3565
lpf(gff)	-0.00018	0.961	0.007987	0.0181	0.002367	0.6077	0.016855	<.0001	-0.0203	-3.6021
lpt(gft)	0.000672	0.6064	-0.00051	0.6983	-0.00409	0.0205	0.002177	0.2267	0.0066	2.8842
lpc(gfc)	0.014688	<.0001	0.010834	<.0001	0.021914	<.0001	0.025443	<.0001	0.0265	8.172
lps(gfs)	-0.03636	<.0001	-0.03467	<.0001	-0.04115	<.0001	-0.01617	<.0001	0.0079	3.3476
lpw(gfw)	-0.00306	0.0293	-0.00392	0.0021	-0.00634	0.0016	-0.00652	0.0041	-0.0087	-3.7526
lpm(gfm)	0.02424	<.0001	0.02029	<.0001	0.0273	<.0001	0.05256	<.0001		
bf	0.004014	0.0063	-0.00695	<.0001	0.003695	0.0468	0.007302	<.0001	0.005	2.5906
pf			0.04244	<.0001						
zf					0.154526	0.0109				

Table 5.7 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value	Dong et al Actual	T-value
II. Coffee										
constant (ac0)	0.065597	<.0001	0.129077	<.0001	0.202503	0.0068	0.11659	<.0001	0.0751	3.8193
hinc (aca)	-3.99E-08	0.2611	-7.28E-09	0.8111	-2.62E-07	0.0001	2.30E-08	0.5283	-0.0078	-4.3928
hs2(acb)	0.012944	<.0001	0.004756	0.028	0.020336	<.0001	0.011339	<.0001	0.0236	6.2644
hs3(acc)	-0.022	<.0001	-0.0232	<.0001	-0.0417	<.0001	-0.02426	<.0001	-0.0319	-6.0685
hs4(acd)	-0.02987	<.0001	-0.02913	<.0001	-0.05379	<.0001	-0.03261	<.0001	-0.0443	-7.7021
hs5(ace)	-0.03579	<.0001	-0.03235	<.0001	-0.06595	<.0001	-0.03915	<.0001	-0.0537	-7.7525
white(acf)	0.018138	<.0001	0.019822	<.0001	0.004924	0.7136	0.015943	0.0001	0.0304	4.4625
black(acg)	-0.02062	<.0001	-0.02618	<.0001	0.019014	0.425	-0.02246	<.0001	-0.041	-4.9089
oriental(ach)	0.001977	0.8188	-0.01509	0.0405	0.046877	0.0998	0.003139	0.7128	0.0111	0.6569
central(aci)	-0.00988	0.0002	-0.01195	<.0001	0.009873	0.4256	-0.01037	<.0001	-0.0186	-4.3013
south(acj)	-0.01742	<.0001	-0.019	<.0001	-0.01579	0.0266	-0.0175	<.0001	-0.0267	-6.2609
west(ack)	0.00872	0.0018	0.003549	0.137	0.036907	0.0014	0.012525	<.0001	0.0097	2.1751
Q1(acl)	-0.00328	0.1846	-0.00225	0.2871	-0.00689	0.1025	-0.00294	0.2337	-0.0055	-1.4365
Q2(acm)	-0.0123	<.0001	-0.00813	0.0001	-0.02353	<.0001	-0.01242	<.0001	-0.0244	-6.0645
Q3(can)	-0.01018	<.0001	-0.00678	0.0014	-0.02055	<.0001	-0.01038	<.0001	-0.0215	-5.3264
lpf(gfc)	0.014688	<.0001	0.010834	<.0001	0.021914	<.0001	0.00149	0.6438	0.0265	8.172
lpt(gtc)	0.001992	0.0411	-0.00272	0.005	0.020093	<.0001	0.000145	0.9211	0.0043	2.5191
lpc(gcc)	-0.06452	<.0001	-0.04213	<.0001	-0.09507	<.0001	-0.0691	<.0001	-0.1117	-40.2576
lps(gcs)	0.010469	<.0001	0.007857	0.0001	0.010647	0.0117	0.000615	0.8444	0.0213	6.4296
lpw(gcw)	0.007774	<.0001	0.004823	<.0001	0.0169	<.0001	0.004692	0.011	0.0188	9.9521
lpm(gcm)	0.029593	<.0001	0.021335	<.0001	0.025515	<.0001	0.01873	<.0001		
bc	-0.00098	0.4119	-0.01559	<.0001	0.000419	0.8376	-0.00351	0.004	0.0026	1.3483
pc			0.056251	<.0001						
zc					-0.08333	0.4306				

Table 5.7 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value	Dong et al Actual	T-value
III. Carbonated Soft Drinks										
constant (as0)	0.196319	<.0001	0.214279	<.0001	0.085908	0.0107	0.143986	<.0001	0.5269	30.7133
hinc (asa)	-3.66E-07	<.0001	-3.23E-07	<.0001	1.27E-07	0.55	-5.38E-07	<.0001	-0.0305	-20.084
hs2(asb)	-0.01101	0.0078	-0.01313	0.0011	-0.01146	0.0115	-0.00582	0.1602	-0.0048	-1.0681
hs3(asc)	0.014921	0.0035	0.011912	0.0165	0.016997	0.0023	0.021685	<.0001	0.0302	4.7479
hs4(asd)	0.007759	0.1472	0.005664	0.2771	0.008014	0.1686	0.015893	0.0033	0.0205	3.055
hs5(ase)	-0.01083	0.0755	-0.01193	0.0444	-0.01132	0.0875	-0.00112	0.8552	-0.0007	-0.0959
white(asf)	-0.01792	0.0077	-0.01726	0.008	-0.03591	0.0002	-0.01207	0.0717	-0.0009	-0.1634
black(asg)	-0.0098	0.2158	-0.01322	0.0859	-0.02779	0.0101	-0.00777	0.3252	0.0105	1.2538
oriental(ash)	-0.09292	<.0001	-0.09465	<.0001	-0.17557	<.0001	-0.09695	<.0001	-0.1154	-6.3265
central(asi)	0.076523	<.0001	0.076186	<.0001	0.095466	<.0001	0.077509	<.0001	0.1151	20.2808
south(asj)	0.055064	<.0001	0.054924	<.0001	0.071278	<.0001	0.055863	<.0001	0.083	15.059
west(ask)	0.038646	<.0001	0.036423	<.0001	0.040851	<.0001	0.029935	<.0001	0.0619	10.2527
Q1(asl)	-0.00229	0.5706	-0.00198	0.6142	-0.00216	0.6234	-0.00325	0.4197	0.0045	0.8924
Q2(asm)	0.035846	<.0001	0.035323	<.0001	0.040252	<.0001	0.036678	<.0001	0.0487	9.4008
Q3(asn)	0.025694	<.0001	0.025005	<.0001	0.029413	<.0001	0.026503	<.0001	0.0331	6.3035
lpf(gfs)	-0.03636	<.0001	-0.03467	<.0001	-0.04115	<.0001	-0.03095	<.0001	-0.0485	-11.6719
lpt(gts)	0.004836	0.0011	0.003496	0.0298	0.005699	0.0415	0.010563	<.0001	0.0079	3.3476
lpc(gcs)	0.010469	<.0001	0.007857	0.0001	0.010647	0.0117	0.017874	<.0001	0.0213	6.4296
lps(gss)	-0.00442	0.3568	-0.00136	0.7693	-0.0083	0.1352	0.009411	0.0658	-0.005	-0.8126
lpw(gsw)	0.01128	<.0001	0.007153	<.0001	0.012525	<.0001	0.019091	<.0001	0.0179	7.3878
lpm(gsm)	0.0142	<.0001	0.017532	<.0001	0.020577	<.0001	0.034187	<.0001		
bs	0.04851	<.0001	0.04169	<.0001	0.052889	<.0001	0.054016	<.0001	0.0514	24.3444
ps			0.032694	<.0001						
zs					0.611293	<.0001				

Table 5.7 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value	Dong et al Actual	T-value
IV. Water										
constant (aw0)	0.051332	<.0001	0.088	<.0001	0.087653	0.0874	0.014262	0.0059	-0.2621	-31.8639
hinc (awa)	2.35E-07	<.0001	3.15E-08	0.1081	5.05E-07	0.0295	1.86E-07	<.0001	0.0223	25.7652
hs2(awb)	-0.01574	<.0001	-0.01305	<.0001	-0.02319	<.0001	-0.01446	<.0001	-0.0369	-14.3957
hs3(awc)	-0.01338	<.0001	-0.01246	<.0001	-0.01941	<.0001	-0.01162	<.0001	-0.0292	-8.6243
hs4(awd)	-0.01789	<.0001	-0.01865	<.0001	-0.02791	<.0001	-0.01593	<.0001	-0.031	-8.1557
hs5(awe)	-0.01834	<.0001	-0.01798	<.0001	-0.02957	<.0001	-0.01608	<.0001	-0.0342	-7.6966
white(awf)	-0.01419	<.0001	-0.00351	0.1186	-0.03622	0.012	-0.01252	<.0001	-0.0366	-8.6262
black(awg)	0.018541	<.0001	0.014063	<.0001	0.041239	<.0001	0.02049	<.0001	0.0267	5.6107
oriental(awh)	0.003686	0.4835	0.007951	0.0911	0.001508	0.9017	0.003302	0.5191	-0.0076	-0.9225
central(awi)	-0.007	<.0001	-0.00109	0.4511	-0.01506	0.0184	-0.00733	<.0001	-0.0192	-6.3875
south(awj)	-0.00246	0.1008	-0.00127	0.3433	-0.00472	0.1212	-0.00222	0.1369	-0.0078	-2.8603
west(awk)	0.001469	0.3899	0.0007	0.6475	0.002552	0.5063	-0.00182	0.2912	0.0005	0.1555
Q1(awl)	-0.00679	<.0001	-0.00373	0.0058	-0.01049	<.0001	-0.00709	<.0001	-0.0196	-6.7781
Q2(awm)	0.003286	0.0303	0.00288	0.034	0.004472	0.0705	0.002744	0.0698	0.0048	2.0111
Q3(awn)	0.007918	<.0001	0.004951	0.0003	0.011802	<.0001	0.007448	<.0001	0.0185	6.9359
lpf(gfw)	-0.00306	0.0293	-0.00392	0.0021	-0.00634	0.0016	0.012044	<.0001	-0.0087	-3.7526
lpt(gtw)	0.002819	<.0001	-0.00075	0.2724	0.011376	<.0001	0.006699	<.0001	0.0095	6.9618
lpc(gcw)	0.007774	<.0001	0.004823	<.0001	0.0169	<.0001	0.010618	<.0001	0.0188	9.9521
lps(gsw)	0.01128	<.0001	0.007153	<.0001	0.012525	<.0001	0.017538	<.0001	0.0179	7.3878
lpw(gww)	-0.03863	<.0001	-0.02198	<.0001	-0.06014	<.0001	-0.03603	<.0001	-0.0864	-54.3579
lpm(gwm)	0.01982	<.0001	0.014676	<.0001	0.025679	<.0001	0.021395	<.0001		
bw	-0.00252	0.0005	-0.00942	<.0001	-0.00289	0.0145	-0.00161	0.0303	0.0132	10.72
pw			0.037163	<.0001						
zw					0.012815	0.8692				

Table 5.7 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value	Dong et al Actual	T-value
V. Milk										
constant (am0)	0.377373	<.0001	0.39274	<.0001	0.283723	<.0001	0.540447	<.0001		
hinc (ama)	-7.10E-07	<.0001	-6.53E-07	<.0001	-3.19E-07	0.0072	-4.26E-07	<.0001		
hs2(amb)	0.032931	<.0001	0.030971	<.0001	0.035746	<.0001	0.02293	<.0001		
hs3(amc)	0.052592	<.0001	0.05096	<.0001	0.056512	<.0001	0.038544	<.0001		
hs4(amd)	0.088858	<.0001	0.088101	<.0001	0.095301	<.0001	0.071789	<.0001		
hs5(ame)	0.112908	<.0001	0.113932	<.0001	0.122047	<.0001	0.092776	<.0001		
white(amf)	0.04869	<.0001	0.053929	<.0001	0.063972	<.0001	0.037542	<.0001		
black(amg)	-0.0427	<.0001	-0.04191	<.0001	-0.08646	<.0001	-0.046	<.0001		
oriental(amh)	0.033911	0.0041	0.036459	0.0013	0.020234	0.1565	0.042021	0.0003		
central(ami)	0.018784	<.0001	0.019153	<.0001	0.031074	<.0001	0.016128	<.0001		
south(amj)	0.018609	<.0001	0.019418	<.0001	0.027692	<.0001	0.017834	<.0001		
west(amk)	0.026552	<.0001	0.025821	<.0001	0.022959	<.0001	0.044341	<.0001		
Q1(aml)	0.015226	<.0001	0.015816	<.0001	0.015902	<.0001	0.016005	<.0001		
Q2(amm)	-0.01613	<.0001	-0.01488	<.0001	-0.01742	<.0001	-0.01899	<.0001		
Q3(amn)	-0.01286	0.0002	-0.01212	0.0002	-0.01379	0.0001	-0.01557	<.0001		
lpf(gfm)	0.02424	<.0001	0.02029	<.0001	0.0273	<.0001	-0.00779	0.0753		
lpt(gtm)	0.004002	0.0056	0.016702	<.0001	-0.00664	0.0062	-0.00715	0.0003		
lpc(gcm)	0.029593	<.0001	0.021335	<.0001	0.025515	<.0001	0.010791	0.0002		
lps(gsm)	0.0142	<.0001	0.017532	<.0001	0.020577	<.0001	-0.01891	<.0001		
lpw(gwm)	0.01982	<.0001	0.014676	<.0001	0.025679	<.0001	0.017642	<.0001		
lpm(gmm)	-0.09185	<.0001	-0.09054	<.0001	-0.09243	<.0001	-0.1471	<.0001		
bm	-0.03648	<.0001	-0.04419	<.0001	-0.03924	<.0001	-0.04407	<.0001		
pm			0.04714	<.0001						
zm					0.634439	<.0001				
gtt	-0.01432	<.0001	-0.01622	<.0001	-0.02644	<.0001				
at0	0.094907	<.0001	-0.06169	<.0001	0.120607	0.0677				

Table 5.7 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value	Dong et al Actual	T-value
VI. Tea										
constant (am0)									-0.0616	-5.6556
hinc (ama)									0.012	12.9142
hs2(amb)									-0.0048	-1.7551
hs3(amc)									-0.0045	-1.199
hs4(amd)									-0.0083	-2.0865
hs5(ame)									-0.007	-1.4864
white(amf)									-0.0035	-1.1978
black(amg)									-0.0058	-1.3305
oriental(amh)									-0.014	-1.4741
central(ami)									-0.0464	-14.5257
south(amj)									-0.0276	-9.4982
west(amk)									-0.0362	-11.0767
Q1(aml)									-0.0001	-0.0249
Q2(amm)									0.0059	2.2412
Q3(amn)									0.0053	1.8728
lpf(gft)									0.0066	2.8842
lpt(gtt)									-0.04	-24.2229
lpc(gtc)									0.0043	2.5191
lps(gts)									0.0079	3.3476
lpw(gtw)									0.0095	6.9618
bm									-0.0145	-10.7511

Table 5.8. Parameter Estimates of Different Censoring Techniques for QUAIDS Estimation

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max Entropy	P-value
I. Fruit Juice								
constant (af0)	0.170773	<.0001	0.207559	<.0001	0.199746	<.0001	0.091257	<.0001
hinc (afa)	0.000001	<.0001	8.59E-07	<.0001	6.06E-07	0.0009	6.93E-07	<.0001
hs2(afb)	-0.017390	<.0001	-0.01771	<.0001	-0.02477	<.0001	-0.01334	<.0001
hs3(afc)	-0.028940	<.0001	-0.029	<.0001	-0.03924	<.0001	-0.02323	<.0001
hs4(afd)	-0.041430	<.0001	-0.04073	<.0001	-0.05481	<.0001	-0.03477	<.0001
hs5(afe)	-0.039830	<.0001	-0.03799	<.0001	-0.05439	<.0001	-0.032	<.0001
white(aff)	-0.035170	<.0001	-0.03501	<.0001	-0.03545	<.0001	-0.03028	<.0001
black(afg)	0.056113	<.0001	0.053298	<.0001	0.056322	<.0001	0.058094	<.0001
oriental(afh)	0.060538	<.0001	0.056488	<.0001	0.070516	<.0001	0.055337	<.0001
central(afi)	-0.054570	<.0001	-0.05479	<.0001	-0.05357	<.0001	-0.05304	<.0001
south(afj)	-0.038350	<.0001	-0.03907	<.0001	-0.03304	<.0001	-0.03858	<.0001
west(afk)	-0.056150	<.0001	-0.05804	<.0001	-0.0425	0.0001	-0.0638	<.0001
Q1(afl)	-0.002610	0.3908	-0.00142	0.6146	-0.00404	0.2956	-0.0025	0.4102
Q2(afm)	-0.012910	<.0001	-0.01061	0.0002	-0.01756	<.0001	-0.01114	0.0003
Q3(afn)	-0.012600	<.0001	-0.01003	0.0004	-0.01682	<.0001	-0.01087	0.0004
lpf(gff)	-0.002840	0.4445	0.007467	0.0277	-0.00181	0.7029	0.017361	<.0001
lpt(gft)	0.001701	0.2005	-0.00287	0.0467	-0.00097	0.5987	0.002311	0.1993
lpc(gfc)	0.010472	<.0001	0.009464	<.0001	0.01337	<.0001	0.025439	<.0001
lps(gfs)	-0.033690	<.0001	-0.03316	<.0001	-0.03672	<.0001	-0.01526	<.0001
lpw(gfw)	-0.001770	0.2158	-0.0033	0.0107	-0.0037	0.0745	-0.00654	0.0039
lpm(gfm)	0.026132	<.0001	0.022391	<.0001	0.029829	<.0001	0.053057	<.0001
bf	0.037883	<.0001	0.015828	0.0012	0.051421	<.0001	0.033981	<.0001
lf	-0.006030	<.0001	-0.00402	<.0001	-0.0082	<.0001	-0.00486	<.0001
pf			0.042948	<.0001				
zf					0.041737	0.4792		

Table 5.8 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max Entropy	P-value
II. Coffee								
constant (ac0)	-0.010190	0.2119	0.062059	<.0001	0.20217	<.0001	0.049244	<.0001
hinc (aca)	0.000000	0.2189	-9.82E-09	0.746	-3.14E-07	<.0001	2.39E-08	0.5107
hs2(acb)	0.012016	<.0001	0.004078	0.0584	0.018907	<.0001	0.010765	<.0001
hs3(acc)	-0.020750	<.0001	-0.02187	<.0001	-0.0396	<.0001	-0.02293	<.0001
hs4(acd)	-0.026110	<.0001	-0.02551	<.0001	-0.04811	<.0001	-0.02933	<.0001
hs5(ace)	-0.030410	<.0001	-0.0273	<.0001	-0.05728	<.0001	-0.03462	<.0001
white(acf)	0.017941	<.0001	0.019864	<.0001	-0.01377	0.1863	0.016803	<.0001
black(acg)	-0.022200	<.0001	-0.02745	<.0001	0.058613	0.0012	-0.02279	<.0001
oriental(ach)	0.004886	0.5699	-0.01238	0.0912	0.095345	0.0001	0.004677	0.5828
central(aci)	-0.010050	0.0001	-0.01207	<.0001	0.031827	0.0005	-0.01027	<.0001
south(acj)	-0.017930	<.0001	-0.0194	<.0001	-0.00469	0.4124	-0.0176	<.0001
west(ack)	0.008250	0.003	0.003125	0.1883	0.057608	<.0001	0.012509	<.0001
Q1(acl)	-0.003590	0.1452	-0.00246	0.242	-0.0068	0.1061	-0.00292	0.236
Q2(acm)	-0.012620	<.0001	-0.0083	<.0001	-0.02372	<.0001	-0.0125	<.0001
Q3(acn)	-0.010690	<.0001	-0.00714	0.0007	-0.02094	<.0001	-0.01061	<.0001
lpf(gfc)	0.010472	<.0001	0.009464	<.0001	0.01337	<.0001	0.002463	0.4438
lpt(gtc)	0.003795	0.0002	-0.00811	<.0001	0.024437	<.0001	0.000403	0.7829
lpc(gcc)	-0.071550	<.0001	-0.04536	<.0001	-0.11061	<.0001	-0.0691	<.0001
lps(gcs)	0.014694	<.0001	0.011647	<.0001	0.022269	<.0001	0.002367	0.4495
lpw(gcw)	0.009988	<.0001	0.006423	<.0001	0.023245	<.0001	0.004639	0.0118
lpm(gcm)	0.032605	<.0001	0.025931	<.0001	0.027291	<.0001	0.019685	<.0001
bc	0.058849	<.0001	0.036218	<.0001	0.098272	<.0001	0.047798	<.0001
lc	-0.010600	<.0001	-0.0091	<.0001	-0.0164	<.0001	-0.00934	<.0001
pc			0.056598	<.0001				
zc					-0.27906	0.0001		

Table 5.8 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max Entropy	P-value
III. Carbonated Soft Drinks								
constant (as0)	0.287611	<.0001	0.326417	<.0001	0.189931	<.0001	0.226915	<.0001
hinc (asa)	0.000000	<.0001	-3.10E-07	<.0001	8.38E-08	0.684	-5.39E-07	<.0001
hs2(asb)	-0.010060	0.0149	-0.0122	0.0024	-0.01038	0.0219	-0.00502	0.2254
hs3(asc)	0.013243	0.0094	0.009599	0.0525	0.015562	0.0051	0.019843	0.0001
hs4(asd)	0.003055	0.5691	-0.00025	0.9616	0.003983	0.4949	0.011321	0.0364
hs5(ase)	-0.017490	0.0043	-0.02001	0.0008	-0.01716	0.0099	-0.00744	0.228
white(asf)	-0.018220	0.0066	-0.01734	0.0073	-0.03579	0.0002	-0.01328	0.0473
black(asg)	-0.008370	0.2893	-0.01153	0.1318	-0.02586	0.0159	-0.00732	0.3533
oriental(ash)	-0.096640	<.0001	-0.09942	<.0001	-0.17657	<.0001	-0.09912	<.0001
central(asi)	0.076621	<.0001	0.07641	<.0001	0.094539	<.0001	0.077371	<.0001
south(asj)	0.055531	<.0001	0.055445	<.0001	0.07091	<.0001	0.056009	<.0001
west(ask)	0.039208	<.0001	0.037076	<.0001	0.042126	<.0001	0.02996	<.0001
Q1(asl)	-0.002000	0.6192	-0.00163	0.6773	-0.00231	0.5984	-0.00328	0.4146
Q2(asm)	0.036151	<.0001	0.035592	<.0001	0.040078	<.0001	0.036796	<.0001
Q3(asn)	0.026234	<.0001	0.025565	<.0001	0.029419	<.0001	0.026828	<.0001
lpf(gfs)	-0.033690	<.0001	-0.03316	<.0001	-0.03672	<.0001	-0.03231	<.0001
lpt(gts)	0.003189	0.0335	0.011352	<.0001	0.001008	0.6803	0.010203	<.0001
lpc(gcs)	0.014694	<.0001	0.011647	<.0001	0.022269	<.0001	0.017885	<.0001
lps(gss)	-0.004450	0.3573	-0.00516	0.2781	-0.00806	0.143	0.006965	0.173
lpw(gsw)	0.009859	<.0001	0.004772	0.0008	0.004875	0.0502	0.019166	<.0001
lpm(gsm)	0.010402	0.0028	0.010546	0.0018	0.016628	<.0001	0.032854	<.0001
bs	-0.023230	0.001	-0.04397	<.0001	-0.01689	0.0276	-0.01757	0.0152
ls	0.012729	<.0001	0.014866	<.0001	0.011634	<.0001	0.013037	<.0001
ps			0.035701	<.0001				
zs					0.5855	<.0001		

Table 5.8 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max Entropy	P-value
IV. Water								
constant (aw0)	0.074274	<.0001	0.10597	<.0001	-0.04179	0.3203	0.038078	<.0001
hinc (awa)	0.000000	<.0001	3.39E-08	0.084	-3.00E-07	0.1127	1.86E-07	<.0001
hs2(awb)	-0.015470	<.0001	-0.01277	<.0001	-0.02256	<.0001	-0.01426	<.0001
hs3(awc)	-0.013780	<.0001	-0.01269	<.0001	-0.0202	<.0001	-0.01207	<.0001
hs4(awd)	-0.019060	<.0001	-0.01949	<.0001	-0.03017	<.0001	-0.01705	<.0001
hs5(awe)	-0.020010	<.0001	-0.01918	<.0001	-0.03295	<.0001	-0.01764	<.0001
white(awf)	-0.014050	<.0001	-0.00319	0.1555	0.012752	0.2821	-0.01281	<.0001
black(awg)	0.019139	<.0001	0.014733	<.0001	0.038107	<.0001	0.020612	<.0001
oriental(awh)	0.002970	0.572	0.007572	0.1075	0.025583	0.0227	0.002776	0.587
central(awi)	-0.006930	<.0001	-0.00099	0.492	0.005837	0.2768	-0.00736	<.0001
south(awj)	-0.002280	0.1278	-0.0011	0.4135	0.001873	0.5111	-0.00218	0.1435
west(awk)	0.001688	0.3225	0.000941	0.5384	-0.00582	0.105	-0.00181	0.2926
Q1(awl)	-0.006680	<.0001	-0.00361	0.0075	-0.01023	<.0001	-0.00709	<.0001
Q2(awm)	0.003392	0.0252	0.002998	0.0272	0.00496	0.0445	0.002775	0.0665
Q3(awn)	0.008086	<.0001	0.00512	0.0002	0.012397	<.0001	0.007529	<.0001
lpf(gfw)	-0.001770	0.2158	-0.0033	0.0107	-0.0037	0.0745	0.011716	<.0001
lpt(gtw)	0.002286	0.0012	0.001362	0.0684	0.010468	<.0001	0.006611	<.0001
lpc(gcw)	0.009988	<.0001	0.025931	<.0001	0.023245	<.0001	0.01062	<.0001
lps(gsw)	0.009859	<.0001	0.004772	0.0008	0.004875	0.0502	0.016937	<.0001
lpw(gww)	-0.039230	<.0001	-0.02252	<.0001	-0.06193	<.0001	-0.03601	<.0001
lpm(gwm)	0.018870	<.0001	0.013257	<.0001	0.027046	<.0001	0.021083	<.0001
bw	-0.020760	<.0001	-0.02365	<.0001	-0.04092	<.0001	-0.0193	<.0001
lw	0.003239	<.0001	0.002482	<.0001	0.006509	<.0001	0.003223	<.0001
pw			0.03713	<.0001				
zw					0.287215	<.0001		

Table 5.8 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max Entropy	P-value
V. Milk								
constant (am0)	0.372864	<.0001	0.408687	<.0001	0.307543	<.0001	0.524688	<.0001
hinc (ama)	-0.000001	<.0001	-6.55E-07	<.0001	-4.50E-07	0.0001	-4.25E-07	<.0001
hs2(amb)	0.032792	<.0001	0.03094	<.0001	0.035132	<.0001	0.022713	<.0001
hs3(amc)	0.052592	<.0001	0.050421	<.0001	0.055931	<.0001	0.039042	<.0001
hs4(amd)	0.089030	<.0001	0.087038	<.0001	0.094797	<.0001	0.073024	<.0001
hs5(ame)	0.113146	<.0001	0.112551	<.0001	0.121474	<.0001	0.094484	<.0001
white(amf)	0.048320	<.0001	0.054143	<.0001	0.059077	<.0001	0.037872	<.0001
black(amg)	-0.043200	<.0001	-0.04154	<.0001	-0.07559	<.0001	-0.04612	<.0001
oriental(amh)	0.033459	0.0046	0.035624	0.0017	0.025358	0.075	0.042608	0.0003
central(ami)	0.018673	<.0001	0.019098	<.0001	0.028144	<.0001	0.016166	<.0001
south(amj)	0.018509	<.0001	0.019602	<.0001	0.025014	<.0001	0.017796	<.0001
west(amk)	0.026246	<.0001	0.025682	<.0001	0.024688	<.0001	0.044334	<.0001
Q1(aml)	0.015126	<.0001	0.015872	<.0001	0.015893	<.0001	0.016014	<.0001
Q2(amm)	-0.016250	<.0001	-0.01486	<.0001	-0.01741	<.0001	-0.01902	<.0001
Q3(amn)	-0.012990	0.0001	-0.01206	0.0002	-0.01378	0.0001	-0.01565	<.0001
lpf(gfm)	0.026132	<.0001	0.022391	<.0001	0.029829	<.0001	-0.00742	0.0903
lpt(gtm)	0.003660	0.012	0.020484	<.0001	-0.00768	0.0002	-0.00706	0.0004
lpc(gcm)	0.032605	<.0001	0.025931	<.0001	0.027291	<.0001	0.010788	0.0002
lps(gsm)	0.010402	0.0028	0.010546	0.0018	0.016628	<.0001	-0.01825	<.0001
lpw(gwm)	0.018870	<.0001	0.013257	<.0001	0.027046	<.0001	0.017622	<.0001
lpm(gmm)	-0.091670	<.0001	-0.09261	<.0001	-0.09312	<.0001	-0.14674	<.0001
bm	-0.032280	<.0001	-0.05595	<.0001	-0.0365	<.0001	-0.02473	<.0001
lm	-0.000740	0.4636	0.002048	0.0294	-0.00036	0.7227	-0.00352	0.0009
pm			0.048897	<.0001				
zm					0.498066	<.0001		
gtt	-0.014630	<.0001	-0.02222	<.0001	-0.02727	<.0001		
at0	0.104673	<.0001	-0.11069	<.0001	0.142398	0.0008		
bt	-0.020460	<.0001	0.071514	<.0001	-0.05538	<.0001		

Table 5.9. Parameter Estimates of Different Censoring Techniques for AIDS Estimation (Unrestricted)

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value
I. Fruit Juice								
constant (af0)	0.120829	<.0001	0.157005	<.0001	0.099739	0.0006	0.128333	<.0001
hinc (afa)	6.94E-07	<.0001	7.59E-07	<.0001	8.22E-07	<.0001	6.93E-07	<.0001
hs2(afb)	-0.01285	<.0001	-0.01396	<.0001	-0.01902	<.0001	-0.01304	<.0001
hs3(afc)	-0.02364	<.0001	-0.02453	<.0001	-0.03258	<.0001	-0.02392	<.0001
hs4(afd)	-0.03615	<.0001	-0.03616	<.0001	-0.04831	<.0001	-0.03647	<.0001
hs5(afe)	-0.03401	<.0001	-0.03285	<.0001	-0.04682	<.0001	-0.03435	<.0001
white(aff)	-0.0307	<.0001	-0.03096	<.0001	-0.04015	<.0001	-0.03073	<.0001
black(afg)	0.058184	<.0001	0.055244	<.0001	0.064239	<.0001	0.058266	<.0001
oriental(afh)	0.054438	<.0001	0.050954	<.0001	0.062202	<.0001	0.054533	<.0001
central(afi)	-0.05312	<.0001	-0.05393	<.0001	-0.06458	<.0001	-0.05309	<.0001
south(afj)	-0.03853	<.0001	-0.03923	<.0001	-0.04437	<.0001	-0.03852	<.0001
west(afk)	-0.06381	<.0001	-0.06516	<.0001	-0.07458	<.0001	-0.06379	<.0001
Q1(afl)	-0.0025	0.4098	-0.00142	0.613	-0.00371	0.3361	-0.00251	0.4083
Q2(afm)	-0.0111	0.0003	-0.00938	0.0009	-0.01483	0.0001	-0.01109	0.0003
Q3(afn)	-0.01077	0.0004	-0.00881	0.0019	-0.01412	0.0003	-0.01074	0.0004
lpf(gff)	0.017989	<.0001	0.024776	<.0001	0.02359	<.0001	0.016855	<.0001
lpt(gft)	-0.01817	0.4817	-0.01714	0.1197	-0.02498	0.3113	0.002177	0.2267
lpc(gfc)	0.026287	<.0001	0.022831	<.0001	0.017041	0.0011	0.025443	<.0001
lps(gfs)	-0.01592	<.0001	-0.0161	<.0001	-0.01605	0.0012	-0.01617	<.0001
lpw(gfw)	-0.00623	0.0063	-0.00616	0.0035	-0.01506	<.0001	-0.00652	0.0041
lpm(gfm)	0.055842	<.0001	0.044247	<.0001	0.070089	<.0001	0.05256	<.0001
bf	0.006872	<.0001	-0.00422	0.0026	0.006963	0.0002	0.007302	<.0001
pf			0.042621	<.0001				
zf					0.162905	0.0129		

Table 5.9 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value
II. Coffee								
constant (ac0)	0.118101	<.0001	0.175941	<.0001	-1.83385	<.0001	0.11659	<.0001
hinc (aca)	2.04E-08	0.5761	3.26E-08	0.2938	9.89E-07	<.0001	2.30E-08	0.5283
hs2(acb)	0.010877	<.0001	0.003393	0.118	0.017825	<.0001	0.011339	<.0001
hs3(acc)	-0.02492	<.0001	-0.02532	<.0001	-0.04426	<.0001	-0.02426	<.0001
hs4(acd)	-0.03337	<.0001	-0.03193	<.0001	-0.05596	<.0001	-0.03261	<.0001
hs5(ace)	-0.03998	<.0001	-0.03576	<.0001	-0.06838	<.0001	-0.03915	<.0001
white(acf)	0.015892	0.0001	0.020714	<.0001	0.317918	<.0001	0.015943	0.0001
black(acg)	-0.02225	<.0001	-0.02536	<.0001	-0.58441	<.0001	-0.02246	<.0001
oriental(ach)	0.003359	0.697	-0.01113	0.1299	-0.51884	<.0001	0.003139	0.7128
central(aci)	-0.01034	<.0001	-0.01277	<.0001	-0.32311	<.0001	-0.01037	<.0001
south(acj)	-0.01744	<.0001	-0.01793	<.0001	-0.19237	<.0001	-0.0175	<.0001
west(ack)	0.012586	<.0001	0.006807	0.0049	-0.2511	<.0001	0.012525	<.0001
Q1(acl)	-0.00295	0.2334	-0.00149	0.4794	-0.00619	0.1416	-0.00294	0.2337
Q2(acm)	-0.01244	<.0001	-0.0081	0.0001	-0.0213	<.0001	-0.01242	<.0001
Q3(can)	-0.01036	<.0001	-0.00688	0.0012	-0.01821	<.0001	-0.01038	<.0001
lpf(gcf)	0.001002	0.7564	-0.00189	0.4933	0.00129	0.8152	0.00149	0.6438
lpt(gct)	0.007957	0.4444	-0.07595	0.0408	0.01551	0.3056	0.000145	0.9211
lpc(gcc)	-0.06924	<.0001	-0.04716	<.0001	-0.10587	<.0001	-0.0691	<.0001
lps(gcs)	0.000199	0.9495	-0.00233	0.3949	-0.00148	0.7837	0.000615	0.8444
lpw(gcw)	0.004663	0.0118	0.002683	0.0924	0.012315	0.0003	0.004692	0.011
lpm(gcm)	0.017495	<.0001	0.006217	0.0637	0.030717	<.0001	0.01873	<.0001
bc	-0.00262	0.0311	-0.01736	<.0001	-0.00384	0.0649	-0.00351	0.004
pc			0.057273	<.0001				
zc					2.91415	<.0001		

Table 5.9 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value
III. Carbonated Soft Drinks								
constant (as0)	0.079025	<.0001	0.107221	<.0001	0.005497	0.8784	0.143986	<.0001
hinc (asa)	-5.35E-07	<.0001	-4.82E-07	<.0001	-3.55E-07	0.097	-5.38E-07	<.0001
hs2(asb)	-0.00562	0.1766	-0.008	0.0484	-0.00621	0.1724	-0.00582	0.1602
hs3(asc)	0.022062	<.0001	0.018411	0.0002	0.023602	<.0001	0.021685	<.0001
hs4(asd)	0.01633	0.0025	0.01341	0.011	0.015771	0.0074	0.015893	0.0033
hs5(ase)	-0.00065	0.9165	-0.00254	0.6732	-0.00217	0.7465	-0.00112	0.8552
white(asf)	-0.01235	0.0666	-0.01226	0.0616	-0.01983	0.0401	-0.01207	0.0717
black(asg)	-0.00804	0.3108	-0.01068	0.1668	-0.01644	0.1277	-0.00777	0.3252
oriental(ash)	-0.09756	<.0001	-0.1009	<.0001	-0.14218	<.0001	-0.09695	<.0001
central(asi)	0.077335	<.0001	0.076828	<.0001	0.089018	<.0001	0.077509	<.0001
south(asj)	0.05575	<.0001	0.05511	<.0001	0.066322	<.0001	0.055863	<.0001
west(ask)	0.029905	<.0001	0.028655	<.0001	0.031656	<.0001	0.029935	<.0001
Q1(asl)	-0.00321	0.4261	-0.00325	0.4084	-0.00325	0.4586	-0.00325	0.4197
Q2(asm)	0.036638	<.0001	0.035345	<.0001	0.04037	<.0001	0.036678	<.0001
Q3(asn)	0.026391	<.0001	0.025037	<.0001	0.029435	<.0001	0.026503	<.0001
lpf(gsf)	-0.02274	<.0001	-0.02046	<.0001	-0.02499	<.0001	-0.03095	<.0001
lpt(gst)	-0.14758	0.4533	0.216142	0.0277	-0.20111	0.3125	0.010563	<.0001
lpc(gsc)	0.024684	<.0001	0.024985	<.0001	-0.09046	<.0001	0.017874	<.0001
lps(gss)	0.013186	0.0278	0.021732	<.0001	0.007004	0.3083	0.009411	0.0658
lpw(gsw)	0.021362	<.0001	0.022064	<.0001	-0.02304	0.0013	0.019091	<.0001
lpm(gsm)	0.060126	<.0001	0.052037	<.0001	0.0636	<.0001	0.034187	<.0001
bs	0.053321	<.0001	0.046191	<.0001	0.058813	<.0001	0.054016	<.0001
ps			0.032445	<.0001				
zs					0.403847	0.0086		

Table 5.9 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value
IV. Water								
constant (aw0)	0.015818	0.0047	0.065179	<.0001	-0.76235	<.0001	0.014262	0.0059
hinc (awa)	1.86E-07	<.0001	-1.02E-08	0.6107	-3.02E-06	<.0001	1.86E-07	<.0001
hs2(awb)	-0.0146	<.0001	-0.01191	<.0001	-0.0204	<.0001	-0.01446	<.0001
hs3(awc)	-0.01181	<.0001	-0.01115	<.0001	-0.01522	<.0001	-0.01162	<.0001
hs4(awd)	-0.01615	<.0001	-0.0174	<.0001	-0.02291	<.0001	-0.01593	<.0001
hs5(awe)	-0.01633	<.0001	-0.01655	<.0001	-0.02375	<.0001	-0.01608	<.0001
white(awf)	-0.01257	<.0001	-0.00026	0.9067	0.177208	<.0001	-0.01252	<.0001
black(awg)	0.02051	<.0001	0.017098	<.0001	0.025008	<.0001	0.02049	<.0001
oriental(awh)	0.003357	0.523	0.010236	0.0294	0.118177	<.0001	0.003302	0.5191
central(awi)	-0.00734	<.0001	-0.00126	0.3823	0.074172	<.0001	-0.00733	<.0001
south(awj)	-0.00222	0.1382	-0.00038	0.7753	0.019595	<.0001	-0.00222	0.1369
west(awk)	-0.0018	0.2971	-0.00135	0.3831	-0.04053	<.0001	-0.00182	0.2912
Q1(awl)	-0.00709	<.0001	-0.0035	0.0093	-0.01085	<.0001	-0.00709	<.0001
Q2(awm)	0.002733	0.0717	0.002549	0.0603	0.004399	0.0757	0.002744	0.0698
Q3(awn)	0.007448	<.0001	0.004583	0.0007	0.011911	<.0001	0.007448	<.0001
lpf(gwf)	0.011791	<.0001	0.007075	<.0001	0.019015	<.0001	0.012044	<.0001
lpt(gwt)	0.010845	0.0562	-0.03585	0.0657	0.016548	0.0126	0.006699	<.0001
lpc(gwc)	0.010443	<.0001	0.005222	<.0001	0.020464	<.0001	0.010618	<.0001
lps(gws)	0.017363	<.0001	0.010045	<.0001	0.026262	<.0001	0.017538	<.0001
lpw(gww)	-0.03605	<.0001	-0.01967	<.0001	-0.05591	<.0001	-0.03603	<.0001
lpm(gwm)	0.020641	<.0001	0.011315	<.0001	0.034418	<.0001	0.021395	<.0001
bw	-0.0014	0.0587	-0.00905	<.0001	-0.00142	0.2387	-0.00161	0.0303
pw			0.038215	<.0001				
zw					1.207913	<.0001		

Table 5.9 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max. Entropy	P-value
V. Milk								
constant (am0)	0.595231	<.0001	0.622839	<.0001	0.476322	<.0001	0.540447	<.0001
hinc (ama)	-4.27E-07	<.0001	-3.70E-07	<.0001	1.46E-07	0.2448	-4.26E-07	<.0001
hs2(amb)	0.023139	<.0001	0.019202	<.0001	0.025623	<.0001	0.02293	<.0001
hs3(amc)	0.038783	<.0001	0.03458	<.0001	0.042178	<.0001	0.038544	<.0001
hs4(amd)	0.072065	<.0001	0.068431	<.0001	0.07762	<.0001	0.071789	<.0001
hs5(ame)	0.093081	<.0001	0.090795	<.0001	0.100767	<.0001	0.092776	<.0001
white(amf)	0.037798	<.0001	0.037661	<.0001	0.06107	<.0001	0.037542	<.0001
black(amg)	-0.04599	<.0001	-0.04899	<.0001	-0.11081	<.0001	-0.046	<.0001
oriental(amh)	0.042321	0.0003	0.038684	0.0006	0.017848	0.2102	0.042021	0.0003
central(ami)	0.016262	<.0001	0.015664	<.0001	0.031461	<.0001	0.016128	<.0001
south(amj)	0.017868	<.0001	0.017408	<.0001	0.031825	<.0001	0.017834	<.0001
west(amk)	0.044288	<.0001	0.043272	<.0001	0.035049	<.0001	0.044341	<.0001
Q1(aml)	0.015966	<.0001	0.01537	<.0001	0.017081	<.0001	0.016005	<.0001
Q2(amm)	-0.01894	<.0001	-0.01925	<.0001	-0.02038	<.0001	-0.01899	<.0001
Q3(amn)	-0.01548	<.0001	-0.0162	<.0001	-0.01669	<.0001	-0.01557	<.0001
lpf(gmf)	-0.01456	0.0008	-0.01668	<.0001	-0.01542	0.0014	-0.00779	0.0753
lpt(gmt)	0.123771	0.4477	-0.23732	0.0311	0.163821	0.3047	-0.00715	0.0003
lpc(gmc)	0.005021	0.0857	0.003202	0.2645	0.098117	<.0001	0.010791	0.0002
lps(gms)	-0.02161	<.0001	-0.02634	<.0001	-0.01838	0.001	-0.01891	<.0001
lpw(gmw)	0.015659	<.0001	0.013294	<.0001	0.05287	<.0001	0.017642	<.0001
lpm(gmm)	-0.1684	<.0001	-0.16844	<.0001	-0.17687	<.0001	-0.1471	<.0001
bm	-0.04416	<.0001	-0.05177	<.0001	-0.04713	<.0001	-0.04407	<.0001
pm			0.049404	<.0001				
zm					0.859325	<.0001		
gtt								
at0	-2.95849	0.4216	4.524465	0.0365	-3.61141	0.2849		
bt								

Table 5.10. Parameter Estimates of Different Censoring Techniques for QUAIDS Estimation (Unrestricted)

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max Entropy	P-value
I. Fruit Juice								
constant (af0)	0.109324	<.0001	0.144883	<.0001	0.029642	0.3045	0.091257	<.0001
hinc (afa)	7.00E-07	<.0001	7.61E-07	<.0001	1.16E-06	<.0001	6.93E-07	<.0001
hs2(afb)	-0.01307	<.0001	-0.01414	<.0001	-0.01889	<.0001	-0.01334	<.0001
hs3(afc)	-0.02353	<.0001	-0.02432	<.0001	-0.03247	<.0001	-0.02323	<.0001
hs4(afd)	-0.03597	<.0001	-0.03538	<.0001	-0.04839	<.0001	-0.03477	<.0001
hs5(afe)	-0.0339	<.0001	-0.03176	<.0001	-0.04725	<.0001	-0.032	<.0001
white(aff)	-0.03155	<.0001	-0.03095	<.0001	-0.04949	<.0001	-0.03028	<.0001
black(afg)	0.057652	<.0001	0.054952	<.0001	0.069414	<.0001	0.058094	<.0001
oriental(afh)	0.053676	<.0001	0.051799	<.0001	0.058822	<.0001	0.055337	<.0001
central(afi)	-0.05407	<.0001	-0.0539	<.0001	-0.07945	<.0001	-0.05304	<.0001
south(afj)	-0.03976	<.0001	-0.03922	<.0001	-0.05598	<.0001	-0.03858	<.0001
west(afk)	-0.0646	<.0001	-0.06508	<.0001	-0.09775	<.0001	-0.0638	<.0001
Q1(afl)	-0.00259	0.3944	-0.00142	0.6133	-0.00429	0.2653	-0.0025	0.4102
Q2(afm)	-0.01092	0.0004	-0.00937	0.0009	-0.01575	<.0001	-0.01114	0.0003
Q3(afn)	-0.01049	0.0006	-0.00886	0.0018	-0.0147	0.0002	-0.01087	0.0004
lpf(gff)	0.017866	<.0001	0.024335	<.0001	0.021496	<.0001	0.017361	<.0001
lpt(gft)	-0.01224	0.0174	-0.00317	0.3059	0.007402	0.0195	0.002311	0.1993
lpc(gfc)	0.026185	<.0001	0.021961	<.0001	0.016225	0.0473	0.025439	<.0001
lps(gfs)	-0.01615	<.0001	-0.01555	<.0001	-0.01749	0.0004	-0.01526	<.0001
lpw(gfw)	-0.00586	0.0102	-0.00617	0.0035	-0.01322	<.0001	-0.00654	0.0039
lpm(gfm)	0.056285	<.0001	0.044377	<.0001	0.070427	<.0001	0.053057	<.0001
bf	0.012664	<.0001	0.008023	0.083	0.017254	<.0001	0.033981	<.0001
lf	-0.00049	<.0001	-0.00235	0.0048	-0.0009	<.0001	-0.00486	<.0001
pf			0.043097	<.0001				
zf					0.282385	<.0001		

Table 5.10 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max Entropy	P-value
II. Coffee								
constant (ac0)	0.111606	<.0001	0.127712	<.0001	-3.6965	<.0001	0.049244	<.0001
hinc (aca)	2.39E-08	0.5129	3.52E-08	0.2566	2.12E-06	<.0001	2.39E-08	0.5107
hs2(acb)	0.010788	<.0001	0.002739	0.2056	0.016188	0.0002	0.010765	<.0001
hs3(acc)	-0.02485	<.0001	-0.0246	<.0001	-0.04581	<.0001	-0.02293	<.0001
hs4(acd)	-0.03326	<.0001	-0.02922	<.0001	-0.05777	<.0001	-0.02933	<.0001
hs5(ace)	-0.03989	<.0001	-0.03195	<.0001	-0.06904	<.0001	-0.03462	<.0001
white(acf)	0.01566	0.0001	0.020842	<.0001	0.609494	<.0001	0.016803	<.0001
black(acg)	-0.02231	<.0001	-0.02611	<.0001	-1.1468	<.0001	-0.02279	<.0001
oriental(ach)	0.003202	0.7105	-0.00788	0.2821	-1.04757	<.0001	0.004677	0.5828
central(aci)	-0.01075	<.0001	-0.01256	<.0001	-0.63723	<.0001	-0.01027	<.0001
south(acj)	-0.018	<.0001	-0.01787	<.0001	-0.35914	<.0001	-0.0176	<.0001
west(ack)	0.01226	<.0001	0.007109	0.0032	-0.52624	<.0001	0.012509	<.0001
Q1(acl)	-0.00295	0.2326	-0.00156	0.4583	-0.00274	0.507	-0.00292	0.236
Q2(acm)	-0.0123	<.0001	-0.00812	0.0001	-0.01645	<.0001	-0.0125	<.0001
Q3(can)	-0.01017	<.0001	-0.0071	0.0008	-0.01511	0.0003	-0.01061	<.0001
lpf(gcf)	0.001021	0.7518	-0.0033	0.2354	0.009152	0.093	0.002463	0.4438
lpt(gct)	0.00671	0.034	-0.01474	<.0001	-0.0192	<.0001	0.000403	0.7829
lpc(gcc)	-0.0693	<.0001	-0.05012	<.0001	-0.16896	<.0001	-0.0691	<.0001
lps(gcs)	0.000103	0.9739	0.000531	0.8455	0.003133	0.5569	0.002367	0.4495
lpw(gcw)	0.004866	0.0087	0.003393	0.0346	0.010092	0.0021	0.004639	0.0118
lpm(gcm)	0.018052	<.0001	0.008222	0.0135	0.030066	<.0001	0.019685	<.0001
bc	0.00031	0.8388	0.026469	<.0001	-0.05622	<.0001	0.047798	<.0001
lc	-0.00024	0.0033	-0.00826	<.0001	0.00499	<.0001	-0.00934	<.0001
pc			0.057371	<.0001				
zc					5.746743	<.0001		

Table 5.10 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max Entropy	P-value
III. Carbonated Soft Drinks								
constant (as0)	0.068912	<.0001	0.200245	<.0001	-0.02957	0.4096	0.226915	<.0001
hinc (asa)	-5.30E-07	<.0001	-4.83E-07	<.0001	-3.85E-07	0.0689	-5.39E-07	<.0001
hs2(asb)	-0.00574	0.1674	-0.00739	0.0666	-0.00577	0.2044	-0.00502	0.2254
hs3(asc)	0.022211	<.0001	0.01619	0.0012	0.024035	<.0001	0.019843	0.0001
hs4(asd)	0.016544	0.0022	0.007406	0.1595	0.016271	0.0057	0.011321	0.0364
hs5(ase)	-0.00047	0.9397	-0.01059	0.078	-0.00202	0.7631	-0.00744	0.228
white(asf)	-0.01283	0.0565	-0.01316	0.0434	-0.01709	0.0755	-0.01328	0.0473
black(asg)	-0.00829	0.2955	-0.01001	0.1921	-0.01775	0.0988	-0.00732	0.3533
oriental(ash)	-0.09801	<.0001	-0.10778	<.0001	-0.14548	<.0001	-0.09912	<.0001
central(asi)	0.076594	<.0001	0.076296	<.0001	0.087166	<.0001	0.077371	<.0001
south(asj)	0.054797	<.0001	0.054834	<.0001	0.066066	<.0001	0.056009	<.0001
west(ask)	0.029277	<.0001	0.027933	<.0001	0.028723	<.0001	0.02996	<.0001
Q1(asl)	-0.00324	0.4212	-0.00324	0.4076	-0.00418	0.3397	-0.00328	0.4146
Q2(asm)	0.036821	<.0001	0.035168	<.0001	0.039021	<.0001	0.036796	<.0001
Q3(asn)	0.026654	<.0001	0.025268	<.0001	0.028575	<.0001	0.026828	<.0001
lpf(gsf)	-0.02331	<.0001	-0.01801	0.0005	-0.03131	<.0001	-0.03231	<.0001
lpt(gst)	-0.11075	<.0001	0.034682	<.0001	0.036379	0.0017	0.010203	<.0001
lpc(gsc)	0.024291	<.0001	0.030132	<.0001	-0.17846	<.0001	0.017885	<.0001
lps(gss)	0.013024	0.0114	0.017326	0.0006	0.005935	0.3213	0.006965	0.173
lpw(gsw)	0.021893	<.0001	0.021237	<.0001	-0.00937	0.1218	0.019166	<.0001
lpm(gsm)	0.059799	<.0001	0.049326	<.0001	0.061522	<.0001	0.032854	<.0001
bs	0.058028	<.0001	-0.03779	<.0001	0.073441	<.0001	-0.01757	0.0152
ls	-0.0004	0.0033	0.015778	<.0001	-0.00134	<.0001	0.013037	<.0001
ps			0.035683	<.0001				
zs					0.399182	0.0086		

Table 5.10 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max Entropy	P-value
IV. Water								
constant (aw0)	0.01936	0.0007	0.080486	<.0001	-0.53882	<.0001	0.038078	<.0001
hinc (awa)	1.84E-07	<.0001	-1.06E-08	0.5983	-2.02E-06	<.0001	1.86E-07	<.0001
hs2(awb)	-0.01453	<.0001	-0.01172	<.0001	-0.0205	<.0001	-0.01426	<.0001
hs3(awc)	-0.01184	<.0001	-0.0114	<.0001	-0.01537	<.0001	-0.01207	<.0001
hs4(awd)	-0.0162	<.0001	-0.01829	<.0001	-0.02303	<.0001	-0.01705	<.0001
hs5(awe)	-0.01636	<.0001	-0.01778	<.0001	-0.02393	<.0001	-0.01764	<.0001
white(awf)	-0.01233	<.0001	-0.00023	0.9189	0.116325	<.0001	-0.01281	<.0001
black(awg)	0.020657	<.0001	0.017496	<.0001	0.029346	<.0001	0.020612	<.0001
oriental(awh)	0.003573	0.4966	0.009248	0.0491	0.083243	<.0001	0.002776	0.587
central(awi)	-0.00706	<.0001	-0.00133	0.3562	0.048269	<.0001	-0.00736	<.0001
south(awj)	-0.00186	0.2154	-0.00039	0.7708	0.012364	0.0007	-0.00218	0.1435
west(awk)	-0.00157	0.3651	-0.00144	0.351	-0.02989	<.0001	-0.00181	0.2926
Q1(awl)	-0.00707	<.0001	-0.00348	0.0097	-0.01115	<.0001	-0.00709	<.0001
Q2(awm)	0.002677	0.0777	0.00256	0.0591	0.004027	0.1041	0.002775	0.0665
Q3(awn)	0.007362	<.0001	0.004679	0.0006	0.01163	<.0001	0.007529	<.0001
lpf(gwf)	0.011816	<.0001	0.007706	<.0001	0.018637	<.0001	0.011716	<.0001
lpt(wt)	0.009538	<.0001	0.016926	<.0001	0.011615	<.0001	0.006611	<.0001
lpc(gwc)	0.010466	<.0001	0.006541	<.0001	0.027848	<.0001	0.01062	<.0001
lps(gws)	0.017426	<.0001	0.008667	<.0001	0.026071	<.0001	0.016937	<.0001
lpw(gww)	-0.03617	<.0001	-0.02022	<.0001	-0.05607	<.0001	-0.03601	<.0001
lpm(gwm)	0.020459	<.0001	0.009754	<.0001	0.03391	<.0001	0.021083	<.0001
bw	-0.00314	0.0007	-0.02299	<.0001	0.000737	0.6314	-0.0193	<.0001
lw	0.000146	0.0038	0.002646	<.0001	-0.00025	0.0069	0.003223	<.0001
pw			0.038072	<.0001				
zw					0.861922	<.0001		

Table 5.10 Continued

Parameters	ITSUR	P-value	Heien & Wessells	P-value	Shonkwiler & Yen	P-value	Generalized Max Entropy	P-value
V. Milk								
constant (am0)	0.583246	<.0001	0.627887	<.0001	0.463968	<.0001	0.524688	<.0001
hinc (ama)	-4.20E-07	<.0001	-3.67E-07	<.0001	8.70E-08	0.4838	-4.25E-07	<.0001
hs2(amb)	0.023129	<.0001	0.019312	<.0001	0.025888	<.0001	0.022713	<.0001
hs3(amc)	0.039104	<.0001	0.034673	<.0001	0.042351	<.0001	0.039042	<.0001
hs4(amd)	0.072474	<.0001	0.068512	<.0001	0.077746	<.0001	0.073024	<.0001
hs5(ame)	0.093482	<.0001	0.09097	<.0001	0.100714	<.0001	0.094484	<.0001
white(amf)	0.037682	<.0001	0.038345	<.0001	0.061616	<.0001	0.037872	<.0001
black(amg)	-0.04585	<.0001	-0.04845	<.0001	-0.10687	<.0001	-0.04612	<.0001
oriental(amh)	0.042235	0.0003	0.038793	0.0005	0.019061	0.1795	0.042608	0.0003
central(ami)	0.015589	<.0001	0.015696	<.0001	0.030425	<.0001	0.016166	<.0001
south(amj)	0.016926	<.0001	0.017466	<.0001	0.031552	<.0001	0.017796	<.0001
west(amk)	0.043699	<.0001	0.043184	<.0001	0.034954	<.0001	0.044334	<.0001
Q1(aml)	0.016005	<.0001	0.015477	<.0001	0.01651	<.0001	0.016014	<.0001
Q2(amm)	-0.01864	<.0001	-0.0192	<.0001	-0.02113	<.0001	-0.01902	<.0001
Q3(amn)	-0.01509	<.0001	-0.01614	<.0001	-0.01709	<.0001	-0.01565	<.0001
lpf(gmf)	-0.01391	0.0014	-0.01598	0.0001	-0.01359	0.0046	-0.00742	0.0903
lpt(gmt)	0.095511	<.0001	0.020502	<.0001	-0.01542	0.0504	-0.00706	0.0004
lpc(gmc)	0.005243	0.0724	0.004566	0.1063	0.197153	<.0001	0.010788	0.0002
lps(gms)	-0.02171	<.0001	-0.02847	<.0001	-0.02053	<.0001	-0.01825	<.0001
lpw(gmw)	0.015828	<.0001	0.011532	<.0001	0.042422	<.0001	0.017622	<.0001
lpm(gmm)	-0.16677	<.0001	-0.17284	<.0001	-0.17347	<.0001	-0.14674	<.0001
bm	-0.03901	<.0001	-0.05375	<.0001	-0.03874	<.0001	-0.02473	<.0001
lm	-0.000440	0.0003	0.000288	0.7597	-0.00072	<.0001	-0.00352	0.0009
pm			0.051058	<.0001				
zm					0.812745	<.0001		
gtt								
at0	-2.293840	<.0001	-0.52751	<.0001	0.263861	0.1244		
bt	0.328340	<.0001	0.323619	<.0001	-0.2258	<.0001		

procedures converged to yielding relatively close parameter estimates. Also, the parameters associated with the quadratic term in the QUAIDS specification are highly significant, suggesting in part the appropriateness of the QUAIDS specification over the AIDS model across estimation procedures and is also true for the unrestricted case. In Table 5.11, we find that the symmetry, homogeneity and the combination of both restrictions are rejected in both AIDS and QUAIDS models.

Expenditure, Uncompensated and Compensated Elasticities

Tables 5.12 to 5.23 present the calculated expenditure, uncompensated and compensated elasticities of non-alcoholic beverages across model specification, estimation techniques and imposition of theoretical restrictions. From the tables, we find that both expenditure elasticities and own-price elasticities were generally similar across model specification, estimation technique and whether the theoretical restrictions were imposed. All of the expenditure elasticities are positive indicating that all non-alcoholic beverages are normal goods. Also, if we look at the compensated cross-price elasticities across model specification, estimation technique and theoretical restriction, we find that almost all of them are positive indicating that the set of non-alcoholic beverages are net substitutes. Similarly, the major substitutes for fruit juice and tea are coffee, carbonated soft drink and milk. On the other hand the major substitutes for coffee are fruit juice, carbonated soft drinks and milk. For carbonated soft drinks the major substitutes are coffee and milk. Coffee, carbonated soft drinks and milk represent the major non-alcoholic beverage substitutes for bottled water. Finally, major commodity substitutes for milk are fruit juice, coffee and carbonated soft drinks.

Table 5.11. Symmetry, Homogeneity and Combination of Symmetry and Homogeneity Restriction Wald Tests

	Symmetry		Homogeneity		Symmetry and Homogeneity	
	χ^2-Statistic	p-value	χ^2-Statistic	p-value	χ^2-Statistic	p-value
A. AIDS model						
ITSUR	671.32	<.0001	367.24	<.0001	755.93	<.0001
H&W	610.79	<.0001	201.58	<.0001	730.66	<.0001
S&Y	561.91	<.0001	177.43	<.0001	624.23	<.0001
B. QUAIDS model						
ITSUR	664.31	<.0001	351.10	<.0001	726.78	<.0001
H&W	623.55	<.0001	745.17	<.0001	1027.90	<.0001
S&Y	594.46	<.0001	392.83	<.0001	1019.80	<.0001

Table 5.12. Expenditure Elasticities[1] of Non-Alcoholic Beverages Using the AIDS System and 1999 ACNielsen Homescan Data

Item	ITSUR Estimate	H&W Estimate	S&Y Estimate	GME Estimate	Dong et al. Actual Estimates	Dong et al. Latent Estimates	Mean	Standard Deviation
Fruit Juice	1.023	0.960	1.021	1.042	1.008	1.027	1.013	0.028
	(0.000)	(0.000)	(0.000)		(0.000)	(0.005)		
Tea	0.733	1.733	0.684	0.741	0.889	0.728	0.918	0.405
	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)		
Coffee	0.991	0.857	1.004	0.968	1.005	1.021	0.974	0.060
	(0.000)	(0.000)	(0.000)		(0.000)	(0.089)		
Carbonated Soft drinks	1.141	1.122	1.154	1.158	1.112	1.156	1.140	0.019
	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)		
Bottled Water	0.934	0.752	0.924	0.958	1.128	1.397	1.016	0.222
	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)		
Milk	0.873	0.847	0.864	0.847	0.864	0.790	0.848	0.030
	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)		

Note: p-values are in brackets

[1] Calculated using sample means

.

Table 5.13. Uncompensated Own- and Cross-Price Elasticity Matrix[1] of Non-Alcoholic Beverages Using the AIDS and the 1999 ACNielsen Homescan Data

								Carbonated		Bottled			
		Fruit Juice		Tea		Coffee		Soft Drinks		Water		Milk	
Fruit Juice	ITSUR	**-1.006**	[.0001]	0.002	[0.8293]	0.081	[.0001]	-0.212	[.0001]	-0.019	[.0196]	0.130	[.0001]
	H&W	**-0.945**	[.0001]	-0.005	[0.5003]	0.068	[.0001]	-0.191	[.0001]	-0.019	[.0136]	0.131	[.0001]
	S&Y	**-0.991**	[.0001]	-0.026	[0.0100]	0.120	[.0001]	-0.236	[.0001]	-0.038	[.0137]	0.150	[.0001]
	Dong et al (actual)	**-1.053**	[.0001]	0.016	[0.0095]	0.079	[.0001]	-0.143	[.0001]	-0.045	[.0001]	0.137	[.0001]
	Dong et al (latent)	**-1.105**	[.0001]	0.037	[0.0013]	0.138	[.0001]	-0.273	[.0001]	-0.040	[.0007]	0.216	[.0001]
	GME (unrestricted)	**-0.912**		0.009		0.141		-0.100		-0.182		0.173	
	Mean	**-1.002**		**0.005**		**0.104**		**-0.193**		**-0.057**		**0.156**	
	Std. Deviation	**0.070**		**0.021**		**0.032**		**0.063**		**0.062**		**0.034**	
Tea	ITSUR	0.071	[.0120]	**-1.279**	[.0001]	0.073	[.0004]	0.148	[.0002]	0.075	[.0001]	0.179	[.0001]
	H&W	-0.188	[.0001]	**-1.306**	[.0001]	-0.178	[.0001]	-0.065	[.0002]	-0.082	[.0001]	0.085	[.0001]
	S&Y	-0.018	[.6308]	**-1.528**	[.0001]	0.513	[.0001]	0.139	[.0194]	0.270	[.0001]	-0.058	[.2577
	Dong et al (actual)	0.035	[.0191]	**-1.298**	[0.0001]	0.050	[.0001]	0.121	[.0001]	0.017	[.0478]	0.186	[.0001]
	Dong et al (latent)	0.075	[.0408]	**-1.763**	[0.0001]	0.126	[.0001]	0.279	[.0001]	0.111	[.0001]	0.445	[.0001]
	GME (unrestricted)	0.231		**-1.242**		0.124		0.205		0.038		0.524	
	Mean	**0.034**		**-1.403**		**0.118**		**0.138**		**0.071**		**0.227**	
	Std. Deviation	**0.137**		**0.204**		**0.224**		**0.115**		**0.117**		**0.220**	
Coffee	ITSUR	0.137	[.0001]	0.019	[0.0325]	**-1.591**	[.0001]	0.098	[0.0001]	0.072	[.0001]	0.275	[.0001]
	H&W	0.134	[.0001]	-0.033	[0.0003]	**-1.363**	[.0001]	0.099	[0.0001]	0.057	[.0001]	0.248	[.0001]
	S&Y	0.200	[.0001]	0.1840	[0.0001]	**-1.873**	[.0001]	0.097	[0.0123]	0.155	[.0001]	0.233	[.0001]
	Dong et al (actual)	0.114	[.0001]	0.014	[0.0204]	**-1.447**	[.0001]	0.092	[.0001]	0.053	[.0137]	0.169	[.0001]
	Dong et al (latent)	0.219	[.0001]	0.036	[0.0047]	**-1.910**	[.0001]	0.163	[.0001]	0.158	[.0001]	0.313	[.0001]
	GME (unrestricted)	0.020		0.004		**-1.628**		0.011		0.045		0.183	
	Mean	**0.137**		**0.037**		**-1.635**		**0.093**		**0.090**		**0.237**	
	Std. Deviation	**0.071**		**0.075**		**0.221**		**0.048**		**0.052**		**0.055**	

Table 5.13 Continued

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Carbonated Soft drinks	ITSUR	-0.136	[.0001]	0.0004	[0.9340]	0.014	[.0300]	**-1.037**	[.0001]	0.025	[.0001]	-0.008	[.3993]
	H&W	-0.131	[.0001]	0.017	[0.0003]	0.003	[.6093]	**-1.027**	[.0001]	0.010	[.0141]	0.006	[.5079]
	S&Y	-0.153	[.0001]	0.000	[0.9949]	-0.010	[.2407]	**-1.033**	[.0001]	0.023	[.0013]	0.019	[.0728]
	Dong et al (actual)	-0.083	[.0001]	0.015	[0.0011]	0.015	[.0213]	**-1.057**	[.0001]	0.042	[.0137]	-0.045	[.0001]
	Dong et al (latent)	-0.120	[.0001]	0.032	[0.0001]	0.038	[.0001]	**-1.089**	[.0001]	0.093	[.0001]	-0.110	[.0001]
	GME (unrestricted)	-0.123		0.016		0.033		**-0.999**		0.047		0.044	
	Mean	**-0.124**		**0.013**		**0.016**		**-1.040**		**0.040**		**-0.016**	
	Std. Deviation	**0.024**		**0.012**		**0.018**		**0.030**		**0.029**		**0.055**	
Bottled Water	ITSUR	-0.066	[.0754]	0.081	[.0001]	0.212	[.0001]	0.308	[.0001]	**-2.013**	[.0001]	0.545	[.0001]
	H&W	-0.043	[.2050]	-0.033	[.0663]	0.168	[.0001]	0.235	[.0001]	**-1.556**	[.0001]	0.478	[.0001]
	S&Y	-0.150	[0.0051	0.308	[.0001]	0.465	[.0001]	0.334	[.0001]	**-2.576**	[.0001]	0.696	[.0001]
	Dong et al (actual)	-0.058	[.0076]	0.093	[.0100]	0.155	[.0001]	0.126	[.0001]	**-1.850**	[.0137]	0.407	[.0001]
	Dong et al (latent)	-0.191	[.0029]	0.305	[.0100]	0.498	[.0001]	0.351	[.0001]	**-3.501**	[.0137]	1.142	[.0001]
	GME (unrestricted)	0.325		0.180		0.284		0.463		**-1.944**		0.577	
	Mean	**-0.031**		**0.156**		**0.297**		**0.303**		**-2.240**		**0.641**	
	Std. Deviation	**0.184**		**0.135**		**0.150**		**0.114**		**0.702**		**0.264**	
Milk	ITSUR	0.111	[.0001]	0.026	[.0001]	0.117	[.0001]	0.070	[.0001]	0.076	[.0001]	**-1.274**	[.0001]
	H&W	0.108	[.0001]	0.050	[.0001]	0.099	[.0001]	0.090	[.0001]	0.065	[.0001]	**-1.258**	[.0001]
	S&Y	0.124	[.0001]	-0.008	[.2041]	0.125	[.0001]	0.079	[.0001]	0.101	[.0001]	**-1.285**	[.0001]
	Dong et al (actual)	0.084	[.0001]	0.011	[.0156]	0.101	[.0001]	0.085	[.0001]	0.056	[.0001]	**-1.200**	[.0001]
	Dong et al (latent)	0.125	[.0001]	0.032	[.0001]	0.184	[.0001]	0.123	[.0001]	0.125	[.0001]	**-1.379**	[.0001]
	GME (unrestricted)	0.004		-0.011		0.056		-0.039		0.070		**-1.456**	
	Mean	**0.093**		**0.017**		**0.114**		**0.068**		**0.082**		**-1.309**	
	Std. Deviation	**0.046**		**0.024**		**0.042**		**0.055**		**0.026**		**0.092**	

Note: p-values are in brackets

[1] Calculated using sample means.

Table 5.14. Compensated Own- and Cross-Price Elasticity Matrix[1] of Non-Alcoholic Beverages Using the AIDS and the 1999 ACNielsen Homescan Data

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Fruit Juice	ITSUR	**-0.827**	[.0001]	0.050	[.0001]	0.193	[.0001]	0.139	[.0001]	0.020	[.0108]	0.425	[.0001]
	H&W	**-0.777**	[.0001]	0.040	[.0001]	0.173	[.0001]	0.139	[.0001]	0.018	[.0149]	0.407	[.0001]
	S&Y	**-0.812**	[.0001]	0.022	[.0245]	0.231	[.0001]	0.114	[.0001]	0.001	[.9528]	0.445	[.0001]
	Dong et al (actual)	**-0.877**	[.0001]	0.064	[0.0001]	0.189	[.0001]	0.202	[.0001]	-0.006	[.1923]	0.428	[.0001]
	Dong et al (latent)	**-0.913**		0.091		0.265		0.065		-0.006		0.498	
	GME (unrestricted)	**-0.730**		0.057		0.255		0.257		-0.142		0.474	
	Mean	**-0.823**		**0.054**		**0.218**		**0.153**		**-0.019**		**0.446**	
	Std. Deviation	**0.066**		**0.023**		**0.038**		**0.068**		**0.061**		**0.034**	
Tea	ITSUR	0.199	[.0001]	**-1.244**	[.0001]	0.153	[.0001]	0.399	[.0001]	0.103	[.0001]	0.390	[.0001]
	H&W	0.115	[.0001]	**-1.224**	[.0001]	0.011	[.5905]	0.530	[.0001]	-0.016	[.2609]	0.585	[.0001]
	S&Y	0.101	[.0073]	**-1.496**	[.0001]	0.587	[.0001]	0.373	[.0001]	0.296	[.0001]	0.139	[.0001]
	Dong et al (actual)	0.190	[.0001]	**-1.256**	[0.0001]	0.147	[.0001]	0.425	[.0001]	0.051	[.0001]	0.442	[.0001]
	Dong et al (latent)	0.210		**-1.725**		0.216		0.519		0.135		0.645	
	GME (unrestricted)	0.361		**-1.207**		0.206		0.257		-0.142		0.474	
	Mean	**0.196**		**-1.359**		**0.220**		**0.417**		**0.071**		**0.446**	
	Std. Deviation	**0.093**		**0.209**		**0.194**		**0.101**		**0.148**		**0.177**	
Coffee	ITSUR	0.310	[0.0001]	0.066	[.0001]	**-1.483**	[.0001]	0.437	[.0001]	0.109	[.0001]	0.560	[.0001]
	H&W	0.284	[0.0001]	0.008	[.3918]	**-1.270**	[.0001]	0.393	[.0001]	0.090	[.0001]	0.495	[.0001]
	S&Y	0.376	[0.0001]	0.231	[.0001]	**-1.764**	[.0001]	0.442	[.0001]	0.193	[.0001]	0.522	[.0001]
	Dong et al (actual)	0.289	[.0001]	0.061	[.0001]	**-1.337**	[.0001]	0.437	[.0001]	0.092	[.0001]	0.459	[.0001]
	Dong et al (latent)	0.409		0.090		**-1.785**		0.500		0.192		0.594	
	GME (unrestricted)	0.189		0.050		**-1.522**		0.343		0.081		0.462	
	Mean	**0.310**		**0.084**		**-1.527**		**0.425**		**0.126**		**0.515**	
	Std. Deviation	**0.077**		**0.077**		**0.213**		**0.053**		**0.052**		**0.054**	

Table 5.14 Continued

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Carbonated Soft drinks	ITSUR	0.064	[.0001]	0.054	[.0001]	0.139	[.0001]	**-0.645**	[.0001]	0.068	[.0001]	0.320	[.0001]
	H&W	0.066	[.0001]	0.069	[.0001]	0.125	[.0001]	**-0.642**	[.0001]	0.053	[.0001]	0.329	[.0001]
	S&Y	0.049	[.0001]	0.054	[.0001]	0.115	[.0001]	**-0.637**	[.0001]	0.067	[.0001]	0.352	[.0001]
	Dong et al (actual)	0.112	[.0001]	0.067	[.0001]	0.137	[.0001]	**-0.676**	[.0001]	0.084	[.0001]	0.276	[.0001]
	Dong et al (latent)	0.096		0.093		0.181	[.0001]	**-0.708**		0.132		0.207	
	GME (unrestricted)	0.080		0.071		0.160		**-0.603**		0.091		0.377	
	Mean	**0.078**		**0.068**		**0.143**		**-0.652**		**0.083**		**0.310**	
	Std. Deviation	**0.023**		**0.014**		**0.024**		**0.036**		**0.028**		**0.061**	
Bottled Water	ITSUR	0.097	[.0089]	0.125	[.0001]	0.314	[.0001]	0.628	[.0001]	**-1.977**	[.0001]	0.814	[.0001]
	H&W	0.088	[.0090]	0.002	[.8978]	0.250	[.0001]	0.493	[.0001]	**-1.527**	[.0001]	0.694	[.0001]
	S&Y	0.011	[.8326]	0.351	[.0001]	0.566	[.0001]	0.651	[.0001]	**-2.541**	[.0001]	0.962	[.0001]
	Dong et al (actual)	0.139	[.0001]	0.146	[.0001]	0.278	[.0001]	0.512	[.0001]	**-1.807**	[.0001]	0.732	[.0001]
	Dong et al (latent)	0.068		0.380		0.670		0.811		**-3.455**		1.525	
	GME (unrestricted)	0.492		0.225		0.389		0.791		**-1.908**		0.853	
	Mean	**0.149**		**0.205**		**0.411**		**0.648**		**-2.203**		**0.930**	
	Std. Deviation	**0.173**		**0.144**		**0.170**		**0.134**		**0.698**		**0.307**	
Milk	ITSUR	0.264	[.0001]	0.067	[.0001]	0.213	[.0001]	0.370	[.0001]	0.109	[.0001]	**-1.023**	[.0001]
	H&W	0.256	[.0001]	0.090	[.0001]	0.192	[.0001]	0.380	[.0001]	0.097	[.0001]	**-1.014**	[.0001]
	S&Y	0.275	[.0001]	0.032	[.0001]	0.219	[.0001]	0.375	[.0001]	0.134	[.0001]	**-1.036**	[.0001]
	Dong et al (actual)	0.235	[.0001]	0.052	[.0001]	0.195	[.0001]	0.381	[.0001]	0.088	[.0001]	**-0.951**	[.0001]
	Dong et al (latent)	0.272		0.074		0.281		0.383		0.152		**-1.162**	
	GME (unrestricted)	0.152		0.040		0.148		0.251		0.102		**-1.211**	
	Mean	**0.242**		**0.059**		**0.208**		**0.357**		**0.114**		**-1.066**	
	Std. Deviation	**0.046**		**0.022**		**0.044**		**0.052**		**0.024**		**0.099**	

Note: p-values are in brackets

[1] Calculated using sample means.

Table 5.15. Expenditure Elasticities[1] of Non-Alcoholic Beverages Using the QUAIDS System and 1999 ACNielsen Homescan Data

Item	ITSUR Estimate	H&W Estimate	S&Y Estimate	GME Estimate	Mean	Standard Deviation
Fruit Juice	0.982	0.932	0.964	1.010	0.972	0.033
	(0.000)	(0.000)	(0.000)			
Tea	0.767	1.601	0.841	0.776	0.996	0.404
	(0.000)	(0.000)	(0.000)			
Coffee	0.879	0.757	0.844	0.872	0.838	0.056
	(0.000)	(0.000)	(0.000)			
Carbonated Soft drinks	1.184	1.171	1.189	1.201	1.186	0.012
	(0.000)	(0.000)	(0.000)			
Bottled Water	1.033	0.828	1.127	1.054	1.011	0.128
	(0.000)	(0.000)	(0.000)			
Milk	0.870	0.855	0.864	0.833	0.856	0.016
	(0.000)	(0.000)	(0.000)			

Note: p-values are in brackets

[1] Calculated using sample means

Table 5.16. Uncompensated Own- and Cross-Price Elasticity Matrix[1] of Non-Alcoholic Beverages Using the QUAIDS and the 1999 ACNielsen Homescan Data

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Fruit Juice	ITSUR	**-0.998**	[.0001]	0.004	[0.6352]	0.084	[.0001]	-0.197	[.0001]	-0.017	[.0344]	0.143	[.0001]
	H&W	**-0.939**	[.0001]	-0.004	[0.5953]	0.070	[.0001]	-0.181	[.0001]	-0.018	[.0137]	0.139	[.0001]
	S&Y	**-0.974**	[.0001]	-0.033	[0.0011]	0.142	[.0001]	-0.214	[.0001]	-0.046	[.0001]	0.160	[.0001]
	GME (unrestricted)	**-0.892**		-0.004		0.160		-0.095		-0.210		0.177	
	Mean	**-0.951**		**-0.009**		**0.114**		**-0.172**		**-0.073**		**0.155**	
	Std. Deviation	**0.046**		**0.016**		**0.044**		**0.053**		**0.093**		**0.018**	
Tea	ITSUR	0.063	[.0454]	**-1.279**	[.0001]	0.070	[.0019]	0.136	[.0002]	0.074	[.0001]	0.170	[.0001]
	H&W	-0.165	[.0070]	**-1.303**	[.0001]	-0.177	[.0001]	-0.005	[.8915]	-0.072	[.0001]	0.120	[.0012]
	S&Y	-0.080	[.0747]	**-1.462**	[.0001]	0.389	[.0001]	0.076	[.1515]	0.289	[.0001]	-0.053	[.2629]
	GME (unrestricted)	0.201		**-1.236**		0.090		0.216		0.048		0.514	
	Mean	**0.005**		**-1.320**		**0.093**		**0.106**		**0.085**		**0.188**	
	Std. Deviation	**0.161**		**0.098**		**0.231**		**0.094**		**0.150**		**0.237**	
Coffee	ITSUR	0.159	[0.0001]	0.025	[0.0059]	**-1.586**	[0.0001]	0.140	[0.0001]	0.078	[.0001]	0.305	[.0001]
	H&W	0.154	[0.0001]	-0.0299	[0.0009]	**-1.357**	[0.0001]	0.137	[0.0001]	0.062	[.0001]	0.277	[.0001]
	S&Y	0.248	[0.0064]	0.143	[0.0001]	**-1.792**	[0.0001]	0.198	[0.0913]	0.132	[.0001]	0.228	[.0001]
	GME (unrestricted)	0.076		-0.002		**-1.579**		0.037		0.034		0.201	
	Mean	**0.159**		**0.034**		**-1.579**		**0.128**		**0.077**		**0.253**	
	Std. Deviation	**0.070**		**0.076**		**0.178**		**0.067**		**0.041**		**0.047**	
Carbonated Soft drinks	ITSUR	-0.145	[.0001]	-0.002	[0.6908]	0.010	[.1178]	**-1.051**	[.0001]	0.023	[.0001]	-0.020	[.0446]
	H&W	-0.141	[.0001]	0.014	[0.0019]	-0.001	[.8488]	**-1.044**	[.0001]	0.007	[.0642]	-0.007	[.4951]
	S&Y	-0.166	[.0001]	0.001	[0.8270]	-0.027	[.0023]	**-1.046**	[.0001]	0.039	[.0001]	0.009	[.4027]
	GME (unrestricted)	-0.143		0.017		0.021		**-1.022**		0.048		0.037	
	Mean	**-0.149**		**0.008**		**0.001**		**-1.041**		**0.029**		**0.005**	
	Std. Deviation	**0.012**		**0.009**		**0.020**		**0.013**		**0.018**		**0.025**	

Table 5.16 Continued

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Bottled Water	ITSUR	-0.089	[0.0171]	0.077	[.0001]	0.204	[.0001]	0.274	[.0001]	**-2.015**	[.0001]	0.517	[.0001]
	H&W	-0.062	[0.06203]	-0.035	[.0510]	0.159	[.0001]	0.210	[.0001]	**-1.556**	[.0001]	0.457	[.0001]
	S&Y	-0.230	[0.5354]	0.375	[.0835]	0.372	[.0001]	0.144	[.0290]	**-2.539**	[.0001]	0.751	[.0001]
	GME (unrestricted)	0.267		0.187		0.232		0.457		**-1.932**		0.558	
	Mean	**-0.029**		**0.151**		**0.242**		**0.271**		**-2.011**		**0.571**	
	Std. Deviation	**0.211**		**0.175**		**0.092**		**0.135**		**0.405**		**0.127**	
Milk	ITSUR	0.114	[.0001]	0.026	[.0001]	0.120	[.0001]	0.069	[.0001]	0.075	[.0001]	**-1.274**	[.0001]
	H&W	0.107	[.0001]	0.051	[.0001]	0.102	[.0001]	0.083	[.0001]	0.064	[.0001]	**-1.261**	[.0001]
	S&Y	0.131	[.0001]	-0.010	[.1081]	0.133	[.0001]	0.079	[.0001]	0.088	[.0001]	**-1.285**	[.0001]
	GME (unrestricted)	0.007		-0.010		0.053		-0.024		0.071		**-1.455**	
	Mean	**0.090**		**0.014**		**0.102**		**0.052**		**0.074**		**-1.319**	
	Std. Deviation	**0.056**		**0.030**		**0.035**		**0.051**		**0.010**		**0.091**	

Note: p-values are in brackets

[1] Calculated using sample means.

Table 5.17. Compensated Own- and Cross-Price Elasticity Matrix[1] of Non-Alcoholic Beverages Using the QUAIDS and the 1999 ACNielsen Homescan Data

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Fruit Juice	ITSUR	**-0.826**	[.0001]	0.050	[.0001]	0.191	[.0001]	0.140	[.0001]	0.020	[.0108]	0.426	[.0001]
	H&W	**-0.776**	[.0001]	0.040	[.0001]	0.172	[.0001]	0.139	[.0001]	0.018	[.0214]	0.408	[.0001]
	S&Y	**-0.805**	[.0001]	0.013	[.2032]	0.247	[.0001]	0.117	[.0001]	-0.009	[.4151]	0.438	[.0001]
	GME (unrestricted)	**-0.716**		0.043		0.270		0.251		-0.172		0.469	
	Mean	**-0.781**		**0.036**		**0.220**		**0.162**		**-0.036**		**0.435**	
	Std. Deviation	**0.048**		**0.016**		**0.046**		**0.060**		**0.092**		**0.026**	
Tea	ITSUR	0.197	[.0001]	**-1.243**	[.0001]	0.154	[.0001]	0.399	[.0001]	0.103	[.0001]	0.391	[.0001]
	H&W	0.115	[.0001]	**-1.228**	[.0001]	-0.002	[.9184]	0.544	[.0001]	-0.011	[.4564]	0.581	[.0001]
	S&Y	0.067	[.0772]	**-1.422**	[.0001]	0.480	[.0001]	0.365	[.0001]	0.321	[.0001]	0.189	[.0001]
	GME (unrestricted)	0.337		**-1.199**		0.175		0.482		0.078		0.737	
	Mean	**0.179**		**-1.273**		**0.202**		**0.447**		**0.123**		**0.475**	
	Std. Deviation	**0.118**		**0.101**		**0.202**		**0.081**		**0.141**		**0.237**	
Coffee	ITSUR	0.313	[0.0001]	0.066	[.0001]	**-1.490**	[.0001]	0.442	[.0001]	0.111	[.0001]	0.558	[.0001]
	H&W	0.286	[0.0001]	0.006	[.5303]	**-1.275**	[.0001]	0.397	[.0001]	0.091	[.0001]	0.495	[.0001]
	S&Y	0.396	[0.0001]	0.182	[.0001]	**-1.700**	[.0001]	0.487	[.0001]	0.164	[.0001]	0.471	[.0001]
	GME (unrestricted)	0.228		0.039		**-1.484**		0.336		0.068		0.452	
	Mean	**0.306**		**0.073**		**-1.487**		**0.415**		**0.108**		**0.494**	
	Std. Deviation	**0.070**		**0.077**		**0.174**		**0.065**		**0.041**		**0.046**	
Carbonated Soft drinks	ITSUR	0.062	[.0001]	0.054	[.0001]	0.139	[.0001]	**-0.644**	[.0001]	0.068	[.0001]	0.321	[.0001]
	H&W	0.064	[.0001]	0.069	[.0001]	0.126	[.0001]	**-0.642**	[.0001]	0.052	[.0001]	0.331	[.0001]
	S&Y	0.042	[.0001]	0.057	[.0001]	0.103	[.0001]	**-0.638**	[.0001]	0.084	[.0001]	0.352	[.0001]
	GME (unrestricted)	0.067		0.073		0.152		**-0.611**		0.094		0.384	
	Mean	**0.059**		**0.064**		**0.130**		**-0.634**		**0.075**		**0.347**	
	Std. Deviation	**0.011**		**0.009**		**0.021**		**0.016**		**0.019**		**0.028**	

Table 5.17 Continued

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Bottled Water	ITSUR	0.092	[.0693]	0.125	[.0001]	0.317	[.0001]	0.628	[.0001]	**-1.976**	[.0001]	0.814	[.0001]
	H&W	0.083	[.0140]	0.004	[.8310]	0.249	[.0001]	0.494	[.0001]	**-1.525**	[.0001]	0.695	[.0001]
	S&Y	-0.033	[.5349]	0.428	[.0001]	0.495	[.0001]	0.530	[.0001]	**-2.496**	[.0001]	1.076	[.0001]
	GME (unrestricted)	0.451		0.236		0.347		0.818		**-1.892**		0.862	
	Mean	**0.148**		**0.198**		**0.352**		**0.618**		**-1.972**		**0.862**	
	Std. Deviation	**0.210**		**0.180**		**0.104**		**0.145**		**0.400**		**0.159**	
Milk	ITSUR	0.266	[.0001]	0.067	[.0001]	0.215	[.0001]	0.367	[.0001]	0.108	[.0001]	**-1.024**	[.0001]
	H&W	0.257	[.0001]	0.091	[.0001]	0.195	[.0001]	0.377	[.0001]	0.096	[.0001]	**-1.015**	[.0001]
	S&Y	0.283	[.0001]	0.031	[.0001]	0.227	[.0001]	0.375	[.0001]	0.120	[.0001]	**-1.036**	[.0001]
	GME (unrestricted)	0.153		0.039		0.145		0.261		0.103		**-1.215**	
	Mean	**0.240**		**0.057**		**0.195**		**0.345**		**0.107**		**-1.072**	
	Std. Deviation	**0.059**		**0.027**		**0.036**		**0.056**		**0.010**		**0.095**	

Note: p-values are in brackets

[1] Calculated using sample means.

Table 5.18. Expenditure Elasticities[1] of Non-Alcoholic Beverages Using the AIDS System and 1999 ACNielsen Homesan Data (Unrestricted)

Item	ITSUR Estimate	H&W Estimate	S&Y Estimate	GME Estimate	Mean	Standard Deviation
Fruit Juice	1.039	0.976	1.040	1.042	1.024	0.032
	(0.000)	(0.000)	(0.000)			
Tea	0.745	1.770	0.715	0.741	0.993	0.519
	(0.000)	(0.000)	(0.000)			
Coffee	0.976	0.841	0.965	0.968	0.937	0.065
	(0.000)	(0.000)	(0.000)			
Carbonated Soft Drinks	1.155	1.135	1.171	1.158	1.155	0.015
	(0.000)	(0.000)	(0.000)			
Bottled Water	0.963	0.762	0.963	0.958	0.911	0.100
	(0.000)	(0.000)	(0.000)			
Milk	0.847	0.820	0.836	0.847	0.838	0.013
	(0.000)	(0.000)	(0.000)			

p-values are in parenthesis

[1] Calculated using sample means

Table 5.19. Uncompensated Own- and Cross-Price Elasticity Matrix[1] of Non-Alcoholic Beverages Using the AIDS and 1999 ACNielsen Homescan Data (Unrestricted)

		Fruit Juice		Tea		Coffee		Carbonated Soft drinks		Bottled Water		Milk	
Fruit Juice	ITSUR	**-0.905**	[.0001]	0.010	[0.3170]	0.145	[.0001]	-0.093	[.0001]	-0.038	[.0035]	0.301	[.0001]
	H&W	**-0.853**	[.0001]	0.016	[0.1106]	0.133	[.0001]	-0.084	[.0001]	-0.034	[.0136]	0.258	[.0001]
	S&Y	**-0.872**	[.0001]	0.000	[0.9705]	0.170	[.0001]	-0.086	[.0022]	-0.059	[.0004]	0.384	[.0001]
	GME	**-0.912**		0.009		0.141		-0.100		-0.182		0.173	
	Mean	**-0.885**		**0.009**		**0.147**		**-0.091**		**-0.078**		**0.279**	
	Std. Deviation	**0.028**		**0.007**		**0.016**		**0.007**		**0.070**		**0.088**	
Tea	ITSUR	0.197	[.0001]	**-1.254**	[.1828]	0.096	[.0004]	0.159	[.0001]	0.036	[.0001]	0.417	[.0001]
	H&W	-0.018	[.0586]	**-1.431**	[.3909]	-0.277	[.0001]	0.019	[.1877]	-0.306	[.0001]	0.991	[.0001]
	S&Y	-0.024	[.0123]	**-1.376**	[.1545]	0.775	[.0001]	0.014	[.2010]	0.416	[.0001]	-0.345	[.0001]
	GME	0.231		**-1.242**		0.124		0.205		0.038		0.524	
	Mean	**0.097**		**-1.325**		**0.179**		**0.099**		**0.046**		**0.397**	
	Std. Deviation	**0.136**		**0.093**		**0.437**		**0.098**		**0.295**		**0.554**	
Coffee	ITSUR	0.014	[0.6441]	0.003	[0.8206]	**-1.632**	[0.0001]	0.003	[0.9226]	0.044	[.0092]	0.172	[.0001]
	H&W	0.018	[0.4823]	0.053	[0.0001]	**-1.415**	[0.0001]	0.032	[0.2068]	0.034	[.0192]	0.092	[.0025]
	S&Y	0.018	[0.7216]	0.0164	[0.4725]	**-2.035**	[0.0001]	-0.019	[0.7024]	0.089	[.0022]	0.297	[.0001]
	GME	0.020		0.004		**-1.628**		0.011		0.045		0.183	
	Mean	**0.017**		**0.019**		**-1.678**		**0.007**		**0.053**		**0.186**	
	Std. Deviation	**0.003**		**0.023**		**0.259**		**0.021**		**0.024**		**0.084**	

Table 5.19 Continued

		Fruit Juice		Tea		Coffee		Carbonated Soft drinks		Bottled Water		Milk	
Carbonated Soft drinks	ITSUR	-0.096	[.0001]	0.0218	[0.0025]	0.053	[.0001]	**-0.968**	[.0001]	0.053	[.0001]	0.103	[.0001]
	H&W	-0.090	[.0001]	-0.004	[0.6689]	0.058	[.0001]	**-0.982**	[.0001]	0.056	[.0001]	0.058	[.0001]
	S&Y	-0.103	[.0001]	0.027	[0.0007]	0.048	[.0001]	**-0.954**	[.0001]	0.051	[.0001]	0.112	[.0001]
	GME	-0.123		0.016		0.033		**-0.999**		0.047		0.044	
	Mean	**-0.103**		**0.015**		**0.048**		**-0.976**		**0.052**		**0.079**	
	Std. Deviation	**0.015**		**0.013**		**0.011**		**0.019**		**0.004**		**0.033**	
Bottled Water	ITSUR	0.317	[0.0001]	0.178	[.0001]	0.279	[.0001]	0.458	[.0001]	**-1.947**	[.0001]	0.560	[.0001]
	H&W	0.239	[0.0001]	0.177	[.0001]	0.163	[.0001]	0.344	[.0001]	**-1.503**	[.0001]	0.351	[.0001]
	S&Y	0.507	[0.0001]	0.302	[.0001]	0.471	[.0001]	0.686	[.0001]	**-2.497**	[.0001]	0.922	[.0001]
	GME	0.325		0.180		0.284		0.463		**-1.944**		0.577	
	Mean	**0.347**		**0.209**		**0.299**		**0.488**		**-1.973**		**0.602**	
	Std. Deviation	**0.113**		**0.062**		**0.127**		**0.143**		**0.407**		**0.236**	
Milk	ITSUR	-0.022	[.1566]	-0.016	[.0234]	0.037	[.0003]	-0.069	[.0001]	0.063	[.0001]	**-1.513**	[.0001]
	H&W	-0.018	[.2153]	0.022	[.0156]	0.031	[.0019]	-0.031	[.0390]	0.057	[.0001]	**-1.545**	[.0001]
	S&Y	-0.024	[.1348]	-0.016	[.0341]	0.044	[.0001]	-0.088	[.0001]	0.070	[.0001]	**-1.544**	[.0001]
	GME	0.004		-0.011		0.056		-0.039		0.070		**-1.456**	
	Mean	**-0.015**		**-0.005**		**0.042**		**-0.057**		**0.065**		**-1.515**	
	Std. Deviation	**0.013**		**0.018**		**0.011**		**0.026**		**0.006**		**0.042**	

Note: p-values are in brackets

[1] Calculated using sample means

Table 5.20. Compensated Own- and Cross-Price Elasticity Matrix[1] of Non-Alcoholic Beverages Using the AIDS and 1999 ACNielsen Homescan Data (Unrestricted)

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Fruit Juice	ITSUR	**-0.723**	[.0001]	0.059	[.0001]	0.259	[.0001]	0.264	[.0001]	0.002	[.9024]	0.600	[.0001]
	H&W	**-0.682**	[.0001]	0.061	[.0001]	0.239	[.0001]	0.251	[.0001]	0.003	[.7842]	0.539	[.0001]
	S&Y	**-0.690**	[.0001]	0.048	[.0002]	0.283	[.0001]	0.271	[.0001]	-0.019	[.2473]	0.683	[.0001]
	GME	**-0.730**		0.057		0.255		0.257		-0.142		0.474	
	Mean	**-0.706**		**0.057**		**0.259**		**0.261**		**-0.039**		**0.574**	
	Std. Deviation	**0.024**		**0.006**		**0.018**		**0.009**		**0.069**		**0.089**	
Tea	ITSUR	0.327	[.0001]	**-1.219**	[.1954]	0.177	[.0001]	0.415	[.0001]	0.065	[.0001]	0.631	[.0001]
	H&W	0.292	[.0001]	**-1.347**	[.4191]	-0.084	[.0001]	0.626	[.0001]	-0.239	[.0001]	1.501	[.0001]
	S&Y	0.102	[.0001]	**-1.342**	[.1649]	0.853	[.0001]	0.259	[.0001]	0.443	[.0001]	-0.139	[.0001]
	GME	0.361		**-1.207**		0.206		0.257		-0.142		0.474	
	Mean	**0.270**		**-1.279**		**0.288**		**0.389**		**0.032**		**0.617**	
	Std. Deviation	**0.116**		**0.076**		**0.398**		**0.174**		**0.302**		**0.677**	
Coffee	ITSUR	0.185	[0.0001]	0.049	[.0003]	**-1.526**	[.0001]	0.338	[.0001]	0.081	[.0001]	0.045	[.0001]
	H&W	0.165	[0.0001]	0.092	[.0001]	**-1.324**	[.0001]	0.320	[.0001]	0.066	[.0001]	0.335	[.0001]
	S&Y	0.187	[0.0002]	0.062	[.0001]	**-1.930**	[.0001]	0.312	[.0001]	0.125	[.0001]	0.575	[.0001]
	GME	0.189		0.050		**-1.522**		0.343		0.081		0.462	
	Mean	**0.181**		**0.063**		**-1.575**		**0.328**		**0.088**		**0.354**	
	Std. Deviation	**0.011**		**0.020**		**0.255**		**0.014**		**0.026**		**0.228**	

Table 5.20 Continued

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Carbonated Soft drinks	ITSUR	0.107	[.0001]	0.076	[.0001]	0.179	[.0001]	**-0.572**	[.0001]	0.097	[.0001]	0.436	[.0001]
	H&W	0.109	[.0001]	0.050	[.0001]	0.182	[.0001]	**-0.593**	[.0001]	0.099	[.0001]	0.449	[.0001]
	S&Y	0.102	[.0001]	0.082	[.0001]	0.175	[.0001]	**-0.552**	[.0001]	0.096	[.0001]	0.450	[.0001]
	GME	0.080		0.071		0.160		**-0.603**		0.091		0.377	
	Mean	**0.099**		**0.070**		**0.174**		**-0.580**		**0.096**		**0.428**	
	Std. Deviation	**0.013**		**0.014**		**0.010**		**0.023**		**0.004**		**0.034**	
Bottled Water	ITSUR	0.486	[.0001]	0.223	[.0001]	0.384	[.0001]	0.789	[.0001]	**-1.910**	[.0001]	0.838	[.0001]
	H&W	0.372	[.0001]	0.213	[.0001]	0.246	[.0001]	0.605	[.0001]	**-1.474**	[.0001]	0.570	[.0001]
	S&Y	0.675	[.0001]	0.347	[.0001]	0.576	[.0001]	0.651	[.0001]	**-2.461**	[.0001]	1.016	[.0001]
	GME	0.492		0.225		0.389		0.791		**-1.908**		0.853	
	Mean	**0.506**		**0.252**		**0.399**		**0.709**		**-1.938**		**0.819**	
	Std. Deviation	**0.125**		**0.064**		**0.135**		**0.095**		**0.404**		**0.185**	
Milk	ITSUR	0.127	[.0001]	0.024	[.0009]	0.129	[.0001]	0.222	[.0001]	0.096	[.0001]	**-1.269**	[.0001]
	H&W	0.125	[.0001]	0.060	[.0001]	0.120	[.0001]	0.250	[.0001]	0.088	[.0001]	**-1.309**	[.0001]
	S&Y	0.122	[.0001]	0.023	[.0001]	0.135	[.0001]	0.199	[.0001]	0.102	[.0001]	**-1.303**	[.0001]
	GME	0.152		0.040		0.148		0.251		0.102		**-1.211**	
	Mean	**0.131**		**0.037**		**0.133**		**0.230**		**0.097**		**-1.273**	
	Std. Deviation	**0.014**		**0.018**		**0.012**		**0.025**		**0.007**		**0.045**	

Note: p-values are in brackets

[1]Calculated using sample means

Table 5.21. Expenditure Elasticities[1] of Non-Alcoholic Beverages Using the QUAIDS System and 1999 ACNielsen Homescan Data (Unrestricted)

Item	ITSUR Estimate	H&W Estimate	S&Y Estimate	GME Estimate	Mean	Standard Deviation
Fruit Juice	1.054	0.956	1.079	1.010	1.025	0.054
	(0.000)	(0.000)	(0.000)			
Tea	0.586	1.547	0.929	0.776	0.959	0.416
	(0.000)	(0.000)	(0.000)			
Coffee	0.988	0.734	0.661	0.872	0.814	0.145
	(0.000)	(0.000)	(0.000)			
Carbonated Soft Drinks	1.162	1.198	1.199	1.201	1.190	0.019
	(0.000)	(0.000)	(0.000)			
Bottled Water	0.943	0.862	0.995	1.054	0.963	0.081
	(0.000)	(0.000)	(0.000)			
Milk	0.854	0.820	0.856	0.833	0.841	0.017
	(0.000)	(0.000)	(0.000)			

Note: p-values are in brackets

[1] Calculated using sample means

Table 5.22. Uncompensated Own- and Cross-Price Elasticity Matrix[1] of Non-Alcoholic Beverages Using the QUAIDS and the 1999 ACNielsen Homescan Data (Unrestricted)

		Fruit Juice		Tea		Coffee		Carbonated Soft drinks		Bottled Water		Milk	
Fruit Juice	ITSUR	**-0.907**	[.0001]	0.051	[0.0001]	0.143	[.0001]	-0.094	[.0001]	-0.037	[.0045]	0.297	[.0001]
	H&W	**-0.850**	[.0001]	-0.028	[0.1028]	0.134	[.0001]	-0.081	[.0001]	-0.034	[.0049]	0.264	[.0001]
	S&Y	**-0.887**	[.0001]	0.021	[0.1310]	0.384	[.0001]	-0.095	[.0007]	-0.038	[.0232]	0.377	[.0001]
	GME	**-0.892**		-0.004		0.160		-0.095		-0.210		0.177	
	Mean	**-0.884**		**0.010**		**0.205**		**-0.091**		**-0.080**		**0.279**	
	Std. Deviation	**0.024**		**0.034**		**0.120**		**0.007**		**0.087**		**0.083**	
Tea	ITSUR	0.210	[.0001]	**-1.686**	[.0001]	0.115	[.0001]	0.160	[.0001]	0.014	[.0001]	0.453	[.0001]
	H&W	0.006	[.7797]	**-1.737**	[.0001]	-0.287	[.0001]	0.140	[.8915]	-0.312	[.0001]	0.978	[.0001]
	S&Y	-0.074	[.0001]	**-1.427**	[.0001]	1.984	[.0001]	0.071	[.0001]	0.520	[.0001]	-0.451	[.0001]
	GME	0.201		**-1.236**		0.090		0.216		0.048		0.514	
	Mean	**0.086**		**-1.521**		**0.475**		**0.147**		**0.067**		**0.373**	
	Std. Deviation	**0.142**		**0.234**		**1.023**		**0.060**		**0.343**		**0.598**	
Coffee	ITSUR	0.012	[0.6903]	0.034	[0.0413]	**-1.634**	[0.0001]	0.003	[0.9165]	0.045	[.0076]	0.170	[.0001]
	H&W	0.094	[0.0004]	0.003	[0.7935]	**-1.410**	[0.0001]	0.052	[0.0386]	0.042	[.0042]	0.139	[.0001]
	S&Y	0.124	[0.0147]	-0.083	[0.0258]	**-3.800**	[0.0001]	0.0003	[0.9951]	-0.067	[.0897]	0.390	[.0001]
	GME	0.076		-0.002		**-1.579**		0.037		0.034		0.201	
	Mean	**0.077**		**-0.012**		**-2.106**		**0.023**		**0.014**		**0.225**	
	Std. Deviation	**0.048**		**0.050**		**1.133**		**0.025**		**0.054**		**0.113**	
Carbonated Soft drinks	ITSUR	-0.097	[.0001]	0.041	[0.0001]	0.052	[.0001]	**-0.970**	[.0001]	0.054	[.0001]	0.103	[.0001]
	H&W	-0.136	[.0001]	0.035	[0.0001]	0.053	[.0001]	**-0.989**	[.0001]	0.051	[.0001]	0.090	[.001]
	S&Y	-0.116	[.0001]	0.051	[0.0001]	0.747	[.0001]	**-0.972**	[.0001]	0.066	[.0001]	0.116	[.0001]
	GME	-0.143		0.017		0.021		**-1.022**		0.048		0.037	
	Mean	**-0.123**		**0.036**		**0.218**		**-0.988**		**0.055**		**0.087**	
	Std. Deviation	**0.021**		**0.014**		**0.353**		**0.024**		**0.008**		**0.034**	

Table 5.22 Continued

		Fruit Juice		Tea		Coffee		Carbonated Soft drinks		Bottled Water		Milk	
Bottled Water	ITSUR	0.321	[0.0001]	0.123	[.0001]	0.282	[.0001]	0.459	[.0001]	**-1.948**	[.0001]	0.565	[.0001]
	H&W	0.227	[0.0001]	0.318	[.0001]	0.167	[.0001]	0.297	[.0001]	**-1.499**	[.0001]	0.354	[.0001]
	S&Y	0.492	[0.0001]	0.307	[.0001]	0.711	[.0001]	0.688	[.0001]	**-2.478**	[.0001]	0.893	[.0001]
	GME	0.267		0.187		0.232		0.457		**-1.932**		0.558	
	Mean	**0.327**		**0.234**		**0.348**		**0.475**		**-1.965**		**0.593**	
	Std. Deviation	**0.116**		**0.095**		**0.247**		**0.161**		**0.401**		**0.223**	
Milk	ITSUR	-0.022	[.1551]	0.003	[.7189]	0.035	[.0004]	-0.066	[.0001]	0.064	[.0001]	**-1.516**	[.0001]
	H&W	-0.016	[.2688]	-0.017	[.1292]	0.035	[.0001]	-0.045	[.0015]	0.060	[.0001]	**-1.526**	[.0001]
	S&Y	-0.029	[.0798]	-0.014	[.0710]	0.149	[.0001]	-0.077	[.0001]	0.080	[.0001]	**-1.557**	[.0001]
	GME	0.007		-0.010		0.053		-0.024		0.071		**-1.455**	
	Mean	**-0.015**		**-0.009**		**0.068**		**-0.053**		**0.069**		**-1.514**	
	Std. Deviation	**0.015**		**0.009**		**0.055**		**0.023**		**0.009**		**0.043**	

Note: p-values are in brackets

[1] Calculated using sample means

Table 5.23. Compensated Own- and Cross-Price Elasticity Matrix[1] of Non-Alcoholic Beverages Using the QUAIDS and the 1999 ACNielsen Homescan Data (Unrestricted)

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Fruit Juice	ITSUR	**-0.723**	[.0001]	0.100	[.0001]	0.258	[.0001]	0.268	[.0001]	0.003	[.8061]	0.600	[.0001]
	H&W	**-0.683**	[.0001]	0.017	[.3264]	0.238	[.0001]	0.246	[.0001]	0.003	[.8260]	0.539	[.0001]
	S&Y	**-0.698**	[.0001]	0.071	[.0001]	0.502	[.0001]	0.275	[.0001]	0.003	[.8804]	0.687	[.0001]
	GME	**-0.716**		0.043		0.270		0.251		-0.172		0.469	
	Mean	**-0.705**		**0.058**		**0.317**		**0.260**		**-0.041**		**0.574**	
	Std. Deviation	**0.018**		**0.036**		**0.124**		**0.014**		**0.087**		**0.093**	
Tea	ITSUR	0.312	[.0001]	**-1.658**	[.0001]	0.179	[.0001]	0.361	[.0001]	0.037	[.0001]	0.622	[.0001]
	H&W	0.276	[.0001]	**-1.207**	[.0001]	-0.119	[.0001]	0.671	[.0001]	-0.254	[.0001]	1.424	[.0001]
	S&Y	0.089	[.0001]	**-1.383**	[.0001]	2.085	[.0001]	0.390	[.0001]	0.555	[.0001]	-0.184	[.0001]
	GME	0.337		**-1.199**		0.175		0.482		0.078		0.737	
	Mean	**0.254**		**-1.362**		**0.580**		**0.476**		**0.104**		**0.650**	
	Std. Deviation	**0.113**		**0.215**		**1.013**		**0.140**		**0.335**		**0.659**	
Coffee	ITSUR	0.185	[0.0001]	0.081	[.0001]	**-1.527**	[.0001]	0.342	[.0001]	0.083	[.0001]	0.454	[.0001]
	H&W	0.163	[0.0001]	-0.163	[.0001]	**-1.330**	[.0001]	0.304	[.0001]	0.070	[.0001]	0.351	[.0001]
	S&Y	0.240	[0.0001]	-0.052	[.1667]	**-3.728**	[.0001]	0.227	[.0001]	-0.042	[.2921]	0.580	[.0001]
	GME	0.228		0.039		**-1.484**		0.336		0.068		0.452	
	Mean	**0.204**		**-0.024**		**-2.017**		**0.302**		**0.044**		**0.459**	
	Std. Deviation	**0.036**		**0.108**		**1.144**		**0.053**		**0.058**		**0.094**	
Carbonated Soft drinks	ITSUR	0.106	[.0001]	0.096	[.0001]	0.178	[.0001]	**-0.572**	[.0001]	0.098	[.0001]	0.438	[.0001]
	H&W	0.110	[.0001]	0.214	[.0001]	0.184	[.0001]	**-0.578**	[.0001]	0.097	[.0001]	0.435	[.0001]
	S&Y	0.093	[.0001]	0.108	[.0001]	0.348	[.0001]	**-0.561**	[.0001]	0.112	[.0001]	0.461	[.0001]
	GME	0.067		0.073		0.152		**-0.611**		0.094		0.384	
	Mean	**0.094**		**0.123**		**0.216**		**-0.580**		**0.100**		**0.429**	
	Std. Deviation	**0.019**		**0.062**		**0.089**		**0.021**		**0.008**		**0.033**	

Table 5.23 Continued

		Fruit Juice		Tea		Coffee		Carbonated Soft Drinks		Bottled Water		Milk	
Bottled Water	ITSUR	0.486	[.0001]	0.167	[.0001]	0.385	[.0001]	0.783	[.0001]	**-1.912**	[.0001]	0.837	[.0001]
	H&W	0.378	[.0001]	0.358	[.0001]	0.261	[.0001]	0.593	[.0001]	**-1.467**	[.0001]	0.602	[.0001]
	S&Y	0.666	[.0001]	0.354	[.0001]	0.819	[.0001]	1.029	[.0001]	**-2.440**	[.0001]	1.180	[.0001]
	GME	0.451		0.236		0.347		0.818		**-1.892**		0.862	
	Mean	**0.495**		**0.279**		**0.453**		**0.806**		**-1.928**		**0.870**	
	Std. Deviation	**0.122**		**0.093**		**0.250**		**0.179**		**0.399**		**0.237**	
Milk	ITSUR	0.128	[.0001]	0.043	[.0001]	0.129	[.0001]	0.227	[.0001]	0.096	[.0001]	**-1.270**	[.0001]
	H&W	0.127	[.0001]	0.022	[.0594]	0.124	[.0001]	0.236	[.0001]	0.091	[.0001]	**-1.290**	[.0001]
	S&Y	0.121	[.0001]	0.026	[.0001]	0.243	[.0001]	0.216	[.0001]	0.112	[.0001]	**-1.311**	[.0001]
	GME	0.153		0.039		0.145		0.261		0.103		**-1.215**	
	Mean	**0.132**		**0.033**		**0.160**		**0.235**		**0.101**		**-1.271**	
	Std. Deviation	**0.014**		**0.010**		**0.056**		**0.019**		**0.009**		**0.041**	

Note: p-values are in brackets

[1] Calculated using sample means

Elasticity Comparisons across Censored Estimation Techniques of Non-Alcoholic Beverages

Table 5.14 presents the AIDS compensated or Hicksian price elasticity matrix of non-alcoholic beverages. We note more variability of cross price elasticities estimates of non-alcoholic beverage that are highly censored. These include tea, coffee and bottled water. On the other hand, relatively less variable cross-price elasticity estimates were observed for commodities with relatively minor censoring issues. For example, in milk, the cross-price elasticity estimates of milk with respect to fruit juice ranged from 0.152 to 0.264. Though not comparable, the cross-price elasticity values for bottled water with respect to fruit juice ranged from 0.011 to 0.492. Also note that associated p-values for all price elasticities are mostly significant. For the QUAIDS specification, we note the same claim that the greater number of censored observations the commodity, the more variable its respective own- and cross-price elasticities are. For milk the compensated price elasticities with respect to fruit juice ranged from 0.153 to 0.283, while for the bottled water, the compensated price elasticities ranged from -0.033 to 0.451 (Table 5.17). On the other hand, the same observation can be made for the AIDS and QUAIDS unrestricted cases. For example the cross price elasticity of milk with respect fruit juice ranged from 0.122 to 0.152 for AIDS and 0.121 to 0.153 for QUAIDS, while the cross price elasticity of bottled water with respect to fruit juice ranged from 0.372 to 0.675 for the AIDS specification and 0.378 to 0.666 for the QUAIDS model (Tables 20 and 23).

Elasticity Comparisons across Model Specification (AIDS vs. QUAIDS)

Table 5.14 and 5.17 present the compensated own- and cross-price elasticity matrices of non-alcoholic beverages of both the AIDS and QUAIDS models. We note relatively similar price elasticity estimates especially with respect to the own price elasticity values of both models. For example for milk, the range of the own price elasticities were from -0.951 to -1.211, whereas for the QUAIDS model, the values ranged from -1.015 to -1.215. Also if we look at a highly censored commodity such as bottled water, the cross price elasticity of bottled water with respect to tea ranged from 0.002 to 0.380 for the AIDS model and 0.004 to 0.428 in the QUAIDS specification. The same findings were also observed for the unrestricted cases of AIDS and QUAIDS where the calculated compensated price elasticities were remarkably similar.

Elasticity Comparisons across Imposition of Theoretical Restrictions

Tables 5.14 and 5.20 show the compensated own- and cross-price elasticity matrices of the AIDS restricted and unrestricted cases. Two notable results were observed; own price elasticity estimates (absolute values) were larger in the restricted case vis-as-vis the unrestricted case. On the other hand compensated cross price elasticities were generally larger in absolute terms in the unrestricted case relative to the values generated in the restricted case. The same result can also be observed for the QUAIDS restricted and unrestricted models (Tables 5.17 & 5.23).

Fit Comparisons across Econometric Techniques

Table 5.24 present the R-square values of the budget share equations from different censoring econometric techniques across demand system specification and imposition of theoretical restrictions. From the estimates, we find that across model specification and theoretical restrictions, the Heien and Wessells approach had the highest R-square values in its budget share equations. On the other hand, R-square values generated by the Shonkwiler and Yen technique registered second if theoretical restrictions are relaxed. Likewise, the ITSUR technique placed last across demand model specifications and theoretical impositions.

Conclusions

We find that the price elasticities especially the compensated price elasticities were robust and relatively similar and statistically significant across model specifications, estimation techniques and restriction impositions. The results of the compensated cross-price elasticities across the three categories were generally positive indicating that the respective non-alcoholic beverages are net substitutes. Comparative analysis show that across estimation techniques, greater variability of compensated cross-price elasticity estimates were observed in highly censored non-alcoholic beverages such as tea, coffee and bottled water. As for the comparison between model specification (AIDS versus QUAIDS), the compensated price estimates were remarkably similar especially for the own-price elasticity values. Finally, the estimates for unrestricted compensated cross price elasticities were generally greater vis-à-vis the restricted cases. The reverse is generally true with regard to the compensated own-price elasticity estimates.

Table 5.24. R-squared Values of Budget Share Equations from Different Censoring Econometric Techniques

Micro-Demand System Model	Econometric Techniques	Fruit Juice w_f	Coffee w_c	Soft Drink w_s	Bottled Water w_w	Milk w_m	Tea w_t
AIDS	ITSUR	0.0622	0.0673	0.0484	0.0764	0.0734	0.0184
	H&W	0.1937	0.3202	0.0966	0.2593	0.1441	0.0038
	S&Y	0.0629	0.0641	0.0479	0.0720	0.0744	0.0133
	GME (unrestricted)	0.0673	0.0695	0.0537	0.0801	0.0937	0.0145
	Dong et. al	0.0139	0.0484	0.0016	0.0676	0.0253	0.0101
QUAIDS	ITSUR	0.0636	0.0732	0.0517	0.0779	0.0734	0.0189
	H&W	0.1956	0.3259	0.1054	0.2602	0.1463	0.0037
	S&Y	0.0643	0.0702	0.0511	0.0740	0.0742	0.0155
	GME (unrestricted)	0.0681	0.0742	0.0571	0.0816	0.0940	0.0150
AIDS (unrestricted)	ITSUR	0.0672	0.0694	0.0532	0.0801	0.0940	0.0035
	H&W	0.1981	0.3257	0.1008	0.2649	0.1699	0.0113
	S&Y	0.0676	0.0697	0.0529	0.0766	0.0944	0.0005
	GME	0.0673	0.0695	0.0537	0.0801	0.0937	0.0145
QUAIDS (unrestricted)	ITSUR	0.0682	0.0697	0.0536	0.0804	0.0946	0.0030
	H&W	0.1995	0.3299	0.1106	0.2656	0.1721	0.0001
	S&Y	0.0696	0.1076	0.0562	0.0768	0.0958	0.0037
	GME	0.0681	0.0742	0.0571	0.0816	0.0940	0.0150

The robustness of both the parameter estimates and the calculated expenditure and price elasticities may be explained in part to the availability of high number of observations (n~30,000). However, since most censored data sets do not usually have this particular characteristic, then studies that simulate the effect of sample size will be beneficial on determining whether robustness will still be observed for parameter estimates and price and expenditures elasticities in the presence of differing sample sizes.

CHAPTER VI

CONCLUSIONS

This dissertation has produced a series of interrelated studies that focused from the examination of selected socio-demographic variables as potential drivers of organic and conventional milk choice, estimation of demand interrelationship of organic and conventional milk, examination of the sorting ability of binary choice models and to the estimation of a demand system that includes milk in a broader non-alcoholic beverage complex. These studies relied on the usage of 1999 and 2004 Nielsen Homescan Panel data.

In Chapter II, an attempt was made to look at the various socio-demographic drivers in terms of explaining household purchase of three milk types namely purchase of organic and conventional milk, purchase of organic milk only and purchase of conventional milk only. This examination was facilitated by the usage of both multinomial logit and multinomial probit models. The findings indicated that increasing household size, the presence of children, increasing educational level of household, hispanic households and those located in the west were identified as the key variables in explaining the likelihood of purchasing organic milk and the combination of organic and conventional milk. The study also found that little differences exist in the magnitudes of the marginal effects for both the multinomial logit and probit models. However the standard errors from the multinomial probit model are higher than the multinomial logit model thus more insignificant marginal effects with the multinomial probit model were observed than from the multinomial logit model.

In chapter III, a Heckman two-step correction was done in order to address the issue of sample selection in estimating the demand for both organic and conventional milk. Results from the first-stage probit analysis indicate that socio-demographic variables such household size, income, educational and employment levels of household head, race, ethnicity and regions were significant in explaining the likelihood of purchasing organic milk. Likewise, once the decision to purchase organic milk has been made, the findings indicate that variables such as household size, presence of children are associated with increased purchases of both organic and conventional milks. Also as household head educational level increases, purchases of organic milk also increases. The same also is true for white and oriental households where purchases of organic are more relative to black households. In terms of race, Hispanic households purchase more organic milk, while those located in the west purchase more organic milk relative to the other regions. Finally, the calculated elasticties indicate that both organic and conventional milks are substitutes. However the relationship is an asymmetric one, where the demand for organic milk is more sensitive to price changes in conventional milk but changes in the price of organic milk has relatively little impact on the demand for conventional milk.

In Chapters II and III, binary choice models were used in evaluating behavioral choices with regard to two alternatives. And because of their usefulness, methods such as the prediction-success contingency tables have been a standard measure in evaluating the ability of models to make correct predictions. However, these types of methods are centered on the assumption of a symmetric loss function with a default cut-off value of 0.5. And a major critique of this method is it does not address the quality of predicted probabilities in that is there is no discrimination whether the predicted probability is 51 percent or 99 percent. Thus, Chapter IV focuses on the assessment of binary choice models through alternative methods such as probability scores. In this chapter both the Brier Score and Yates Brier Score Decomposition were used. Results show that when important socio-demographic variables are omitted, scatter and minimum variance values are significantly reduced. An intuitive explanation for this change might lie in the variability reduction of the predicted probabilities. Also the removal of important socio-demographic variables resulted in a weakened ability to sort between events that occurred and did not occur.

Finally in Chapter V, the study estimated both censored AIDS and QUAIDS demand systems involving non-alcoholic beverages such as fruit juice, tea, coffee, carbonated soft drinks, bottled water and tea. The highlight of the study involved the usage of different estimation techniques that addressed censoring in demand systems. These include two-step estimation techniques such as the Heien and Wessells (1990) and Shonkwiler and Yen (1999) approaches, general maximum entropy and the Dong, Gould and Kaiser (2004a) methods. The study also included the use of ITSUR without adjustments for censoring as a means of acting as a base estimator relative to the other techniques. The results show that the estimated elasticities bear little difference with the estimates from past studies and most of the commodities in the non-alcoholic beverage complex are net substitutes. Likewise across censoring techniques, variability of cross price elasticities was observed especially for those beverages that are highly censored such as tea, coffee and bottled water. On the other hand when comparisons are made across model types, compensated price elasticities especially the own price elasticities were remarkably very similar. Also compensated cross price estimates from the unrestricted AIDS and QUAIDS models were relatively greater compared to the restricted cases, but the reverse is true with regards to the compensated own price elasticities.

From a marketing standpoint, the implications for organic milk are clear, that the results of the dissertation particularly those of Chapters II and III imply crucial inputs in terms of designing marketing strategies that can target demographic groups such as single person, college educated head, Hispanic households. However, since the data were compiled from a 2004 data set, a more updated database might provide richer insights as to whether significant changes have occurred with regards to organic milk preference.

In terms of methodological implications, the chapter on Brier score provides valuable insights in using alternative techniques such as the Yates partition in complementing the use of prediction-success tables. More importantly, binary choice specifications that omit important drivers may achieve some noise reduction but at the cost of weakening the ability of models to sort alternative events. And finally, since many censored data sets do not have the luxury of very high sample sizes, a future area of research might be determining robustness through simulation of different levels of sample sizes and its effect on the estimated elasticities in a censored demand system framework. Also one can simulate alternative error term specifications and determine whether robustness still holds in all of the considered techniques that address censored demand systems.

REFERENCES

Alvarez, M. and J. Nagler. 1994. “Correlated Disturbances in Discrete Choice Models: A Comparison of Multinomial Probit Models and Logit Models.” Division of the Humanities and Social Sciences, California Institute of Technology, Working paper No 914.

Alviola IV, P. and O. Capps, Jr. 2009. “Household Demand Analysis of Organic and Conventional Fluid Milk in the United States Based on the 2004 Homescan Panel.” Dept. of Agricultural Economics, Texas A&M University, Research Report.

Arndt, C. 1999. “An Entropy Approach to Treating Binding Non-Negativity Constraints in Demand Systems Estimation and an Application to Herbicide Demand in Corn.” *Journal of Agricultural and Resource Economics* 24: 204-221.

Arndt, C., S. Liu and P.V. Preckel. 1999. “On Dual Approaches to Demand Systems Estimation in the Presence of Binding Quantity Constraints.” *Applied Economics* 31(8), 999-1008.

Baltagi, B.H. 2005. *Econometrics*. New York: Springer-Verlag Berlin.

Banks, J., R. Blundell and A. Lewbel. 1997. “Quadratic Engel Curves and Consumer Demand.” *The Review of Economic and Statistics* 79: 527-539.

Bernard, D.J. and A. Mathios. 2005. “Factors Affecting Consumer Choice and Willingness to Pay for Milk Attributes.” Selected Paper at the American Association of Agricultural Economics Annual Meeting, Providence, Rhode Island, July 24-27.

Bessler, D.A. and R. Ruffley. 2004. “Prequential Analysis of Stock Market Returns.” *Applied Economics* 36: 399-412.

Breslaw, J.A. 2002. “Multinomial Probit Estimation without nuisance parameters.” *Econometrics Journal* 5: 417-434.

Bunch, D.S. 1991. “Estimability in the Multinomial Probit Model.” *Transportation Research* 25B: 1-12.

Cameron, A.C., and P.K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Cameron, A.C. and P.K. Trivedi. 2008. *Microeconometrics using STATA*. STATA Press, StataCorps LP, College Station, Texas.

Capps, Jr. O., and R.A. Kramer.1985. “Analysis of Food Stamp Participation Using Qualitative Choice Model.” *American Journal of Agricultural Economics* 67: 49-59.

Capps, Jr. O., H.A. Love, G.W. Williams, and W.L. Adams. 1999. “Examining Packer Choice of Slaughter Cattle Procurement and Pricing Methods.” *Agricultural and Resource Economics Review* 28: 12-25.

Cheng, S. and J.S. Long. 2007. “Testing IIA in the Multinomial Logit Model.” *Sociological Methods and Research* 35: 583-600.

Cox, T.L. and M.K. Wohlgenant. 1986. “Price and Quality Effects in Cross-Sectional Demand Analysis.” *American Journal of Agricultural Economics* 68: 908-919.

Deaton, A., and J. Muelbauer. 1980. “An Almost Ideal Demand System.” *American Economic Review* 70: 312-326.

Demeritt, L. 2004. “Organic Pathways.” *[N]sight Magazine* VI(2):16-21.

Dhar, T and J.D. Foltz. 2004. “Is Soy Milk? The Economics of the Soy Milk Market.” Paper presented at the American Association of Agricultural Economics Meeting, Denver, USA.

Dhar, T and J.D. Foltz. 2005. “Milk by Any Other Name Consumer Benefits from Labeled Milk.” *American Journal Agricultural Economics* 87: 214-218.

Dimitri, C. and C. Greene. 2002. “Recent Growth Patterns in U.S. Organic Market.” *Agriculture Information Bulletin,* No. 777. Washington, D.C.: US Department of Agriculture, Economic Research Service.

Dimitri, C. and K.M. Venezia. 2007. “Retail and Consumer Aspects of the Organic Milk Market.” Outlook Report No. LDP-M-155-01. U.S. Department of Agriculture, Economic Research Service.

Dong, D., B.W. Gould and H.M. Kaiser. 2004a. “Food Demand in Mexico: An Application of the Amemiya-Tobin Approach to the Estimation of a Censored Food System.” *American Journal Agricultural Economics* 86: 1094-1107.

Dong, D., C. Chung, and H. Kaiser. 2004b. “Modeling Milk Purchasing Behavior with a Panel Data Double-Hurdle Model.” *Applied Economics* 36: 769-779.

Dong, D., J.S. Shonkwiler, and O. Capps, Jr. 1998. “Estimation of Demand Functions using Cross-Sectional Household Data: The Problem Revisited.” *American Journal of Agricultural Economics* 80: 466-473.

Dorfman, J.H. 1996. “Modeling Multiple Adoption Decisions in a Joint Framework.” *American Journal Agricultural Economics* 78: 547-557.

Dow, J. and J. Endersby. 2004. “Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research.” *Electoral Studies* 23: 107-122.

Dubin, J.A. and D.L. McFadden. 1984. “An Econometric Analysis of Residential Electric Appliance Holdings and Consumption.” *Econometrica* 52: 345-362.

DuPuis, E.M. 2000. “Not in My Body: rBGH and the Rise of Organic Milk.” *Agriculture and Human Values* 17: 285–295.

Gan, L. 2007. Lecture Notes for Econometrics III, Department of Economics, Texas A&M University, College Station.

Glaser, L.K. and G.D. Thompson. 2000. “Demand for Organic and Conventional Beverage Milk.” Paper presented at Western Agricultural Economics Association Annual Meeting, Vancouver, Canada.

Golan, A., J.M Perloff and E.Z. Shen. 2001. “Estimating a Demand System with Non-negativity Constraints: Mexican Meat Demand.” *The Review of Economics and Statistics* 83: 541-550.

Gould, B. 1996. “Factors Affecting U.S. Demand for Reduced Fat Fluid Milk.” *Journal of Agricultural and Resource Economics* 21: 68-81.

Green, R., and J. Alston. 1990. “Elasticities in the AIDS Models.” *American Journal of Agricultural Economics* 72: 442-445.

Greene, W.T. 2008. *Econometric Analysis*. 6th Edition. Upper Saddle River, New Jersey: Prentice Hall.

Hammarlund, R. 2001. “A Study of Marketing Issues with Organic Milk.” Unpublished M.S. Thesis, Kansas State University, Manhattan.

Hausman, J. A. and D. McFadden. 1984. “Specification Tests for the Multinomial Logit Model.” *Econometrica* 52: 1210-1240.

Hausman, J. A. and D. Wise. 1978. “A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences.” *Econometrica* 46: 403-426.

Heckman, J. 1976. “The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variable and a Simple Estimator for Such Models.” *Annals of Economic and Social Measurement* 5: 475-492.

Heckman, J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 47: 153-161.

Heien, D. and C.R. Wessells. 1990. “Demand Systems Estimation with Microdata: A Censored Regression Approach.” *Journal of Business and Economic Statistics* 8: 365-371.

Hill, H. and F. Lynchehaun. 2002. “Organic Milk: Attitudes and Consumption Patterns.” *British Food Journal* 104: 526-542.

Hruschka, H. 2007. “Using a Heterogeneous Multinomial Probit Model with a Neural Net Extension to Model Brand Choice.” *Journal of Forecasting* 26: 113-127.

Jepsen, C. 2008. “Multinomial Probit Estimates of College Completion at 2-year and 4-year Schools.” *Economics Letters* 98: 155-160.

Judd, K. 1998. *Numerical Methods in Economics.* Cambridge, Massachusetts: The MIT Press.

Kropko, J. 2008. "Choosing Between Multinomial Logit and Multinomial Probit Models for Analysis of Unordered Choice Data." Paper presented at the Annual Meeting of the MPSA Annual National Conference, Palmer House Hotel, Chicago, Illinois, April 03.

Lewbel, A. 1997. “Constructing Instruments for Regressions with Measurement Error when No Additional Data are Available, with an Application to Patents and R&D.” *Econometrica* 65: 1201-1213.

Li, J., L. Zepeda, and B. Gould. 2007. “The Demand for Organic Food in the US: An Empirical Assessment”. *Journal of Food Distribution Research* 38: 54-69.

Long, J.S. and J. Freese. 2006. *Regression Models for Categorical Dependent Variables Using STATA*, 2nd Edition. College Station, Texas: Stata Press.

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

McFadden, D. 1973. “Conditional Logit Analysis of Qualitative Choice Behavior”. In P. Zarembka, ed, *Frontiers in Econometrics.* New York: Academic Press, pp. 105-142.

McFadden, D. 1974a. “The Measurement of Urban Travel Demand.” *Journal of Public Economics* 3: 303-328.

McFadden, D. 1974b. "Conditional Logit Analysis of Qualitative Choice Behavior." In P. Zarembka, ed. *Frontiers in Econometrics*. New York: Academic Press, pp. 105-144.

McFadden, D. 1978. "Modeling the Choice of Residential Location." In A. Karlquist, L. Lundquist, F. Snickars and J.L. Weibull, eds. *Spatial Interaction Theory and Planning Models.* North Holland, Amsterdam: Elsevier Science Ltd, pp. 75-96.

McKnight, H. 2007. "Organic Milk: Consumers and Their Purchasing Patterns." Unpublished M.S. Thesis, Virginia Polytechnic Institute and State University, Blacksburg.

Meyerhoefer, C., C.K. Ranney and D.E. Sahn. 2005. "Consistent Estimation of Censored Demand Systems using Panel Data." *American Journal Agricultural Economics* 87: 660-672.

Miller, J.J. and D.P. Blayney. 2006. *Dairy Backgrounder*. Outlook Report No. LDP-M-145-01. U.S. Department of Agriculture, Economic Research Service.

Mutuc, M.E., S. Pan, and R.M. Rejesus. 2007. "Household Vegetable Demand in the Philippines: Is There an Urban-Rural Divide." *Agribusiness: An International Journal* 23: 511-527.

Nakamura, A. and M. Nakamura. 1998. "Model Specification and Endogeneity." *Journal of Econometrics* 83: 213-237.

Nobile, A., C. Bhat and E. Pas. 1996. "A Random Effects Multinomial Probit Model of Car Ownership Choice." Research Paper of Duke University and University of Massachusetts at Amherst.

Olvera, G.C. and D.A. Bessler. 2006. "Probability Forecasting and Central Bank Accountability." *Journal of Policy Modeling* 28: 223-234.

Organic Valley. 2005. “Organic Valley Culminates 2005 with Launch of ‘Generation Organic’,” http://organicvalley.coop/newsroom/article.html?/andid=201, on February 2, 2006.

Park, J. and G. Davis. 2001. “The Theory and Econometrics of Health Information in Cross-Sectional Nutrient Demand Analysis.” *American Journal of Agricultural Economics* 83: 840-851.

Park, J. and O. Capps, Jr. 1997. “The Demand for Prepared Meals by US Households.” *American Journal of Agricultural Economics* 79: 814-824.

Pittman, G. 2004. “Drivers of Demand, Interrelationships, and Nutritional Impacts Within the Non-Alcoholic Beverage Complex.” PhD Dissertation, Texas A&M University, College Station.

SAS Institute Inc. 2008. *SAS/ETS ® 9.2 User’s Guide*. Cary, North Carolina: SAS Institute Inc.

Saha, A., O. Capps, Jr. and P.J. Byrne. 1997. “Calculating Marginal Effects in Dichotomous-Continuous Models.” *Applied Economic Letters* 4: 181-185.

Schmit, T., C. Chung, D. Dong, H. Kaiser, and B. Gould. 2002. “Identifying the Effects of Generic Advertising on the Household Demand for Fluid Milk and Cheese: A Two-Step Panel Data Approach.” *Journal of Agricultural and Resource Economics* 27:165-186.

Shonkwiler, J.S. and S.T. Yen. 1999. “Two-Step Estimation of a Censored System of Equations.” *American Journal of Agricultural Economics* 81: 972-982.

Small, K. and C. Hsiao. 1985. “Multinomial Logit Specification Tests.” *International Economic Review* 26: 619-627.

Stock J. H and M.W. Watson. 2007. *Introduction to Econometrics*. Boston Massachusetts: Pearson-Addison Wesley.

Stockton, M.C., O. Capps, Jr and D. Dong. 2007. “A Micro Level Demand Analysis of Selected Non-Alcoholic Beverages with Emphasis on Container Size.” Department of Agricultural Economics, Texas A&M University, Research Report.

StataCorp. 2005. Stata Statistical software: Release 9. College Station, Texas: StataCorp LP.

Thompson, G.D. 1998. “Consumer Demand for Organic Foods: What We Know and What We Need to Know,” *American Journal of Agricultural Economics* 80: 1113-1118.

Tiffin, R., and M. Aquiar. 1995. “Bayesian Estimation of an Almost Ideal Demand System for Fresh Fruit in Portugal.” *European Review of Agriculture Economics* 22: 469-80

Tomek, W.G and K.L Robinson. 2003. *Agricultural Product Prices*, 4th Edition. Ithaca, New York: Cornell University Press.

Train, K. 2003. *Discrete Choice Methods with Simulation*. New York: Cambridge University Press.

Vergara, O., K. Coble, T. Knight, G. Patrick and A, Baquet. 2004. “Cotton Producers’ Choice of Marketing Techniques.” *Agribusiness: An International Journal* 20: 465-479.

Weeks, M. 1997. “The Multinomial Probit Model Revisited: A Discussion of Parameter Estimability, Identification and Specification Testing.” *Journal of Economic Surveys* 11: 297-320.

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data Method*. Cambridge, Massachusetts: The MIT Press.

Yates, F.T. 1982. "External correspondence: Decomposition of the Mean Probability Score." *Organizational Behavior and Human Performance* 30: 132-156.

Yates, F.T. 1988. "Analyzing Accuracy of Probability Judgments for Multiple Events: An Extension of the Covariance Decomposition." *Organizational Behavior and Human Performance* 41: 281-299.

Yen, S.T., B. Lin, and D.M. Smallwood. 2003. "Quasi and Simulated Likelihood Approaches to Censored Demand Systems: Food Consumption by Food Stamp Recipients in the United States." *American Journal of Agricultural Economics* 85: 458-478.

VITA

Name:	Pedro Aya-ay Alviola IV
Address:	Agricultural Economics and Agribusiness AGRI 217 1 University of Arkansas Fayetteville, AR 72701
Email Address:	petealviola@yahoo.com
Education:	B.S., Economics, University of the Philippines, 1991 M.A., Economics, University of the Philippines, 1997 Ph.D., Agricultural Economics, Texas A&M University, 2009