THE FACTORS INFLUENCING ORGANIC PRODUCE DEMAND: EVIDENCE FROM HOUSEHOLD SCANNER PANEL DATA

by

Scott Michael Albrechtsen

A Thesis Submitted to the Faculty of the

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2 0 0 7

STATEMENT BY AUTHOR

SIGNED: ____________________

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

____________________ __________

Dr. Gary D. Thompson — Date
Professor of Agricultural and Resource Economics

ACKNOWLEDGEMENTS

I would like to thank Dr. Gary D. Thompson for his guidance, encouragement, and understanding throughout this project. I appreciate his sincere concern for my academic development and his insightful suggestions for study and research. I owe a great deal to him both academically and professionally. Thanks to my committee members, Dr. Russell Tronstad and Dr. Satheesh Aradhyula for their time. I would also like to thank my parents, Paul and Mary-Lou Albrechtsen for their unwavering support for everything that I have done in my life.

TABLE OF CONTENTS

TABLE OF CONTENTS – *Continued*

TABLE OF CONTENTS – *Continued*

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Existing literature analyzing consumer demand for organically grown produce has primarily consisted of studies that contain self-reported data from a specific region in the US. Research utilizing direct purchasing behavior of organic produce at the household level is almost non-existent. With the use of ACNielsen household scanner data, consumer demand and loyalty to organic fresh products are analyzed as a function of own price, income, conventional prices, household composition, and demographic attributes. While organic products still maintain a small market share in the produce industry, an empirical model must correctly account for corner solutions where a high percentage of households choose conventional over organic items. In controlling for unobserved heterogeneity of households in a panel model, empirical estimates suggest that price effects, the presence of adolescent children, larger households, income levels, and female head of household demographics are statistically significant influences on organic produce purchase decisions at the household level.

CHAPTER 1

INTRODUCTION

Historically, consumer demand studies of organic products have relied on data that is based on self-reporting of purchase behavior and attitudes (Thompson, 1998). Differences between recalled consumption, reported willingness to pay, and actual consumption within such self-reported studies often lead to unreliable estimates and results. Direct observation of purchase activity as it is related to organic purchases is almost non-existent within current literature (Thompson, 1998). As household scanner data steadily improves (or becomes available) throughout time, the ability to reliably measure household organic purchase activity has recently emerged. Though organic products still maintain a relatively small market share, household panel studies have now been able to decipher between organic and non-organic purchases and thus provide the means for analyzing organic consumption. Now researchers have the opportunity to analyze purchase activity as affected by own price, cross-prices, income, and other household characteristics.

This study examines what factors may influence household decisions as to purchase organic rather than non-organic fresh fruit and vegetable items. One may hypothesize that those who consume organically grown fruit and vegetable items could have a strong loyalty to these items due to personal beliefs concerning disparities between organic and non-organic fresh produce. If, for example, consumers believe that organic produce provides increased health benefits from diminished exposure to pesticide or herbicide contaminants, these beliefs could enhance some element of the consumer's value for organic produce and thus influence organic produce demand.

On the other hand, undoubtedly there are individuals who base their decision to purchase organic versus non-organic produce purely on price. These individuals might prefer to consume organic versus non-organic items; however, due to budget constraints, higher organic prices may limit purchases of organic produce. This study aims to address these demand effects.

The emergence of improved data collection methods in the last decade, especially at the household level, has allowed industry managers the ability to successfully conduct analyses of trade promotions and price elasticities on consumer products (Bucklin, 1999). This study will contribute to the rather recent literature regarding organic produce purchase decisions and perhaps encourage further study regarding this subject.

1.1 Why Do Consumers Buy Organic?

Within the last decade, market share for organic produce has steadily increased. With increased consumer consciousness about potential harmful pesticide and herbicide contaminants, more consumers are interested in produce that is produced organically. As will be discussed later, intrinsically related values and concerns seem to have a strong effect on household purchases of organic produce regardless of income.

According to The Packer (2007), the primary reason consumers choose organic products over conventional is to avoid chemicals in their food. In a 2006 study, 62 % of organic shoppers explained that the avoidance of potentially harmful chemicals was their main reason for purchasing organics.

Consumers also derive added utility from organic products for other personal reasons. More than one-third of those surveyed by The Packer (2007) reported buying organic for the improved nutrient content over conventional produce. Also, forty-three percent of consumers said that organic produce helps promote their personal health while 40% buy organic for perceived superior taste.

Finally, issues of social and environmental accountability also drive the decision to purchase of organic produce for some consumers. Over 37% of survey participants reported buying organics because they feel a strong social responsibility to protect the environment.

1.2 Organic Price Premiums

Studies have shown that some consumers are willing to pay significant price premiums for organic produce compared to their conventional counterparts. Of those shoppers who purchased organic products last year, 37% reported paying a higher premium of 10% to 24% while an additional 9% of respondents payed between 25% and 49% more than conventional products (The Packer, 2007).

For other consumers, these price premiums prove to be too high. For many, organic produce purchases are never observed regardless of availability. The high price ratio between organic and conventional items primarily motivates why organic producers still observe low market share percentages within the produce industry.

1.3 Problem Statement

With a large household consumer panel, purchase activity as it relates to produce consumption can be observed throughout time. In estimating a panel model for purchases that accounts for a high volume of zero censored observations while controlling for unobserved heterogeneity, the effects of price, income, and demographics on a household's decision to buy organic over conventional items maybe estimated. In comparing empirical results with other studies based on self-reported and nationally aggregated data, the results can be scrutinized.

1.4 Thesis Organization

The subsequent parts of this thesis are organized as follows: Chapter Two discusses the preceding economic literature regarding organic demand factors and offers potential contributions of this study. Chapter Three examines the theoretical framework behind censored panel analysis. Chapter Four discusses the data used for empirical study and the rather involved process of data manipulation for analysis. Chapter Five offers summary statistics and graphical interpretations related to the data. Chapter Six discusses the empirical models and estimation techniques used.

Chapter Seven reports empirical results and the final chapter provides a summary of empirical results and provides concluding remarks regarding the study.

CHAPTER 2

LITERATURE REVIEW

2.1 Existing Consumer Demand Analysis for Organic Produce

At present it would seem that current investigation and publication of studies on organic produce demand have not kept up with growing demand trends themselves. Few studies exist with a rather inclusive examination of price responsiveness, income effects, household characteristic effects, and store effects as they influence organic produce demand. Studies that address some of these factors exist; however, limitations in data have precluded comprehensive study of multiple effects. The studies that do exist have become quite dated and more recent literature is to be desired.

2.1.1 Price Elasticities

The existing literature on price elasticities of organic food products is quite short. Park and Lohr (1996) estimate equilibrium farm price and quantity for the wholesale organic produce industry. In their model, farmers represent the supply side while wholesalers represent the demand side. They argue that wholesalers exhibit effects of factors within the market and therefore reflect demand from retailers and ultimately consumers. Using a nonlinear two-stage least squares (2SLS) model, they estimate equilibrium farm price elasticities for organic wholesalers. They also estimate price margin elasticities where the difference in price between conventional and organic wholesale prices affects equilibrium output. While the study makes the argument that wholesaler demand reflects consumer demand, the study does not discuss own- and cross-price elasticities for the organic produce items discussed.

Thompson (1998) notes the wide potential for expansion into demand analysis for organic produce. He explains that little has been done in terms of estimation of price elasticities for organic produce. He highlights works (1998 and prior) that have attempted to capture organic consumption trends, their means of data generation and explains that self-reported data, rather than direct observation of consumer behavior in retail markets, dominate these studies. He comments that price elasticity analysis is needed as non-reported price values in previous studies have hindered estimation.

More recent unpublished works by Glaser and Thompson have used demand system estimation to determine uncompensated price elasticities for related organic and conventional items. In Glaser and Thompson (1998), demand estimation is done for organic and conventional frozen vegetables, (2000) for organic and conventional beverage milk, and (2001) for organic and conventional baby food. Data used for analysis in these studies are based on supermarket scanners from ACNielsen's Marketing Research and Information Resources Inc(IRI) and are of a different type than the household level data used for this study. In these studies, price and expenditure elasticities are estimated using non-linear almost ideal demand systems (AIDS) for (1998) and (2000) and the quadratic almost ideal system (QUAIDS) for (2001). These studies are some of the first to utilize actual purchase data from ACNielsen and IRI for analysis of retail organic demand and exhibit the potential for future research and publication. The range of food products analyzed was limited to frozen vegetables, fluid milk, and baby food.

In the AIDS model, Glazer and Thompson express budget share for the i^{th} organic item in the functional form for the AIDS demand function:

$$w_i = \alpha_i + \sum_j \gamma_{ij}\ log\ p_j + \beta_i log(x/P) \tag{2.1}$$

where w_i is the budget share of good i, α_i, β_i, and γ^*_{ij} are parameters and P is a price index defined as:

$$log\ P = \alpha_0 + \sum_k \alpha_k\ logp_k + \frac{1}{2}\sum_j\sum_k \gamma_{kj}\ logp_k\ logp_j \tag{2.2}$$

Considering the conditions for adding-up, for homogeneity, and for Slutsky symmetry, (2.1) represents a system of demand functions used to derive uncompensated or Marshallian elasticities of demand for organic own-prices and cross-prices of conventional related items.

2.1.2 Income

Adhering to conventional economic theory of demand, one would expect that consumers with higher income levels would be more likely to be purchasers of organic produce. Curiously, many modern studies of organic produce demand seem to diverge in their results regarding income effects on organic purchases.

Verifying theory, Park and Lohr (1996) found the effect of personal income on demand for organic broccoli and lettuce to be large, positive, and significant. Within the broccoli and lettuce organic retail market, their study found that consumers with higher incomes were able to pay more for organic produce. They also found that the effect of income on organic carrot demand was insignificant and close to zero however. Thompson and Kidwell (1998) found that households within multiple income brackets had a positive and significant likelihood of purchasing organic. Increasingly large classification levels of income were found to have a higher likelihood of purchasing organic. Jolly (1991) found that willingness to pay for organic produce increases with income. Conversely, the study also found that the level of household income was not a significant explanatory variable in explaining divergent buying behavior between organic buyers and non-buyers.

Contrary to theory, multiple studies found income effects to be not significant in influencing organic produce demand. Ott (1990) found that income was not important in explaining shoppers' willingness to pay for Certified Pesticide-Residue-Free (CPRF) produce. Also, Buzby and Skees (1994) found that in willingness to pay for increased reductions in pesticide exposure in produce, income had no effect. Based on evidence from Baker and Crosbie (1993) and Thompson and Kidwell (1998), Thompson (1998) mentions that higher income does not necessarily mean a high likelihood of organic purchases. This concept suggests that income related lifestyles, values, and concerns have key roles to play in the organic food purchasing decision.

As mentioned by Thompson (1998), "some consumer segments with relatively lower incomes seem to be more entrenched buyers of organic products and tend to have shopped for organics at retail outlets" (1115).

2.1.3 Household Size

In exploring the effect that the size of the household has on consumption of organic produce, mixed results have been discovered. Jolly (1991) found that two-person households with a household income in excess of $30,000 had the highest willingness to pay for organic produce while willingness to pay decreased as household size increased.

On the other hand, Huang (1996) found that willingness to buy organic produce increased with household size based on a household sample in Georgia at a statistically significant level. Furthermore, Thompson and Kidwell (1998) found that the number of children under the age of 18 had a positive, significant effect on the probability of purchasing organic produce.

2.1.4 Occupation

Literature exploring the effects of occupation on organic produce consumption still is quite limited. Jolly (1991) found that white-collar and service occupations were more likely to be buyers of organic produce. As a possible explanation, Jolly rationalizes that white-collar occupations are more independent of consumer reference groups than other occupations. These individual consumers tend to be more innovative in their purchases. Thus, organic purchases become more of a viable purchase decision for these groups rather than others.

2.1.5 Age

Some studies have verified that an individual's age has significant effects on organic produce consumption while others have maintained that it is not a significant factor in demand. Misra et al. (1991) revealed that individuals over 60 years old were willing to pay more for CPRF produce than those within the age of 36 to 60 years old with statistical significance. Interestingly, Buzby and Skees (1994) find that younger respondents are willing to pay more for reductions in pesticide exposure than older respondents. Conversely, Jolly (1991) found that no trend existed between willingness to pay for organics and age. The study found that the effect of different age brackets were not statistically different from zero.

2.1.6 Education

In terms of willingness to pay, Jolly (1991) found that those with post-secondary degrees and graduate degrees were willing pay the most for organic produce. Similarly, Huang (1996) found that the number of years of education had a positive and significant effect on preference for organically grown produce.

Thompson and Kidwell (1998) observed conflicting results regarding the influence of education on consumption. The study found that having a college degree had positive, yet insignificant effect on the likelihood of purchasing organic produce. In contrast, the study found that having a graduate or professional degree had a negative effect on the probability of an organic purchase at the 10 % significance level. Additionally, Misra et al. (1991) found that college-educated consumers were more price elastic to pesticide-free produce with statistical significance.

2.1.7 Race

Ott (1990) found that whites showed more propensity to purchase CPRF produce rather than non-whites. In the study, 69% of whites were willing to pay higher prices for CPRF produce while only 54% of non-whites would pay higher prices. Furthermore, Misra et al. (1991) found that consumers of white background were less price elastic than non-whites with a strong statistical significance. Buzby and Skees (1994) found that race had no effect on willingness to pay for increased pesticide reductions.

2.1.8 Martial Status

With respect to martial status and its effect on organic demand, the literature is rather short. Jolly (1991) reports that married couples are less likely to pay for organic produce. He reports that singles with less than $ 30,000 income are willing to pay 44% more while singles with incomes $ 50,000 and above are willing to pay 46.1 % more for organic produce. In comparison to married couples, willingness to pay never exceeds 37.6 % more for organics.

2.1.9 Geographic Scope and Survey Methods

Many previous studies have had a rather limited scope in terms of geography and have relied heavily on self-reporting via telephone, mail surveys, and in-store surveys at point of purchase. Often, limited geographical scope has hindered the applicability of these studies to a national scale. Furthermore, one may question the validity that self-reporting may have on accurate estimates especially within the context of mail and telephone surveys. Shown in table 2.1 is the geographic scope of the studies discussed.

2.1.10 Store Effects

Studies have attempted to estimate the effects that purchasing at specific stores had on consumption patterns. Thompson and Kidwell (1998) found considerable disparity of consumption patterns between the two retail venues used within their analysis. In comparison of purchases at a specialty grocery store as opposed to a food cooperative, effects on organic purchases were quite different. Purchases at a specialty store had a strong negative effect on the likelihood of an organic purchase at a statistically significant level. Thompson (1998) suggests that as long as organic products remain being regularly unavailable in most mainstream supermarkets, store choice is a critical variable in determining whether an organic purchase will occur.

Table 2.1: Geographic Scope of Previous Organic Produce Studies

Year[a]	Author	Geographic Scope
1987	Jolly	Three Counties, California
1989	Misra, Huang and Ott Huang	Statewide, Georgia
1989	Park and Lohr	California
1989	Morgan *et al.*	Statewide, New Jersey
1990	Ott	Atlanta, Georgia
1994	Thompson and Kidwell	Tucson, Arizona
1994	Buzby and Skees	National
1988-1996	Glaser and Thompson	National

[a]Year implies the year of data collection

2.1.11 Organic Produce Seasonality

A large number of analyses related to the demand of perishable produce items have examined seasonal and regional variations as a source of changes in consumption over time. Thompson and Wilson (1999) note that seasonality in perishable food consumption is most apparent in econometric models employing scanner data with high frequency observations similar to the data utilized in this study. For such models, they acknowledge the imperative to properly account for seasonality and provide evidence of demand influences in seasonality trends.

Often studies model seasonality but do not address why these trends occur. Thompson and Wilson note that seasonality is usually cast as a supply phenomenon related to climate effects but note that various demand characteristics likely have a strong influence on seasonal changes in consumption. They refer to a physiological link between eating patterns and climate or weather that influence demand patterns. They refer to Brewerton et al. (1994) who highlights the tendency to

Table 2.2: Survey Methods of Previous Organic Produce Studies

Author	Survey method	Sample Size[a]
Jolly	Mail	955 (54%)
Misra, Huang and Ott	Mail	389 (67%)
Park and Lohr	Mail	955 (54%)
Morgan	Mail	552 (?)
Ott	Interview Questionnaire	315 (?)
Thompson and Kidwell	In-Store Interview Questionnaire	360 (?)
Buzby and Skees	Telephone Survey followed by Mail	2,098 / 1,671 (65%)/ (76%)
Glaser and Thompson	In-Store Scanner Data	3,000[b] / 13,000[c]

[a]Response rate in parentheses (eg. 54% response rate resulted in a sample size of 955)
[b] ACNeilsen Marketing Research Store Data
[c] Information Resources Inc. Supermarket Data

consume "lighter" foods such as fruits and vegetables during the summer where "heavier" foods are consumed in more abundance during the winter. Also, location differences were found to affect seasonal consumption patterns of bagged salads as more northern locations displayed more variation across season than those closer to the equator.

Such observations will become more interesting in the latter parts of this study as they could potentially provide explanation for observed seasonal consumption trends for organic fresh produce.

2.2 Contribution to the Literature

2.2.1 Tracking National Trends

No current studies exist that use direct observations of price and quantity information to examine national trends in household organic produce consumption. While Glaser and Thompson (1998) use national-level retail scanner data, the influence of household level characteristics and demographic factors on organic frozen vegetables could not be examined. Due to the rather limited geographic focus of other existing studies, their results are not suitable for tracking national consumption trends. Based on region, likelihood of purchases of organic produce across America is quite different (The Packer, 2007).

With the existence of ACNielsen household scanner data used in this analysis, consumption patterns across the country can be examined. Due to ACNielsen efforts to balance their panel to represent the household population of the mainland US, the existing data set allows for applicable examination and estimation across the country. Controlling for location with the ability of estimation across 52 distinct market areas clearly offers an avenue of research that has yet to be explored within the literature.

2.2.2 Direct Organic Purchasing Behavior

Shown in table 2.2 are the survey methods, number of observations and the response rates for the studies discussed. As mentioned, a large proportion of existing literature employs self-reported data from questionnaires and surveys. This leads to multiple empirical problems associated with respondent recall, truthfulness and reliability. As foreseen by Krissoff (1998), as organic produce sales have increased in supermarkets over the years, the availability of scanner data, both at the retail and household level have become accessible for further research. The aim of this study is to add a consumer demand study of organic produce to the existing literature with more precision than studies that involve self-reporting for empirical analysis.

2.2.3 Increased Store Effect Analysis

The reporting of store, market, and channel information within the ACNielsen panel allows for almost unlimited study of store effects on consumption. Though a potentially laborious task of categorizing store information would exist, creative incorporation could greatly contribute to the literature as it pertains to store type and specific purchasing channel.

As will be discussed in Chapter 4, the construction of the household panels used for analysis indirectly captures store pricing effects as they pertain to sales of both conventional and organic produce items at local chains, produce markets, and cooperatives. In this analysis, a nationally reaching examination of retail venues who offer both conventional and organic produce provides insight into store pricing effects on household consumption of organic fresh produce.

While Thompson and Kidwell (1998) were able to capture the effects that specialty stores had on purchases rather than a food cooperative, the ability to expand to a wider range of stores is possible with the data in this analysis.

2.2.4 Increased Household Composition Analysis

Few studies have examined the possible effects that presence of children and their respective age would have on household organic purchase activity. While Thompson and Kidwell (1998) find that the number of children in the household increases the probability that a household will choose organic produce, no studies report age effects of children.

Due to the high detail level of Homescan^{TM} household data, information on the number of children in age brackets of under 6 years, 6-12 years, and 13-17 years is provided and allows for increased analysis of children's age effects on household organic purchase volumes.

Furthermore, Homescan^{TM} household data allows the researcher to study consumption patterns of larger households without children. The effects of additional members living within the household excluding the female and male head (if present) provides more detailed analysis of household composition effects.

Clearly potential exists for insightful economic analysis based on the breadth of scanner data available. The researcher must, however, fully develop the appropriate theoretical foundations in order to specify reasonable econometric models to achieve plausible results. The next chapter will address the economic theory behind the empirical models that will be used later in this study.

CHAPTER 3

THEORETICAL MODEL

Any study which aims to model organic purchase activity amongst a diversified data set of produce purchases will logically encounter problems with high frequencies of zero expenditures within the organic niche of the industry. Any subsequent estimates based on that data without accounting for a high number zero expenditures will consequently yield biased and inconsistent coefficient estimates.

Wales and Woodland (1983) and Tobin (1958) derive models with an economic interpretation of zero expenditure as well as an appropriate method to specify the econometric model. In order to construct a model of consumer demand which handles significant zero observations, consider the demand theory based on the Kuhn and Tucker (1951) conditions for utility maximization and non-random preference selection by Tobin (1958) for the case of a single demand equation[1].

Both microeconomic models are extended to accommodate the case of panel data. This type of data helps to control for individual heterogeneity and includes an individual unobserved effect in the model. Time-series and cross-sectional studies not controlling for this heterogeneity run the risk of obtaining biased coefficient estimates. The biggest advantage of panel estimation is the ability to control for unobserved heterogeneity.

3.1 Kuhn-Tucker Non-Negativity Constraints for Corner Solutions

Wales and Woodland (1983) consider the problem of estimating consumer demand for samples that contain a high proportion of observations with zero consumption of one or more goods. Their econometric model follows that of maximizing a random

[1]See "Single Equation vs. Demand System" Section as to why a single demand equation is of interest.

direct utility function subject to budget constraints whereas Kuhn-Tucker conditions determine the set of non-consumed goods.

Consider traditional theory of consumer demand which assumes that an individual will maximize utility subject to their budget constraint. The utility function is assumed to be a continuously differentiable, quasi-concave, increasing function. $G(x)$ denotes an individual's utility function over a set of non-negative quantities $x = (x_1, ..., x_m)$ which satisfy the budget constraint $v^T x \leq 1$ where v represents a vector of normalized prices[2], $v = (v_1, ..., v_m)^T > 0$. The maximization problem is:

$$H(v) \equiv \max_x \{G(x) : v^T x \leq 1, x \geq 0\} \tag{3.1}$$

The necessary and sufficient Kuhn-Tucker conditions for a solution are:

$$\begin{aligned} G_i(x) - \lambda v_i \leq 0 \leq x_i, \quad i = 1, ..., M, \\ v^T x - 1 \leq 0 \leq \lambda, \end{aligned} \tag{3.2}$$

where λ is the Lagrange multiplier associated with the budget constraint and M is the number goods consumed. When given the double inequality constraints in the above conditions, multiplied together they yield zero. The first condition in (3.2), therefore can be written as $x_i(G_i(x) - \lambda v_i) = 0$ which is the necessary condition for optimality. The second condition which can be written as $\lambda(v^T x - 1) = 0$ is the complementary slackness condition.

Since the utility function, $G(x)$, is assumed to be increasing as mentioned, the consumer will spend all his or her income[3]. Therefore λ is assumed to be positive and at least one good will be consumed. Given the Kuhn-Tucker conditions in (3.2), λ can be written as $\lambda = G_1(x)/v_1$. Given the Kuhn-Tucker conditions, (3.1) can be rewritten as:

$$v_1 G_i(x) - v_i G_1(x) \leq 0 \leq x_i, \quad i = 2, ..., M,$$

[2]Normalized prices designate $v_i = p_i/m$ where p_i is the price of good i and m is income.
[3]No saving is assumed in this model

$$v^T x = 1 \tag{3.3}$$

According to (3.3), if $x_i > 0$ then $v_1 G_i(x) - v_i G_1(x) = 0$ as the marginal rate of substitution between goods i and 1 along the indifference curve at the solution is equal to the price ratio, $G_i(x)/G_1(x) = v_i/v_1$. If $x_i = 0$ as is pertinent for this study, the marginal rate of substitution between goods i and 1 usually is less than the price ratio ($MRS_{1,i} < -\frac{v_i}{v_1}$) given the Kuhn Tucker conditions.

The Kuhn-Tucker conditions above explain the non-use of goods for which the market price is too high. The consumer sits at a corner solution of zero in this situation. Such theory can be later used to explain high frequencies of non-purchases of organic produce given the high price ratio of organic to conventional fresh produce.

Adding a random component (u_i) to marginal utilities $G_i(x)$ to handle differing tastes among a heterogenous random population can be written as:

$$G_i(x, u_i) = \bar{G}_i(x) + u_i, \quad i = 1, ..., M. \tag{3.4}$$

The random utility function can therefore be written as $G_i(x, u_i) = \bar{G}_i(x) + u^T x$ where u is a vector of random components for the sample, $u = (u_1, ..., u_M)$. Applying the random marginal utility $G_i(x, u_i)$ to (3.3) will yield:

$$(v_1 u_i - v_i u_1) + [v_1 \bar{G}_i(x) - v_i \bar{G}_1(x)] \leq 0 \leq x_i, \quad i = 2, ..., M,$$

$$v^T x = 1 \tag{3.5}$$

Because the u vector is unknown to the researcher and considered a random draw, $G(x, u)$ has an incorporated stochastic element. From the consumers' point of view, the utility function $G(x, u)$ is nonstochastic since their u vector is known to them, thus their optimal consumption vector is obtained by solving (3.5) for x. This study utilizes panel estimation in an effort to capture part of u in an econometric model whereas the stochastic element is decomposed into two parts to control for heterogeneity.

3.2 The Tobin Interpretation

Tobin (1958) derived an econometric model in which household expenditures on various categories of goods differ with income levels. For goods categorized as "luxuries" for example, high populations of zero expenditure can be observed in household studies. The result is that a straight line is not appropriate in representing the Engel curve for differing household income levels where a corner solution will occur for some households. In assuming that the consumer attempts to maximize a utility function subject to his or her budget constraint, the determined shares are assumed to lie between zero and infinity. Due to random disturbances such as errors of maximization by the consumer, errors of measurement of the observed shares, and other factors that influence the consumer's decisions, the observed shares will not coincide with the determined shares of the model. The typical solution would be to add a normally distributed stochastic error element to the deterministic shares implying that these shares follow a normal distribution. Tobin points out that there is nothing in this formulation to ensure that these values lie between zero and unity and a proper econometric model should be considered in order to avoid biased estimates.

3.3 Econometric Model Specification

In the econometric specification of the microeconomic model, the error term given, ε can be decomposed as to reflect $\varepsilon = \mu + \epsilon$ where μ is the individual random effect term which corresponds to an unobserved effect that handles econometric problems of heterogeneity while ϵ is the remainder error term. The model must accommodate left censoring at zero as prescribed by the microeconomic theory above. In the context of panel data, consider organic quantity purchased as a function of organic and conventional prices, various time-varying and time-invariant demographic traits, time-varying income categories, seasonality dummies, the time-invariant individual unobserved effect, and the error term.

$$ORGQ_{it} = f(ORGPrice_{it}, CONPrice_{it}, Demo_{it}, Demo_i, Inc_{it}, Seasonal_{it}, \mu_i, \epsilon_{it})$$

$$i = 1, 2, ...N \quad t = 1, 2, ...T$$

where i denotes households and t represents time periods.

3.4 Single Equation vs. Demand System: Is a System Interesting Here?

Over the last 3 decades, a wealth of microeconomic and econometric literature on demand systems has accumulated. Systems have a strong basis in microeconomic theory as homogeneity, symmetry, and adding-up restrictions may be satisfied and systems are useful for the estimation of the demands of several goods of the same industry. However, even though this study examines several organic and conventional goods, the decision was made to estimate several single demand equations as apposed to a demand system. Undoubtedly the question as to why a demand system should not be used to estimate price elasticities and demand effects begs to be asked. In this case, the answer is two-fold.

First, due to the panel nature of the data in this study and the high population of left-censored organic purchases in this case, computation and programming of a demand system in a panel context would be quite difficult.

Second, though it is possible to implement a system for this study, the results would not be of great interest. The pivotal application of this study is to examine the effects of demand for organic produce as a function of conventional price, demographics and household attributes and a system is not necessary here. In terms of elasticities, clearly own-price and cross-price elasticities of the same type of organic and conventional produce are interesting whereas a system that estimates elasticities across multiple conventional and organic products is not. Estimating cross-price elasticities, for example, for organic peppers and conventional tomatoes is not very interesting per se and it is not clear how policy makers of industry participants would use such information.

CHAPTER 4

THE DATA

4.1 Why Panel Data?

A panel, or longitudinal, data set is comprised of a sample of individuals who are observed over time, and therefore provide repeated observations on each individual in the sample throughout a given time horizon. Depending on the nature of how the panel was generated, the data could be balanced or unbalanced. A balanced panel contains an equal number of temporal observations for each individual. Unbalanced panels, or also referred to as incomplete panels, consist of data that do not have equal time periods across each individual. While both balanced and unbalanced panels are feasible for estimation, with incomplete panels the researcher must be cautious to ensure that missing cross-sectional observations are random as to avoid systematic errors that could effect results.

There are many advantages that panel data models offer over more common cross-sectional or time series samples. The more noticeable advantage is that panel analysis endows regressions with both a spatial and temporal dimension. With repeated observations on enough cross-sections, panel analysis permits the researcher to study the dynamics of change even with a short time series. The ability to utilize a specific individual's history within a model allows the researcher to estimate a specific item of interest in different subintervals of the life cycle (Hsiao, 2003).

4.1.1 Avoidance of Heterogeneity Bias

Because most economic data do not come from simple controlled experiments but from complicated processes in everyday life, it is almost impossible to include all the factors affecting the outcome for all individuals in a model specification. Since the researcher is mainly interested in capturing the essential forces affecting an outcome, typically the factors which cannot be measured or are believed to be insignificant are left out of the model. If important individual-specific factors are left out of the model, the usual assumption is that the estimated parameter vector is identical for all individuals. Often this is not a realistic assumption which can lead to parameter heterogeneity. In this case, including or omitting variables in the model would greatly affect parameter estimates. If heterogeneity is present and not accounted for, the parameter estimates are inconsistent(Hsiao, 2005).

Panel analysis measures not only the effects that observable variables have on the dependent variable, but also can account for the effects of relevant unobservable influences. How these unobservable influences are incorporated into the model depends upon whether a fixed-effect (FE) or random-effects model (RE) is used in estimation. In the FE model, the unobservable effect is an individual-specific intercept shifter within the model that accounts for unobserved heterogeneity of that individual. In the RE model, the unobservable effects across cross-section units are assumed to be best characterized as random variables.

Especially dealing with large, complicated data files like the one that will be utilized later in the analysis, the issue of heterogeneity bias is unquestionably important. By explicitly modeling the unobserved heterogeneity, the problem of inconsistent estimates that would come from alternative models may be avoided.

4.2 The ACNielsen Homescan^{TM} Panel

4.2.1 Panel History

In the early 1970s, scanner technology at the retail level began to emerge. In 1974, the first checkout scanner system was installed in the US while Portable Data Capture Units (PDCUs) surfaced around the same time (Baron and Lock, 1995). In the early years, PDCUs took the form of light-pen barcode readers and were used for stock reordering purposes. After years of problematic IT, computing, and data storage setbacks, scanner technology, especially PDCUs began to improve in the late 1980s. In 1987, ACNielsen's first household panel, National Electronic Household Panel (NEHP), was developed in which household participants used PDCUs to record all UPC-coded purchases.

As ACNielsen's NEHP panel expanded, so did the panel's capacity for increased households and breadth of recording and data storage. Currently, Homescan^{TM}, formerly part of NEHP, is a multi-outlet panel that captures all consumer package goods purchase information, as well as non-UPC coded random weight perishable products. Store, price, quantity, and promotion information is recorded with the products scanned and the information is transmitted weekly to the company. Participants use a UPC scanner device to scan or input perishable product information which is subsequently stored within the device and transferred to ACNielsen at a later date.

Currently multiple collaborative efforts are underway with the Economic Research Service (ERS) and ACNielsen to provide scanner data for analysis. Such efforts have spawned a great deal of study in economics and marketing.

4.2.2 Data Generation Process

Households examined for this study consist of what is referred to by ACNielsen as the Homescan^{TM} "random-weight" panel. The panel consists of approximately 12,000 households where participants scan both fixed-weight products (items with a universal product code, UPC) and random-weight products (meat and poultry, fruit and vegetables). Once selected by ACNielsen[1], a participating household is given a scanner device capable of recording purchase information that is kept at home. Each time a household member shops and returns home, he or she uses the scanner to record each item purchased. For items without a universal product code (UPC), such as fresh produce, respondents are given a vocabulary code book to identify the item, then they provide the price and weight information. Just like fixed-weight products, store, brand, date, and channel information is subsequently stored within the memory of the scanner device. Once a week participants send their information to ACNielsen via telephone lines (National Research Council of the National

[1]See 4.2.3 for ACNielsen's representative selection criteria

Academies, 2005). Besides purchase history, ACNielsen subsequently maintains an active database on the specific demographic characteristics of the household.

The Homescan^{TM} panelists that transmit data on scanned purchases once a week receive points that can be redeemed for prizes for each data transmittal. Typical response rates for the Homescan^{TM} panel are around 85 percent (National Research Council of the National Academies, 2005).

4.2.3 Selection Criteria

The Homescan^{TM} "random-weight" panel is a stratified random sample based on both demographic and geographic targets. The selection is done to ensure that the sample proportionately matches the US Census. The weighting of the number of households in the panel reflects the demographic population distribution described by the strata. Membership in the Homescan^{TM} program has been through random invitations in the mail or e-mail. Acceptance into the program is based on the current needs of the company. Attempts are made by ACNielsen to continually balance their panel to represent the overall population of the mainland US.

For some demographic segments, attaining a representative sample has been historically difficult. ACNielsen has reported having trouble recruiting some groups to participate in the survey, specifically, young single adults, people in low-income households, and hispanics (National Research Council of the National Academies, 2005).

As a check on representativeness, the data used in this study are compared with US Census figures. Specifically, the random-weight panel from 1998 to 2001 obtained from the Economic Research Service is compared to the US Census 2000 to examine the representativeness of the sample to the US population. Table 4.1 shows the differences between the ACNielsen panel and the US Census 2000 in terms of composition (US Census Bureau, 2001). The figures verify that white households were over-represented while hispanics and Asians were under-represented. Also, the Homescan^{TM} sample households tend to have higher incomes than the representative US and had a higher likelihood of being married. While almost no data samples will prove to be exactly representative to a large diverse population such as the US population, such differences should be considered when analyzing empirical results.

Table 4.1: Comparison of Homescan^{TM} Sample with the US Census 2000

Stat	Sample[a]	US Census Figure
Racial Population Demographics		
White	80.7%	75.1%
Black	12.3%	12.3%
Asian	2.2%	3.6%
Hispanic	7.8%	12.5%
Household Income Demographics		
\$0 to \$19,999	13.1%	22.1%
\$20,000 to \$39,999	28.5%	25.3%
\$40,000 to \$59,999	26.3%	19.7%
\$60,000 to \$99,999	25.0%	20.6%
\$100,000 or over	7.9%	12.3%
Household Demographics		
Married	62.0%	54.3%
Divorced	14.7%	9.7%
Average Household Size	2.67	2.59
Sample Size	13,790	

[a]Unique Households in the Sample 1998-2001

4.3 Data Manipulation

The data from ACNielsen contain purchase activity information which is comprised of two components. The first component includes demographic-specific variables that are for the most part, time-invariant across the panel. Clearly, some demographical variables would be time-varying throughout the time horizon of the panel;

however, many that denote ethnicity, race, and location remain constant throughout. Also, because many of these demographical identifiers are recorded as categorical and not as continuous variables, little change exists within specific households in the panel. These demographic variables consist of income, age, and household composition elements that correspond to each household participating in the study. Each household is denoted with a unique identification number in which the second component of the data set, the purchase activity data, may be merged.

The variables in the purchase activity data set consist of specific store, price, quantity, product size, and promotion information relevant to all the corresponding purchases made by a specific household through time. Date of purchase information is also recorded as well as abbreviated product descriptions, brand, product type, and channel information.

4.3.1 Differentiating Organic vs. Non-Organic Fresh Purchases

Within the product description variable, specifics about the item purchased are represented. One acronym used within this description denotes the fact that an item is either organic or non-organic by description. In matching items with identical names with the exception of a changed organic (ORG) or non-organic (N-ORG) acronym, like items may be differentiated. Fresh purchases such as 'RW FRT BANANA ORG LWAS' and 'RW FRT BANANA N-ORG LWAS ' were of particular interest due to their high incidences of purchase activity within the panel. In aggregating non-organic and organic price and purchase activity for each household throughout time, a panel was assembled.

Given the ACNielsen random-weight data, specific produce items were first selected for analysis due to their high purchase frequencies. Three raw fresh items were selected due to their uniform descriptions between organic and conventional types (besides the ORG/N-ORG designation) which would facilitate appropriate analysis.

4.3.2 Aggregation of Data and the Price Proxies

The aggregation process entailed summing up both organic and conventional purchases occurring within the given time period in order to obtain total weight (lbs) and count values of purchases. For counts, the ACNielsen scanner device maintains the "event" date information for each purchase even though the data is transmitted weekly.

Price information pertaining to a specific item was averaged across that specific time period[2]. Because households infrequently buy both conventional and organic fresh items of the same produce type during the month, a proxy price value was constructed. Infrequent purchases mean, for example, that if a household buys conventional bananas but no organic bananas in a given week, no unit price for organic bananas is recorded; the organic banana price in that week would be missing. Without a price observation in that week, the other price and quantity information from that week could not be used for estimation. Hence, we formulate the following method to obtain a price proxy.

First, purchases were segmented by the geographic region designated by the data. In the Homescan^{TM} data, the variable MarketId designates 52 specific US markets in which the sample households reside. Subsequent purchases made at local store chains or produce markets within those geographic regions by other panel members were used to fill missing price values. Average price values for the given month from a specific venue in a specific market area were used to fill missing organic or conventional price information (see figure 4.1 for an example).

By utilizing local chain price information based on other household purchases in the same area during the same month, the problem of missing price values within the data is avoided. Without a price proxy value for non-purchases, few households would remain for estimation and nonrandom sample problems would certainly arise.

[2]Weight values are modeled with a t time period of one month. Count values are modeled with a t time period of one quarter because of relatively infrequent purchases.

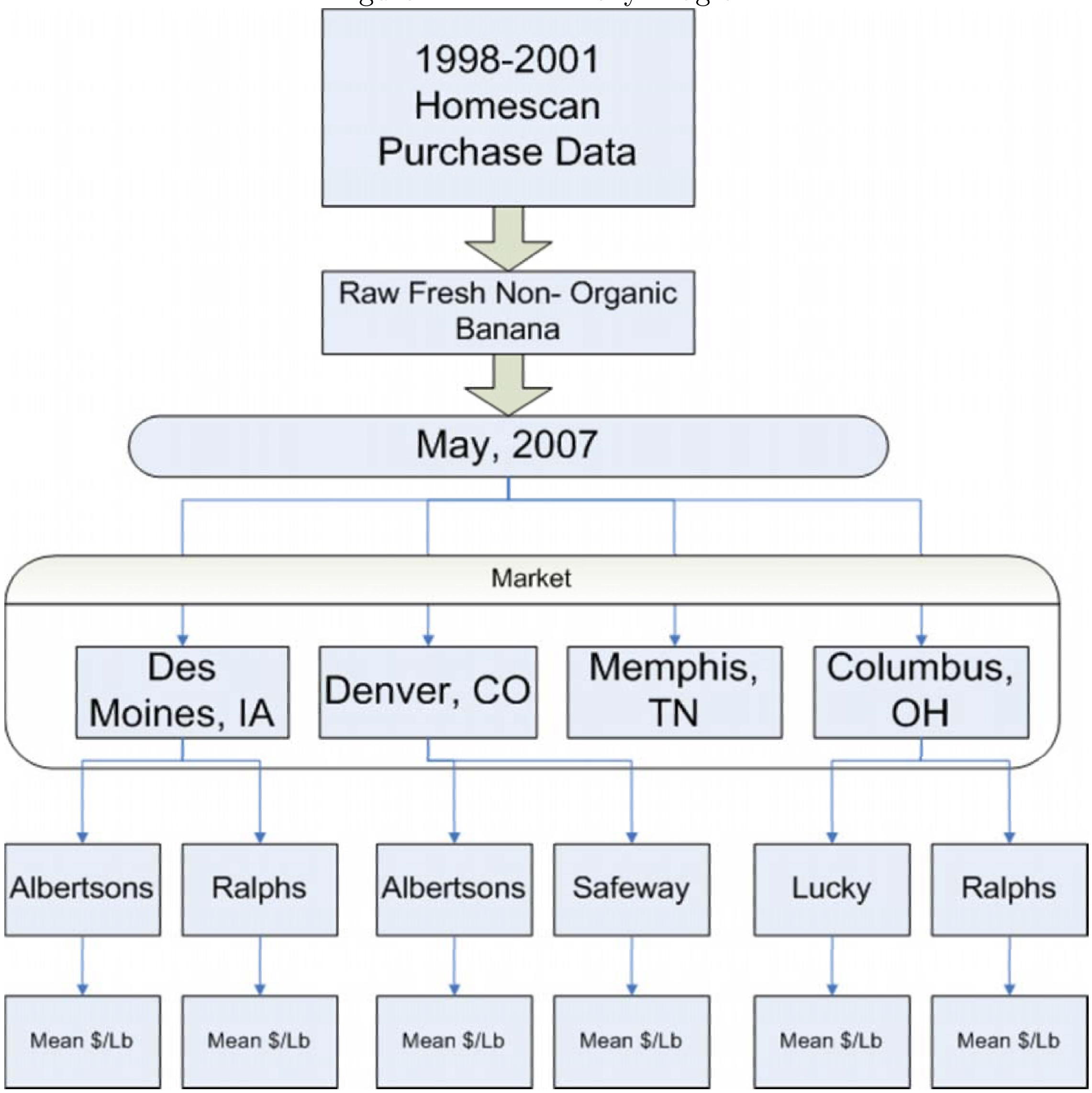
4.3.3 Panels Used for Estimation

The panels constructed for estimation of organic produce quantity consist of the total weight of three fresh produce items -bananas, peppers, and tomatoes- purchased by a household in a given month. An addition panel consists of counts of organic items purchased by a household in a given quarter. Because unbalanced panels are feasible for estimation, an observation is only recorded for a household if a fresh item, either organic or conventional, is purchased during a given time period between January 1998 and December 2001. Unless a household has made both organic and conventional purchases of the same type (i.e. fresh raw tomatoes) during a given time period, an average price value for the item not purchased will be missing from the data. To facilitate price comparison, a price proxy is introduced to the panel as mentioned.

Note on Availability

Unquestionably, availability of organic or conventional items at a given store is important to the validity of price comparison between organic and conventional items. If a consumer that is shopping at a store does not have a choice between conventional or organic items, price comparison at that store is not relevant. To reduce the effect that no availability would have on empirical results, only purchase information based on store chains and market names in a given geographic area (i.e. Columbus, OH) that displayed purchase activity of both organic and conventional items of a specific type were included in the panel.

Figure 4.1: Price Proxy Diagram



CHAPTER 5

SUMMARY STATISTICS

5.1 Overall Organic Trends

Based on overall ACNielsen Homescan^{TM} panel purchases, Figure 5.1 represents the volume of organic sales throughout time. The organic items represented are any fresh or prepackaged fruit or vegetable produce items marked as organic. Notice the seasonality trend towards higher sales in the summer months and lower volumes during the winter. Similarly, Figure 5.2 shows a general availability measure over time where the number of different stores that sold organic products are shown. Figure 5.3 shows the volumes of purchases by the four classified regions in the US (East, Central, South, West). Between 1998 and 2001, the South region made the most organic food purchases on the aggregate followed closely by the West region of the US.

5.2 Samples for Estimation

Tables 5.1 and 5.2 include descriptions of variables used in empirical analysis. Tables 5.3 and 5.4 include the means and standard deviations of variables used for estimation. Methods of obtaining these samples from the overall panel were discussed in Chapter 4 and are specific to three specific types of raw fresh produce.

As discussed in Chapter 1, organic produce commands higher price premiums than those of similar conventional produce. Figure 5.4 represents average price premiums of organic produce within each estimation sample across the 4 year time period. Figure 5.5 represents the percentage of organic purchases by produce type across the same period. Notice the existence of a seasonality trend on the percentage of organic purchases by type. Percentage of organic tomato purchases relative to the other organics seem to increase during the summer months. Figure 5.6 displays the percentage of overall (conventional and organic) purchases by produce type that are organic and would be non-zero (uncensored) in the model. Notice that bananas have a much smaller percentage of organic relative to conventional purchases while organic tomatoes hold a much higher share of purchases.

Figure 5.7 shows the percentage of conventional and organic purchases across racial and / or ethnicity types[1]. In comparing the racial summary statistics to Table 4.1, the minority groups, specifically Blacks, seem to consume smaller amounts of produce per capita while Whites consume a higher per capita level. The minority groups, however, consume a higher percentage of organic items when they do purchase produce.

Figure 5.8 shows the distribution of the estimated Poisson RE model's dependent variable. Quarterly household organic produce purchases are count values which inspires the count panel model. Figure 5.8 includes a histogram and a fitted normal density curve of quarterly purchase counts.

[1]These are ACNielsen designated racial/ethnicity classifications.

Table 5.1: Variables Used in Analysis

Variable	Definition
Organic Purchases	Number of Monthly / Quarterly Purchases of Organic Raw Fresh Produce
Organic Quantity	Weight in lbs of Monthly Fresh Organic Produce Purchases
Organic Price	Average Monthly $/lb of Raw Fresh Produce Marked as Organic (Nominal)
Conventional Price	Average Monthly $/lb of Raw Fresh Produce Marked as Non-Organic (Nominal)
Household Size	Number of Individuals in the Household
# Children Under 6	Number of Children in the Household Under 6 Years
# Children 6-12	Number of Children in the Household Between 6 and 12 Years
# Children 13-17	Number of Adolescents in the Household Between 13 and 17 Years
Extra Members[a]	Number of Adults in the Household Excluding Male and Female Heads
ageM	A Count Value for every 5 years of age of the Male Household Head has starting at 25 years $(1 =< 25, 2 = 25-29, 3 = 30-34, ...)$
ageF	A Count Value for every 5 years of age of the Female Household Head has starting at 25 years $(1 =< 25, 2 = 25-29, 3 = 30-34, ...)$
No Kids	Binary Variable for a Household with no Children Under 18
noHeadF	Binary Variable for a Household with no Female Head
noHeadM	Binary Variable for a Household with no Male Head
Dad FT	Binary Variable for Male Head with +35 hours/week job
Mom FT	Binary Variable for Female Head with +35 hours/week job
Dad College	Binary Variable for Male Head with a College Degree
Mom College	Binary Variable for Female Head with a College Degree
dadGrad	Binary Variable for Male Head with a Post College Degree
momGrad	Binary Variable for Female Head with a Post College Degree

[a]Extra members of households with no children

Table 5.2: Variables Used in Analysis-Continued

Variable	Definition
Dad Pro	Binary Variable for Male Head with "professional" occupation
Mom Pro	Binary Variable for Female Head with "professional" occupation
Married	Binary Variable for an Existing Marriage
Divorced	Binary Variable for a Divorced Household
Hispanic	Binary Variable for the Household being "Hispanic"
White	Binary Variable for the Household being "White"
Black	Binary Variable for the Household being "Black"
Oriental	Binary Variable for the Household being "Oriental"
Inc20to40	Binary Variable for Annual Household Income of $20,000 to $39,999
Inc40to60	Binary Variable for Annual Household Income of $40,000 to $59,999
Inc60to70	Binary Variable for Annual Household Income of $60,000 to $69,999
Inc70to100	Binary Variable for Annual Household Income of $70,000 to $99,999
HighInc	Binary Variable for Annual Household Income of $100,000 or over
Q1-Q3	Seasonal Dummies for Annual Quarter 1-3

Figure 5.1: Number of Homescan Purchases of Any Products Labeled as Organic from 1998-2001

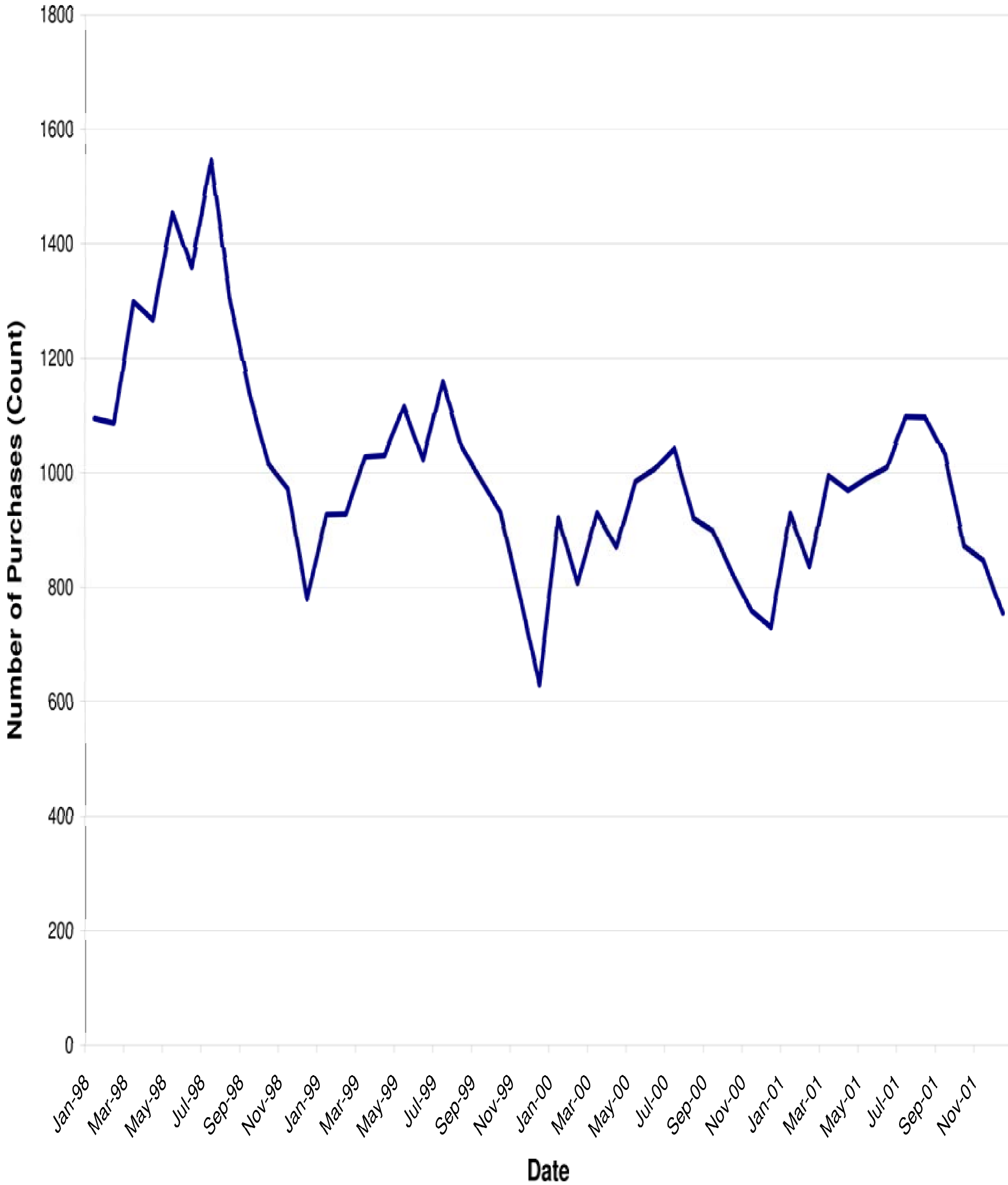


Figure 5.2: Number of Unique Stores where Organic Products Available were Purchased 1998-2001

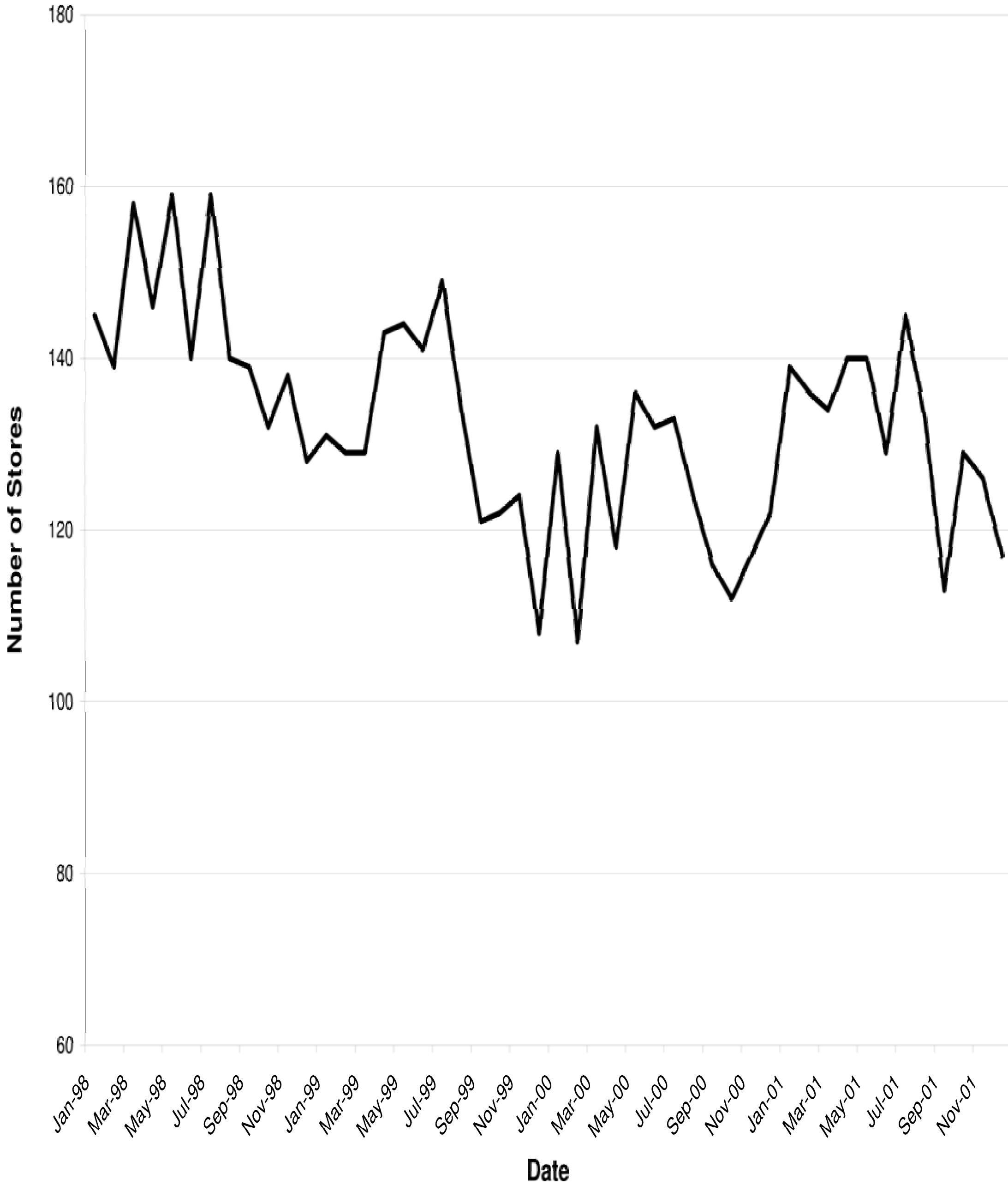


Figure 5.3: Number of Homescan Purchases of Organic Products by Region from 1998-2001

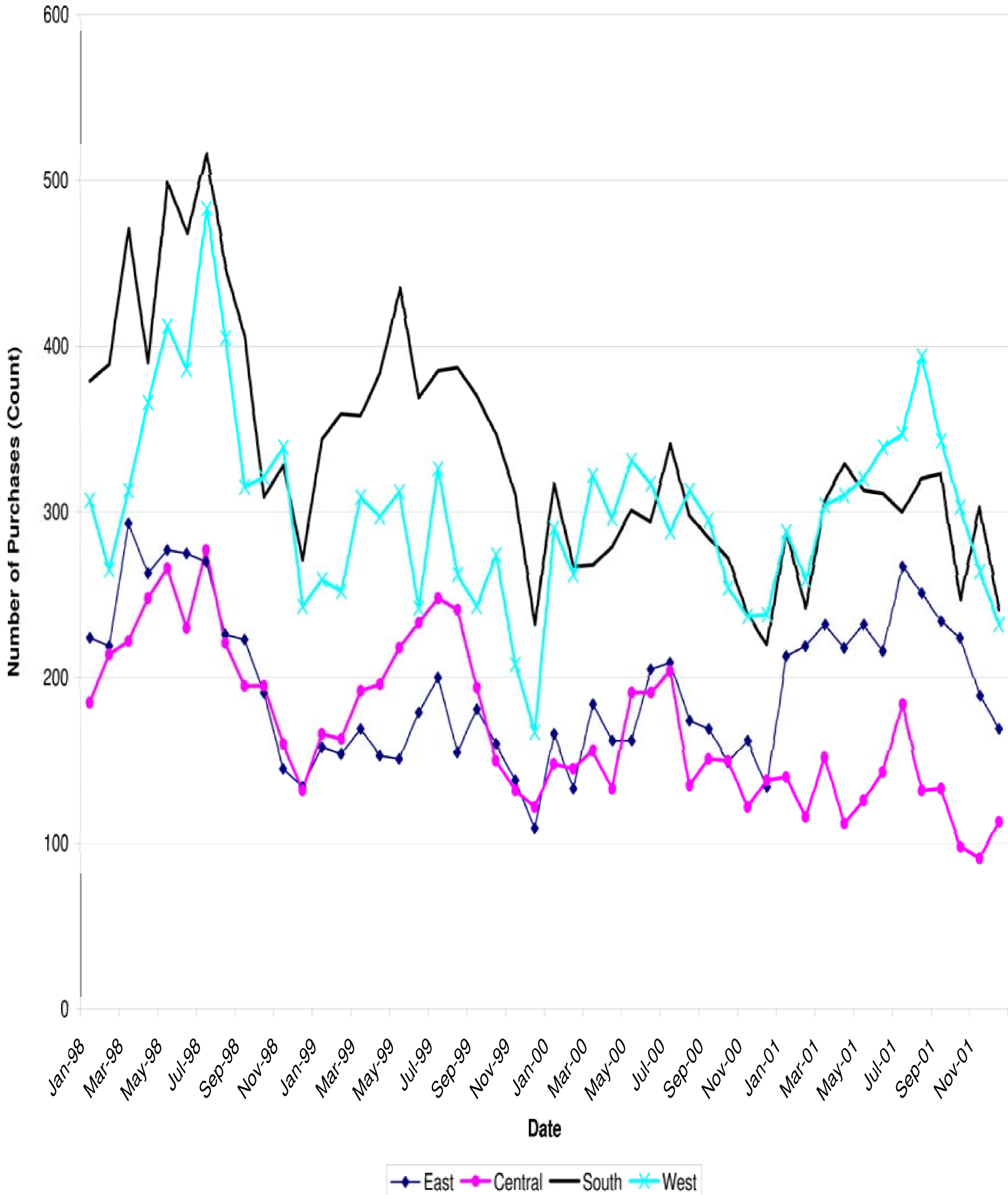


Table 5.3: Variable Means and Standard Deviations

	Bananas[a]		Peppers[a]		Tomatoes[a]		Counts[b]
	Overall	Non-Zero[c]	Overall	Non-Zero[c]	Overall	Non-Zero[c]	(Poisson)
Organic Purchases	0.13	1.36	0.21	1.41	0.26	1.23	0.32
	(0.50)	(1.01)	(0.67)	(1.16)	(0.70)	(1.06)	(1.03)
Organic Quantity	0.23	2.43	0.14	0.91	0.30	1.45	
	(0.98)	(2.25)	(0.44)	(0.79)	(0.75)	(1.01)	
Organic Price	0.73	0.72	1.71	1.48	1.66	1.59	1.16
	(0.54)	(0.55)	(1.10)	(1.06)	(0.96)	(0.95)	(0.95)
Conventional Price	0.62	0.64	1.33	1.53	1.56	1.60	1.01
	(0.35)	(0.34)	(0.79)	(0.76)	(0.89)	(0.81)	(0.76)
Household Size	2.65	3.03	2.70	2.97	2.67	2.60	2.64
	(1.30)	(1.51)	(1.32)	(1.46)	(1.34)	(1.40)	(1.33)
# Children Under 6	0.17	0.24	0.15	0.20	0.15	0.14	0.17
	(0.54)	(0.69)	(0.51)	(0.67)	(0.52)	(0.57)	(0.54)
# Children 6-12	0.26	0.39	0.26	0.29	0.25	0.24	0.27
	(0.64)	(0.82)	(0.64)	(0.70)	(0.66)	(0.70)	(0.67)
# Children 13-17	0.24	0.34	0.27	0.36	0.26	0.27	0.24
	(0.65)	(0.78)	(0.68)	(0.86)	(0.69)	(0.71)	(0.65)
Extra Members	0.22	0.31	0.23	0.33	0.25	0.24	0.23
	(0.58)	(0.72)	(0.58)	(0.74)	(0.63)	(0.62)	(0.59)
ageF	5.89	5.37	5.92	5.45	5.93	5.62	5.83
	(2.59)	(2.70)	(2.40)	(2.26)	(2.55)	(2.62)	(2.58)
ageM	5.44	5.24	5.52	5.17	5.52	5.18	5.24
	(3.11)	(2.95)	(2.95)	(2.91)	(3.09)	(3.12)	(3.16)
nokids	0.69	0.59	0.70	0.65	0.70	0.71	0.69
	(0.46)	(0.49)	(0.45)	(0.48)	(0.46)	(0.45)	(0.46)
noHeadF	0.07	0.10	0.06	0.05	0.07	0.09	0.07
	(0.26)	(0.30)	(0.23)	(0.21)	(0.26)	(0.28)	(0.26)
noHeadM	0.17	0.15	0.15	0.17	0.17	0.20	0.19
	(0.38)	(0.36)	(0.36)	(0.37)	(0.28)	(0.40)	(0.40)
dadFT	0.57	0.59	0.61	0.62	0.57	0.56	0.56
	(0.50)	(0.49)	(0.49)	(0.49)	(0.50)	(0.50)	(0.50)
momFT	0.41	0.42	0.43	0.51	0.42	0.45	0.42
	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)
dadColl	0.42	0.37	0.45	0.38	0.41	0.42	0.40
	(0.49)	(0.48)	(0.50)	(0.49)	(0.49)	(0.49)	(0.49)
momColl	0.40	0.33	0.42	0.41	0.38	0.39	0.39
	(0.49)	(0.47)	(0.50)	(0.49)	(0.49)	(0.49)	(0.49)
dadGrad	0.16	0.11	0.19	0.16	0.16	0.15	0.15
	(0.37)	(0.31)	(0.39)	(0.36)	(0.37)	(0.35)	(0.35)
momGrad	0.12	0.10	0.16	0.17	0.12	0.14	0.12
	(0.33)	(0.30)	(0.36)	(0.38)	(0.33)	(0.35)	(0.33)

Standard Deviations in parentheses.
Based on samples used for estimation.
[a] means measured on the monthly level.
[b] means measured on the quarterly level.
[c] Means when organic quantities are positive.

Table 5.4: Variable Means and Standard Deviations -Continued

	Bananas[a]		Peppers[a]		Tomatoes[a]		Counts[b]
	Overall	Non-Zero[c]	Overall	Non-Zero[c]	Overall	Non-Zero[c]	(Poisson)
dadPro	0.27	0.23	0.30	0.24	0.26	0.24	0.26
	(0.45)	(0.42)	(0.46)	(0.43)	(0.44)	(0.43)	(0.44)
momPro	0.24	0.21	0.26	0.26	0.23	0.22	0.24
	(0.43)	(0.41)	(0.44)	(0.44)	(0.42)	(0.41)	(0.43)
married	0.73	0.73	0.76	0.75	0.74	0.68	0.71
	(0.44)	(0.44)	(0.43)	(0.43)	(0.44)	(0.46)	(0.45)
divorced	0.09	0.09	0.09	0.09	0.09	0.11	0.11
	(0.29)	(0.29)	(0.28)	(0.29)	(0.29)	(0.32)	(0.31)
Hispanic	0.09	0.15	0.09	0.16	0.11	0.11	0.09
	(0.27)	(0.36)	(0.29)	(0.37)	(0.31)	(0.31)	(0.29)
White	0.85	0.72	0.86	0.72	0.87	0.84	0.84
	(0.36)	(0.45)	(0.35)	(0.45)	(0.34)	(0.37)	(0.37)
Black	0.09	0.17	0.08	0.17	0.07	0.09	0.09
	(0.28)	(0.37)	(0.27)	(0.38)	(0.26)	(0.29)	(0.29)
Oriental	0.02	0.02	0.02	0.02	0.02	0.01	0.02
	(0.15)	(0.16)	(0.13)	(0.15)	(0.14)	(0.10)	(0.15)
Inc20to40	0.23	0.28	0.19	0.19	0.21	0.22	0.23
	(0.42)	(0.44)	(0.39)	(0.40)	(0.40)	(0.42)	(0.42)
Inc40to60	0.27	0.26	0.27	0.26	0.27	0.24	0.27
	(0.44)	(0.44)	(0.45)	(0.44)	(0.45)	(0.43)	(0.45)
Inc60to70	0.12	0.10	0.11	0.13	0.12	0.10	0.12
	(0.32)	(0.29)	(0.32)	(0.33)	(0.33)	(0.31)	(0.32)
Inc70to100	0.20	0.18	0.21	0.21	0.21	0.22	0.19
	(0.40)	(0.39)	(0.40)	(0.41)	(0.40)	(0.41)	(0.40)
HighInc	0.13	0.09	0.17	0.14	0.13	0.13	0.12
	(0.34)	(0.28)	(0.38)	(0.34)	(0.34)	(0.33)	(0.33)
Q1	0.26	0.27	0.28	0.28	0.22	0.21	0.25
	(0.44)	(0.45)	(0.45)	(0.45)	(0.41)	(0.41)	(0.44)
Q2	0.26	0.28	0.26	0.25	0.32	0.32	0.27
	(0.44)	(0.45)	(0.44)	(0.43)	(0.47)	(0.47)	(0.44)
Q3	0.24	0.24	0.26	0.27	0.29	0.29	0.25
	(0.43)	(0.43)	(0.44)	(0.44)	(0.45)	(0.45)	(0.44)
nT	23,556	2,189	3,987	591	7,990	1,666	20,020
Households	4,285		1,546		2,527		5,181

Standard Deviations in parentheses.
Based on samples used for estimation.
[a] means measured on the monthly level.
[b] means measured on the quarterly level.
[c] Means when organic quantities are positive.

Figure 5.4: Average Price Premiums of Organic vs. Conventional Produce Through Time (Estimation Samples)

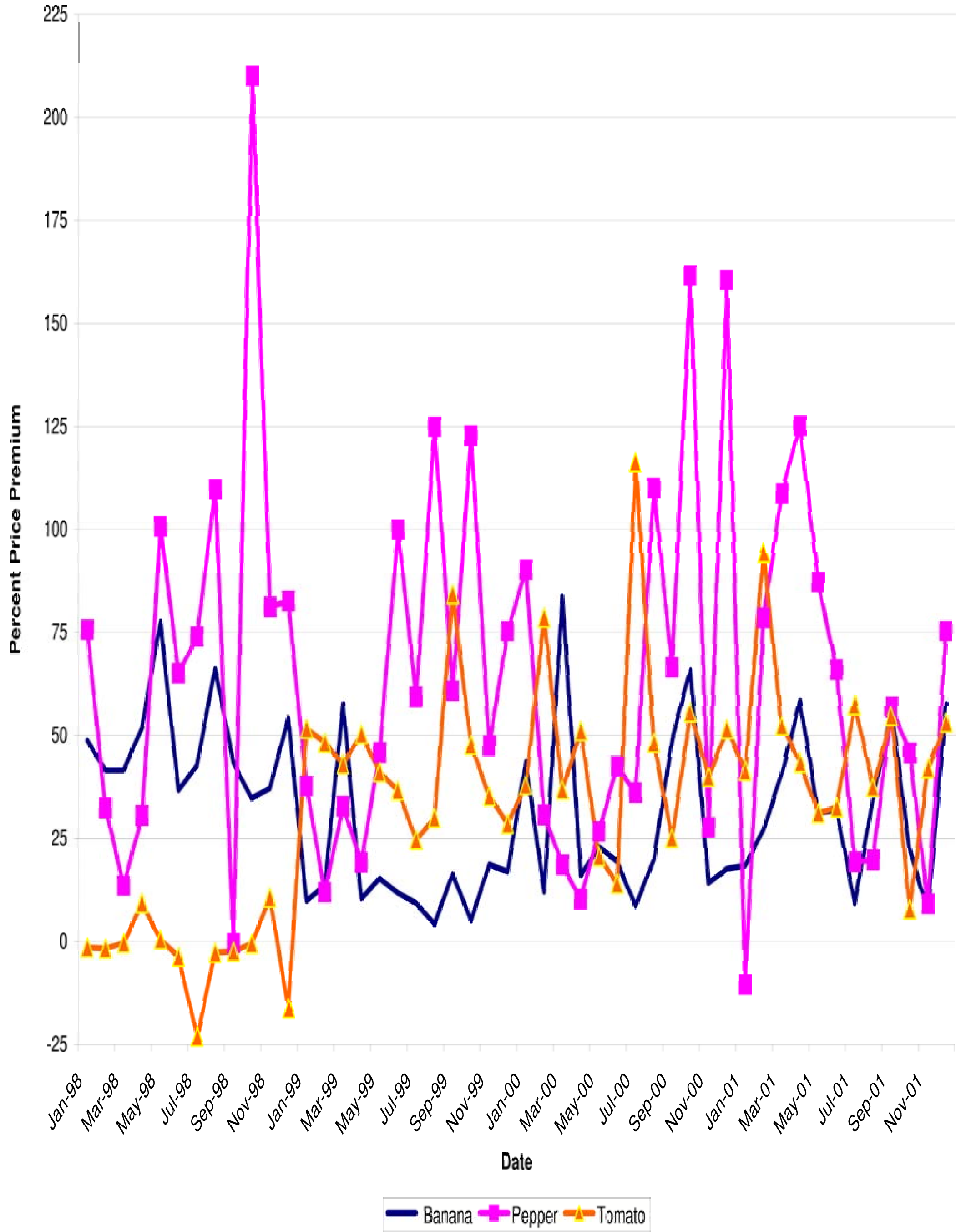


Figure 5.5: Percentage of Organic Purchases by Produce Type Over Time

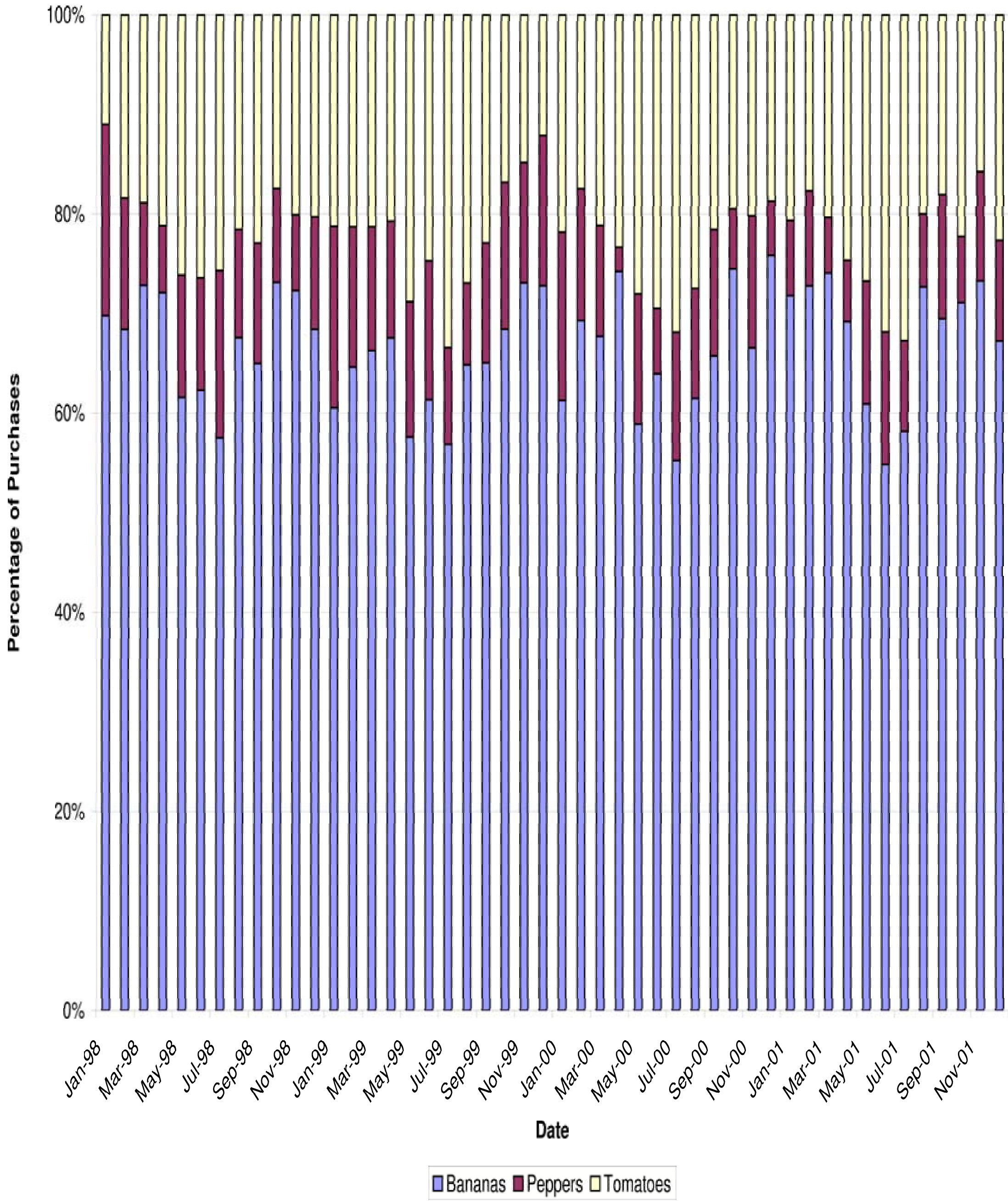


Figure 5.6: Percentage of Overall Purchases by Produce Type that are Organic

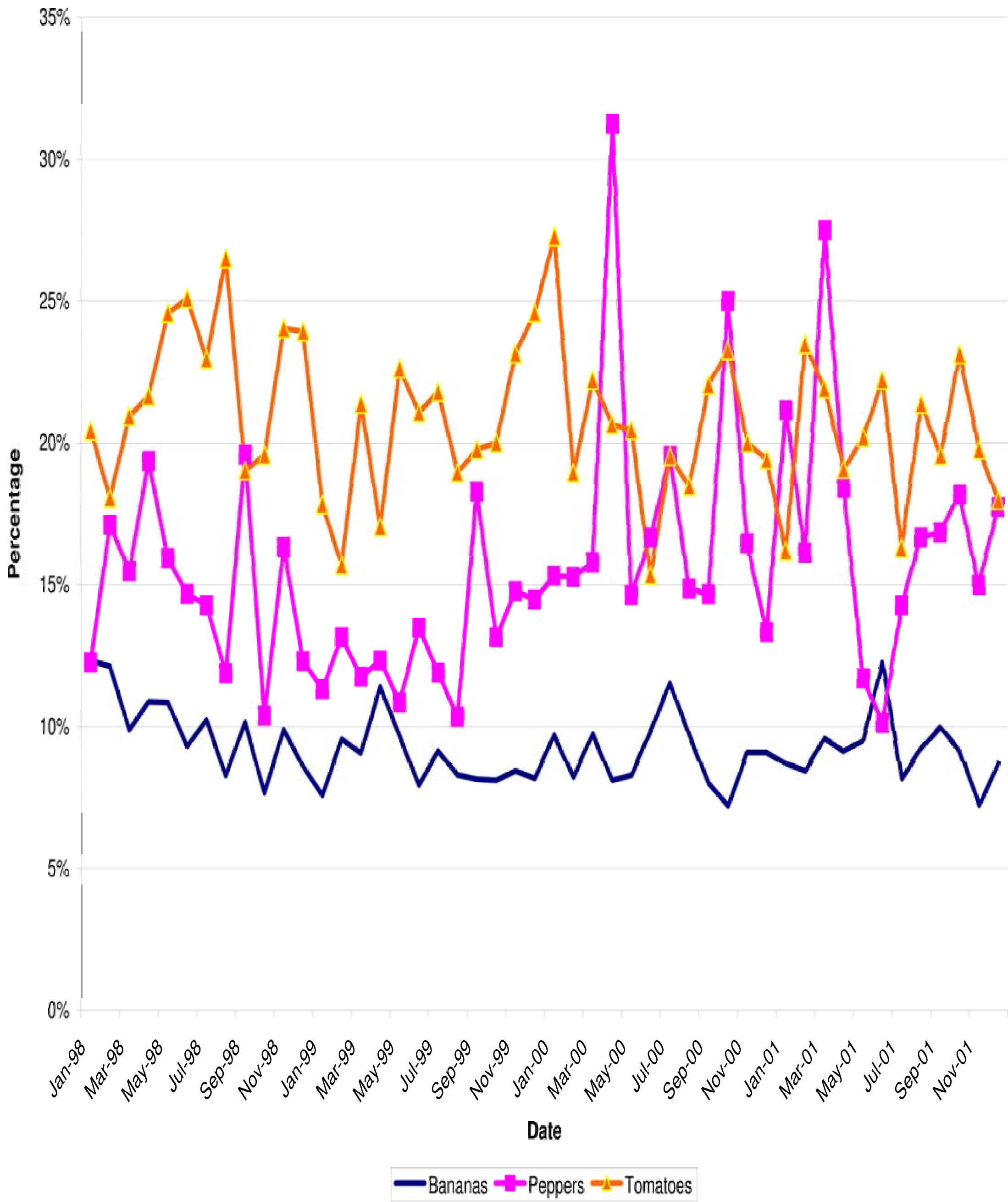


Figure 5.7: Conventional and Organic Purchases by Race / Ethnicity (1998-2001)

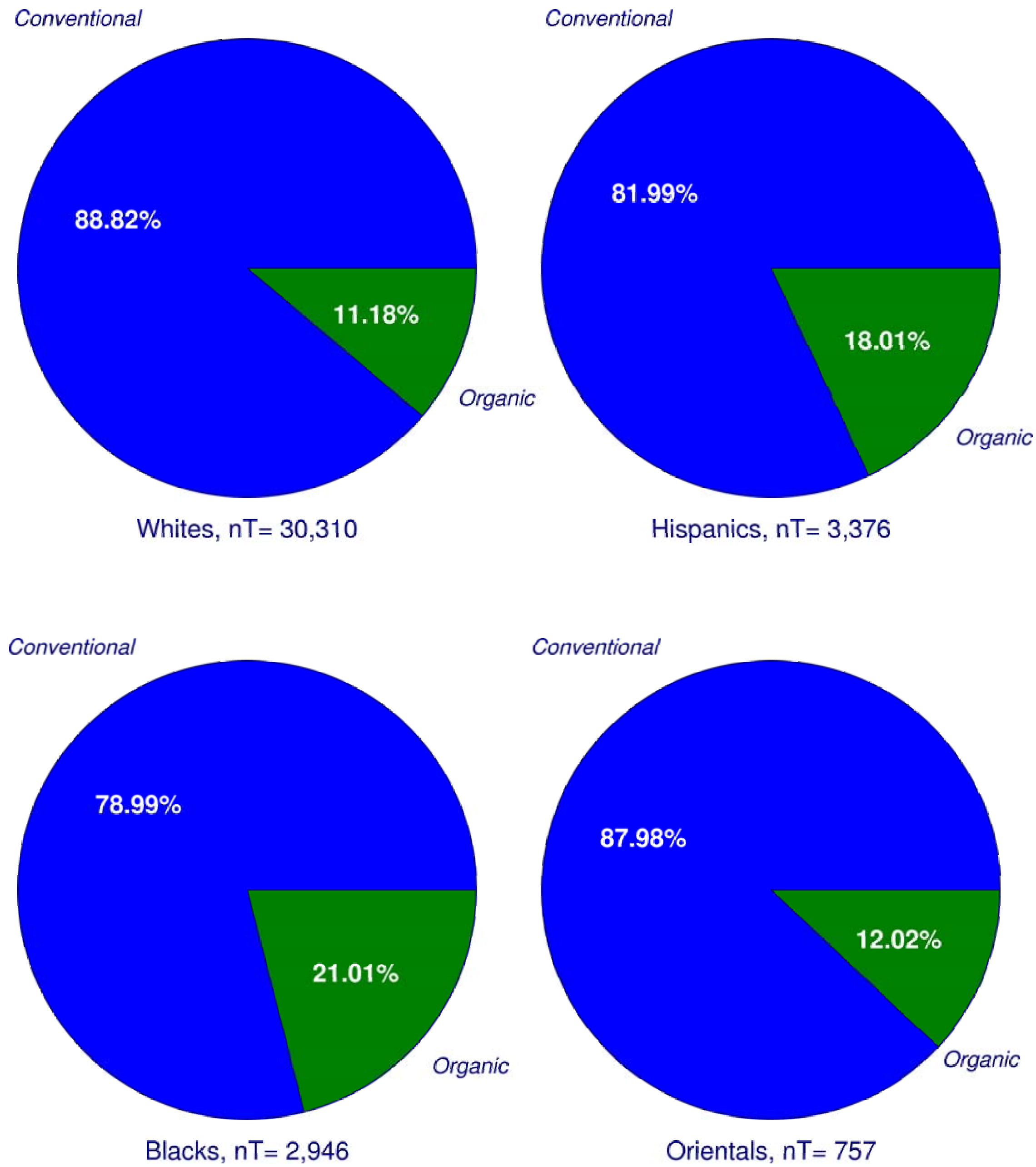
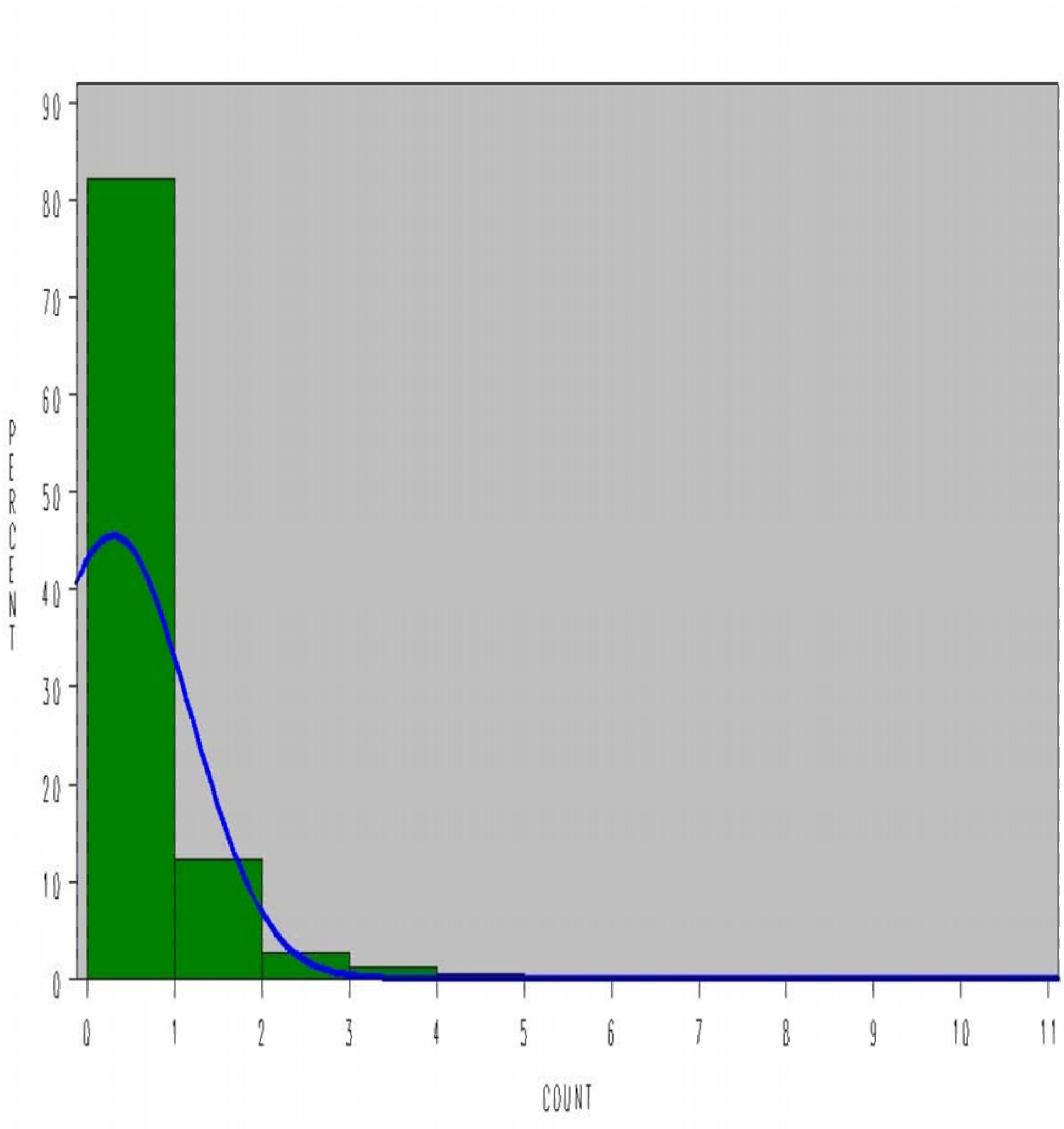


Figure 5.8: Distribution of Household Quarterly Counts of Organic Produce Purchases (Poisson Random Effects Model)



CHAPTER 6

EMPIRICAL MODEL

6.1 Model Specification

Multiple empirical models are examined throughout this study. The effort is to accurately estimate purchase behavior as a function of different market and household characteristics. Models with and without a dynamic element are explored. In dealing with non-linear panel models, estimation can become quite rigorous and computationally intensive yet shed very interesting results.

The economic demand theory discussed in Chapter 4 motivates an empirical model that can accommodate corner solutions at zero. Suitably, a static censored panel model is estimated to examine effects on organic quantity demanded using weighted maximum likelihood. Also, a dynamic panel model that includes a lagged dependent variable and controls for unobserved heterogeneity is explored but not estimated due the fact that it still remains at the frontier of the literature.

Outside of the economic theory in Chapter 4, a panel model that accounts for count outcomes is estimated to establish price, income, and demographic effects on repeated organic purchases. Though a theoretical justification for such a model is withheld, a count model provides determinants for frequency of purchase.

6.2 Tobit Model With Random Effects

Fresh organic produce continues to gain retail market share within the industry though it still only maintains a small percentage of retail sales. In order to estimate organic purchase activity, consideration must be made that within a large heterogeneous household panel, one will observe a small proportion of organic purchases compared to overall produce purchases. With a large frequency of purchases being non-organic, the sample is censored in that organic purchases would be reported as zero for households who do not buy organic. To account for a censored panel, a regression model known as a Tobit model (Tobin, 1958) can be used. A Tobit model allows for a discrete censored point(s) followed by a continuous range of observation values on the dependent variable. For this study, a censored panel model is appropriate due to the fact that the proposed dependent variable of organic produce quantities purchased has a high percentage of left-censored observations at zero while the dependent variable is continuous to the right.

Ignoring time and cross-sectional subscripts, the first part of the Tobit model incorporates the expected value of y where it is uncensored in the regression:

$$E(y|y > 0, x) = x\beta + \sigma\lambda(x\beta/\sigma) \tag{6.1}$$

where $\lambda(c) = \frac{\phi(c)}{\Phi(c)}$ is the ratio between the standard normal pdf and the cdf each evaluated at c. The error term is assumed $\epsilon \sim N(0, \sigma^2)$. The second part includes the probability that the observation will be observed as a positive value if censoring occurs at zero. Given both, the expected value of y given x for the censored model is equal to:

$$\begin{aligned} E(y|x) &= Prob(y = 0) \cdot 0 + Prob(y > 0|x) \cdot E(y|y > 0, x) \\ &= \Phi(x\beta/\sigma) \cdot E(y|y > 0, x) \end{aligned} \tag{6.2}$$

Now combining (6.1) and (6.2) yields:

$$E(y|x) = \Phi(x\beta/\sigma)x\beta + \sigma\phi(x\beta/\sigma) \tag{6.3}$$

As seen in the summary statistics section, the dependent variable's highly censored nature at zero suggests that a Tobit model with random individual effects using weighted maximum likelihood estimation could ensure a vector of parameter estimates that are unbiased, consistent and efficient. A least-squares estimator would provide biased and inconsistent estimates in this case due to the dependence between the error term and the regressors in the model where censoring occurs (Hsiao, 2003).

In the Tobit random-effects (RE) panel model, the conditional density of the unobserved effect is assumed. To control for the presence of unobserved heterogeneity, consider the model as:

$$y_{it}^* = x_{it}\beta + \mu_i + \epsilon_{it},\ t = 1...,T,\ i = 1...,n, \tag{6.4}$$

$$y_{it} = \begin{cases} y_{it}^* & \text{if } y_{it}^* > 0, \\ 0 & \text{if } y_{it}^* \leq 0. \end{cases}$$

Within the RE model, μ_i and ϵ_{it} are randomly distributed composite error terms. The individual random effect term, μ_i, is assumed to be distributed $IID(0, \sigma_\mu^2)$ across i where the unobserved effect remains constant across time. The error term, ϵ_{it}, is $IID(0, \sigma_\epsilon^2)$ across i and over t.

Under the assumption that μ_i is randomly distributed with the density function $g(\mu)$, the likelihood function of the standard Tobit random effects model for censored data takes the form:

$$\prod_{i=1}^{N} \int \left[\prod_{t\epsilon c_i} F(-x_{it}\beta - \mu_i) \prod_{t\epsilon \bar{c}_i} f(y_{it} - \mu_i - x_{it}\beta) \right] g(\mu_i) d\mu_i \tag{6.5}$$

$f(\cdot)$ denotes the density function of the error term, ϵ_{it}, and $F(a) = \int_{-\infty}^{a} f(\epsilon) d\epsilon$. Also, $c_i = \{t | y_{it} = 0\}$ and $\bar{c}_i$ denotes its complement. Maximizing the likelihood function with respect to the parameters of interest yields consistent and asymptotically normal distributed estimators (Hsiao, 2003).

The log likelihood is a straightforward generalization of (6.5) in the form:

$$\begin{aligned} logL = const. + \sum_{i=1}^{N} \sum_{t=1}^{T} & [(1 - \delta_{it})\ log\Phi\ (\frac{-\mu_i - x_{it}\beta}{\sigma}) \\ & + \delta_{it}[-\frac{1}{2}\ log\sigma^2\ - \frac{1}{2}(\frac{y_{it} - \mu_i - x_{it}\beta}{\sigma})^2] \end{aligned} \tag{6.6}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal density and distribution functions, respectively. If $y_{it}^* > 0$ then $\delta_{it} = 1$, $\delta_{it} = 0$ otherwise (Nerlove, 2002).

6.2.1 Marginal Effects for the Tobit Model With Random Effects

Given the censored nature of the dependent variable, the results of the partial derivative of $E[y_{it}^*|x_{it}, \mu_i]$ with respect to x_{it} is not useful because y_{it}^* is censored at zero.

$$\frac{\partial E[y_{it}^*|x_{it}, \mu_i]}{\partial x_{it}}$$

To examine the decomposition of the Tobit marginal effect consider:

$$\frac{\partial E[y_{it}|x_{it}, \mu_i]}{\partial x_{it}} = \beta * \Phi_{it}[1 - \lambda_{it}(\alpha_{it} + \lambda_{it})] + \phi_{it}(\alpha_{it} + \lambda_{it})] \tag{6.7}$$

where $\alpha_{it} = \beta' x_{it}$, $\Phi_{it} = \Phi(\alpha_{it})$ and $\lambda_{it} = \phi_{it}/\Phi_{it}$ (McDonald and Moffitt, 1980). As a result, the following is more useful (Greene, 2003):

$$\frac{\partial E[y_{it}|x_{it}, \mu_i]}{\partial x_{it}} = Prob[y_{it} > 0]\frac{\partial E[y_{it}|x_{it}, \mu_i, y_{it} > 0]}{\partial x_{it}} + E[y_{it}|x_{it}, \mu_i, y_{it} > 0]\frac{\partial Prob[y_{it} > 0]}{\partial x_{it}} \tag{6.8}$$

The marginal effect shown in 6.6 is two-fold. First, a change in the explanatory variable affects the conditional mean of y_{it} where it is positive within the distribution. Second, a change in x affects the probability that the observation be observed as a positive value and will fall into that part of the distribution.

6.2.2 Estimation of Price Elasticities

Given the marginal effects of the model above, the price elasticity measure can be estimated within the model as:

$$\xi_d = \frac{\partial E[y_{it}|x_{it}, \mu_i]}{\partial x_{it}} . \frac{x_{it}}{E[y_{it}|x_{it}, \mu_i]} \tag{6.9}$$

Since the elasticities vary for every observation, it is desirable to report a summary measure. A convenient summary measure is to evaluate the elasticity at the sample means of the explanatory variables. In the Tobit case, evaluating elasticities at "sample means" is interpreted as the mean of the x_{it} observations in the sample at $E[y_{it}|x_{it}, \mu_i]$. In this case, elasticities are evaluated at mean organic and conventional prices across i and t.

6.2.3 Testing for Panel-Level Heteroscedasticity

Testing for heteroscedasticity is especially important when cross-sectional data is used for empirical analysis. When panel data is used, the cross-sectional element combined with a time dimension creates increased difficulty in testing for heteroscedastic variances but does not reduce the need to adequately test for the presence of heteroscedasticity.

The above model assumes that the regression disturbances are homoscedastic with the same variance across households. This assumption may be restrictive, however, as the cross-sectional units vary in size and may exhibit differing variation. Assuming homoscedastic variation when heteroscedasticity is present still creates consistent coefficients but these estimates would not be efficient with minimum variance (Baltagi, 2005). With a heteroscedastic variance, the standard errors are biased and robust standard errors correcting for heteroscedastic disturbances should be computed if possible.

For panel data, the many standardized tests for heteroscedasticity that exist for cross-sectional data cannot be directly applied. Due to the possibility of heteroscedasticity of both the individual effect, μ_i, and the time-varying error term, ϵ_{it}, a more limited literature for tests of the null hypothesis of homoscedasticity for either term exist.

Lejeune (1996) deals with maximum likelihood estimation and Lagrange multiplier (LM) testing of general heteroscedastic one-way error component models. He derives a model where the variance functions are parametrically specified in a context that allows for unbalanced panels. With incomplete panels, Lejeune (1996) derives two joint LM tests for no individual effects and homoscedasticity of the error term. His first test involves a random effects one-way error model where

$\mu_i \sim IID(0, \sigma_\mu^2)$ and the remainder term that is heteroscedastic $\epsilon_{it} \sim N(0, \sigma_{\epsilon_{it}}^2)$ with $\sigma_{\epsilon_{it}}^2 = \sigma_\epsilon^2 h_\epsilon(Z_{it}\theta_1)$. His hypothesis is $H_0 : \theta_1 = \sigma_\mu^2 = 0$ where the restricted MLE is OLS. He argues that since one is testing that σ_μ^2 is zero, there is no need to consider a variance function for μ_i (Baltagi, 2005).

Holly and Gardiol (2000) derive a score test for homoscedasticity in one-way error components model. They derive a score test for the null hypothesis of homoscedasticity of the individual random effects assuming homoscedasticity of the remainder error term.

Also, Baltagi et al. (2005) derive a joint LM test for heteroscedasticity in a one-way error component model ($H_0 = \theta_1 = \theta_2$). Unlike Lejeune where the restricted MLE is OLS, Baltagi et al. (2005) provide a restricted MLE that is a one-way homoscedastic error component model. The benefit of this study is where heterogeneity across households is allowed to exist which allows for $\sigma_\mu^2 > 0$. This model is similar to Holly and Gardiol (2000); however, it does not assume homoscedasticity of the remainder error term. Of particular interest is the addition that derives an LM test for the null hypothesis of homoscedasticity of the remainder error term while assuming homoscedasticity of the individual effects.

6.2.4 Testing for Panel-Level Autocorrelation

Given the temporal element of panel models, a greater potential for autocorrelation exists. Wooldridge (2002) derives a simple test for autocorrelation in panel models. Based on the residuals of the first-difference (FD) estimator for the pooled regression of Δy_{it} on $\Delta x_{it} \quad t = 2..., T;\ i = 1, 2, ..., N$, the regression is based on $T - 2$ time periods:

$$\hat{e}_{it} = \hat{\rho}\hat{e}_{i,t-1} + error_{it}, \quad t = 3, 4, ..., T \quad i = 1, 2, ..., N \tag{6.10}$$

The t statistic on $\hat{\rho}$ is the test statistic and the test is only applicable where $T > 2$. Obviously, with $T = 2$ the test is not available because the first differencing requires 2 time periods. The process involves computing the robust variance matrix of the FD estimator and a significant test statistic of $\hat{\rho}$ would indicate panel level autocorrelation.

6.3 Computational Methods

Without recent versions of statistical and data analysis computer programs, manipulating the raw data and programming a random effects Tobit estimator would be extremely difficult. Even with recent software, data manipulations and empirical results required long computation times. Jensen and Schroeter (1992) comment of their inability to program such an estimator in their study where it would be appropriate. They comment, "random effects Tobit procedures would require optimization of a likelihood function involving multivariate normal integrals evaluated using numerical quadrature, a procedure that would be computationally very costly when applied to problems with the dimensions of ours."

Furthermore, Butler and Moffitt (1982) note that there is a "widespread feeling among those working in the area that one possible method of evaluation, the use of quadrature techniques, is in principle possible but is in practice computationally too burdensome to consider."

Thankfully, new techniques combined with advances in computer hardware and software have empowered researchers to further the related literature where computational barriers once existed. The random effects Tobit estimator using Gaussian quadrature, now programmed in STATA SE 9, was used successfully to obtain estimates for the model.

6.3.1 Gaussian Quadrature

Gaussian quadrature is considered to be quite efficient compared to standard quadrature techniques such as trapezoidal integration and allows for computational feasibility with modern computers. Given (6.3), evaluation of the integral using conventional quadrature procedures would be extremely burdensome. "Gaussian quadrature, on the other hand, is a much more sophisticated procedure that requires the evaluation of the integrand at many fewer points in the domain of μ, thus achieving gains in computational efficiency of several orders of magnitude" (Butler and Moffitt, 1982).

The formula in the Gaussian quadrature that evaluates the necessary integral is the Hermite integration formula, $\int_{-\infty}^{\infty} e^{-Z^2} g(Z)\ dZ = \sum_{j=1}^{G} w_j\ g(Z_j)$. Where G is the number of evaluation points, w_j is the weight given to the j^{th} evaluation point, and $g(Z_j)$ is $g(Z)$ evaluated at the j^{th} evaluation point of Z. In this case, $g(Z)$ is the product of T univariate cdf's and the normal density function $f(\cdot)$ in equation (6.3) contains a term of the form $exp(-Z^2)$.

Though Butler and Moffitt report coefficient estimates for a one-factor multinominal probit model, they state that the technique is applicable to other specific limited-dependent-variable models such as single and double-bound Tobit.

6.4 Tobit Random Effects Models in Related Literature

Though censored panel estimation remains relatively new in directly related organic produce literature, some studies have done work with the empirical model of particular interest in this study, the Tobit Random Effects (RE) model.

6.4.1 Tobit RE in Beef Demand Panel

Of relevance to this study, Jensen and Schroeter (1992) encounter a very similar data structure and estimation problem in their examination of television advertising effects on beef demand. In their study, they use a household scanner panel similar to that in this study. When participating households purchased a fresh beef product at a participating store, relevant data were read by a designated UPC scanner and automatically updated the household's computerized purchase log. In interesting fashion, Jensen and Schroeter control for advertising exposure within the data. All households subscribed to a cable television system with the capability of transmitting different commercial advertisements. Each household was then assigned to one of three different panels characterized by different intensities of ad exposure. The resulting data set is a panel for tracking the beef purchases of 1782 households for

23 four-week periods.

The censoring nature of the dependent variable of the study is analogous to that of this study of organic produce demand. Due to the frequent occurrence of zero purchases of beef for a given household in a given four-week period, a model of household beef demand is introduced to accommodate the limited dependent variable aspect of the data's structure.

The linear, single equation structure of their model is quite similar to the model used in this study. Also, many of their techniques of aggregating price and purchase information are similar to those used here. Furthermore, because direct price substitute information was not present in their scanner data, they attach regional price data from secondary sources. This study on the other hand, had the luxury of an extremely extensive original data set which did not require developing price proxies from secondary sources.

Jensen and Schroeter were unable, however, to directly apply a random effects Tobit model due to computational restrictions of optimizing a likelihood function using numerical quadrature. Their estimates include random-effects, fixed-effects, and a standard Tobit model. Later, Benson et al. (2002) would build on the work by Jensen and Schroeter as they confront the same problem with the same data but successfully implement a random effects Tobit model with Bayesian Inference and provide new empirical results.

6.4.2 Tobit RE in Dealership Panel

Utilizing a new database for dealer profit information, Harless and Hoffer (2002) aimed to empirically model whether price discrimination by new vehicle dealers results in women paying higher prices for new vehicles. In their regression, Harless and Hoffer model effects on dealer gross profit and financed vehicle purchases where a high degree of left-censored zeros occur. To avoid mixing censored and uncensored zeros, they employ a Tobit RE model. They also provide estimates of σ_μ and σ_ϵ but do not provide evidence of possible heteroscedasticity of either.

6.5 Poisson Random Effects Panel Model for Count Data

The above Tobit RE model assumes a censored point at zero followed by a continuous observation of quantity values to the right for the dependent variable. Since fresh organic produce items are perishable, quantity purchased may not be of as much interest as the number of times a household made an organic purchase during a given time period.

Due to the fact that a high level of variation in quantity purchased may not be observed due to issues of spoilage, a count model with number of organic purchases as the dependent variable should be explored. In the interest of modeling the determinants for frequency of purchase, a count panel model that accommodates heterogeneity is estimated.

Although the Poisson RE model makes the assumption that the heterogeneity term is uncorrelated with the regressors in the model, the fixed-effects alternative carries other flaws (a large number of incidental parameters) that make RE estimation more feasible in this case.

The Poisson parameter is normally denoted as λ where the conditional mean is $log\ \lambda_{it} = X_{it}\beta + \alpha_i$ (Greene, 2003). The basic Poisson probability specification is

$$pr(n_{it}) = f(n_{it}) = \frac{e^{-\lambda_{it}}\lambda_{it}^{n_{it}}}{n_{it}!} \tag{6.11}$$

where n_{it} is the observed event count for household i in time period t. The first two moments for this distribution are the same as $E(n_{it}) = \lambda_{it}$ and $V(n_{it}) = \lambda_{it}$.

For random household specific effects, $\tilde{\lambda}_{it} = \lambda_{it}\tilde{\alpha}_i$ is specified where $\tilde{\alpha}_i$ is the unobserved effect. In this case, the Poisson parameter $\tilde{\lambda}_{it}$ is a random variable rather than a function of X_{it}. Because $\tilde{\lambda}_{it}$ needs to be positive, it can be rewritten as

$$\tilde{\lambda}_{it} = \lambda_{it}\tilde{\alpha}_i = e^{X_{it}\beta+\mu_0+u_i} \tag{6.12}$$

where μ_i is the household specific effect and μ_0 is the overall intercept (Hausman et al., 1984). Given (6.11), the Poisson probability becomes

$$pr(n_{it}|X_{it}, \mu_i) = \frac{e^{-\lambda_{it} e^{\mu_i}} (\lambda_{it} e^{\mu_i})^{n_{it}}}{n_{it}!}. \tag{6.13}$$

The joint density function for all counts and the household specific effect is

$$\begin{aligned} pr(n_{i1}, ..., n_{iT}, \mu_i | X_{i1}, ..., X_{iT}) &= pr(n_{i1}, ..., n_{iT} | X_{i1}, ..., X_{iT}, \mu_i) g(\mu_i) \\ &= \prod_t \frac{\lambda_{it}^{n_{it}}}{n_{it}!} e^{-e^{\mu_i} \Sigma_t \lambda_{it}} e^{\mu_i \Sigma_t n_{it}} g(\mu_i) \end{aligned} \tag{6.14}$$

where $g(\mu_i)$ is the pdf of μ. Given (6.14), μ_i is an unobservable random variable which is assumed to be randomly distributed across i. Because μ_i is unobservable, it must be integrated out from the joint density function. This is done by assuming that $\alpha_i = e^{\mu_i}$ is distributed as a gamma random variable with parameters (δ, δ), so that $E\alpha_i = 1$ and $V\alpha_i = 1/\delta$. By integrating by parts, you find

$$\begin{aligned} & pr(n_{i1}, ..., n_{iT} | X_{i1}, ..., X_{iT}) \\ &= \int_0^\infty \prod_t \left[\frac{\lambda_{it}^{n_{it}}}{n_{it}!} \right] e^{-\alpha_i \Sigma_t \lambda_{it}} \alpha_i^{\Sigma_t n_{it}} f(\alpha_i) d\alpha_i \\ &= \prod_t \left[\frac{X_{it}^{n_{it}}}{n_{it}!} \right] \left[\frac{\delta}{\Sigma \lambda_{it} + \delta} \right]^\delta (\Sigma \lambda_{it} + \delta)^{-\Sigma n_{it}} \frac{\Gamma(\Sigma n_{it} + \delta)}{\Gamma(\delta)} \end{aligned} \tag{6.15}$$

where $\Gamma(\cdot)$ is the gamma function, $\Gamma(z) = t^{z-1} e^{-t}$ for $z > 0$ (Hausman et al., 1984).

The subsequent marginal effects for this model are easily attained where

$$\frac{\partial\, E(n_{it}|X_{it})}{\partial\, X_{it}} = \lambda_{it} \beta.$$

These marginal effects can be evaluated at sample means for the predicted number of events assuming $\mu_i = 0$ and approximate standard errors can be obtained by bootstrapping.

6.6 Censored Panel Models with Dynamics

With the interest of measuring loyalty and habits of organic purchases, the hypothesis is examined in that consumption in the current period, consists of a function of consumption in the previous period, as well as a vector of other elements. Dynamics, coupled with other variables included within a dynamic censored model, allow for feasible estimates of organic demand in conjunction with the theoretical model. Recent contributions to econometric literature have enabled the exploration of different dynamic estimators.

The remainder of this chapter will introduce a proposed dynamic censored panel model but it will not be used for estimation due to programming difficulties and time constraints. An overview of the estimator is provided and future work is encouraged.

6.6.1 Dynamic Models Controlling for Heterogeneity

Consider a dynamic model with individual effects within a censored context. Due to censoring, we only observe this model:

$$y_{it}^{*} = max\{0, \gamma y_{i,t-1}^{*} + \boldsymbol{\beta' x}_{it} + \mu_i + \epsilon_{it}, \},\ t = 1..., T,\ i = 1..., n, \tag{6.16}$$

Considering (6.16), newer econometric models have been established to estimate a dynamic model which corrects for censoring.

6.6.2 The Non-Linear GMM: With Correction for Censoring

The GMM estimator is a two-stage estimator due to Hu (2002) which is based on orthogonality conditions which explicitly account for the censoring problem. This model takes a fixed-effects approach to estimation in order to avoid the random-effects model and not to make any distributional assumptions on the individual effect which could lead to biased estimates (Hu, 2002). Within this estimator, Hu suggests "differencing away" α_i, the "nuisance parameter" which corresponds to the individual fixed effect of the model[1]. After "trimming," the moment conditions can be used to formulate the GMM estimator.

To illustrate how the trimmed GMM estimator works, assume the case of a panel with $T = 2$:

$$y_{i1}^* = \gamma_0 y_{i0}^* + x_{i1}\beta_0 + \alpha_i + \varepsilon_{i1}$$

$$y_{i2}^* = \gamma_0 y_{i1}^* + x_{i2}\beta_0 + \alpha_i + \varepsilon_{i2}$$

Whereas ε_{i1} and ε_{i2} are independent and identically distributed conditional on $\{y_{i0}^*, x_{i1}, x_{i2}, \alpha_i\}$, then on the same variables, $(\varepsilon_{i1}, \varepsilon_{i2})$ has the same distribution as $(\varepsilon_{i2}, \varepsilon_{i1})$. Another representation could be made in which $(\varepsilon_{i1}, \varepsilon_{i2})$ is symmetrically distributed around the 45^0 line through the origin. Followed by $(y_{i1}^*, y_{i2}^* - y_{i1}^*)$ which is symmetrically distributed around the 45^0 line through $(\gamma_0 y_{i0}^* + x_{i1}\beta_0 + \alpha_i,\ x_{i2}\beta_0 + \alpha_i)$. Given that symmetry holds for any value of the nuisance parameter, α_i, then conditional on $\{y_{i0}^*, x_{i1}, x_{i2}\}$, $(y_{i1}^*, y_{i2}^* - y_{i1}^*)$ is symmetrically distributed around the 45^0 line through $(\Delta w_i\theta_0, 0)$ where $\Delta w_i\theta_0 \equiv (x_{i1} - x_{i2})\beta_0 + \gamma_0 y_{i0}^*$. However, symmetry is destroyed by censoring in this case.

6.6.3 The Destruction of Symmetry

In the case above, symmetry is destroyed by censoring whereas

$$\tilde{y}(\gamma) = y_2 - (\gamma)y_1. \tag{6.17}$$

In the $T = 2$ example, where $\gamma = \gamma_0$, we have $\tilde{y}(\gamma_0) = max(-\gamma_0 y_1,\ y_2^* - \gamma_0 y_1^*)$. In comparing the pair $(y_1, \tilde{y}(\gamma_0))$ and $(y_1^*, y_2^* - \gamma_0 y_1^*)$, one can notice that they are similar except that the first element y_1 is y_1^* censored from below by zero, while in the latter, $\tilde{y}(\gamma_0)$ is $y_2^* - \gamma_0 y_1^*$ is censored from below by $-\gamma_0 y_1$, conditional on $y_1 > 0$. Therefore, because of the two asymmetric censoring points, censoring destroys symmetry between y_1^* and $y_2^* - \gamma_0 y_1^*$.

[1]Symmetric trimming used in the estimation of censored models can be found in Powell (1986), Honoré (1993), and Honoré and Powell (1994).

6.6.4 Trimming the Estimator

The insight into the Hu GMM estimator comes with the restoration of symmetry by "trimming" observations outside the region, S, which is:

$$1\{S\} \equiv 1\{y_0 > 0,\ y_1 > 0,\ y_2 > 0\}$$

$$\cdot[1\{\Delta w\theta_0 \geq 0\}\ \cdot 1\{y_1 > \Delta w\theta_0 - \gamma_0(\tilde{y}(\gamma_0) + \Delta w\theta_0)\} \cdot 1\{\tilde{y}(\gamma_0) > -\Delta w\theta_0\} \tag{6.18}$$

$$+1\{\Delta w\theta_0 < 0\} \cdot 1\{\tilde{y}(\gamma_0) > -\Delta w\theta_0\}]$$

where $1\{\cdot\}$ is the indicator function.

Conditional on being in the region S, the pair of variables $(y_{t-1}, \tilde{y}_t(\gamma_0))$ is symmetrically distributed around the 45^0 line through $(\Delta w\theta_0, 0)$. The parameter θ_0 can be estimated based on the orthogonality conditions.

6.6.5 Construction of the Moment Conditions

Estimation of model is complicated by the fact that the nuisance parameter, α_i, enters the model nonlinearly due to censoring. Therefore, unlike in the linear models, it is in general not possible to eliminate α_i by first-differencing the data. Hu (2002) proposes a method that exploits the idea of symmetric trimming to "difference away" α_i and construct a set of moment conditions that can be used to estimate the parameters of interest.

The standard GMM estimator is defined as:

$$\hat{\theta}_{GMM} = arg\ \min_{\theta}\ m_n(\theta)' \cdot A_n \cdot m_n \tag{6.19}$$

where $m_n(\theta)$ is the sample analog of the orthogonality conditions

$$m_n(\theta) = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n}[1\{S\} \cdot sgn(y_{i,t-1} - \Delta_{it}(\gamma, \delta_t) - \tilde{y_{it}}(\gamma))]\ \ t = 3, ...T \\ \frac{1}{n}\sum_{i=1}^{n}[1\{S\} \cdot sgn(y_{i,t-1} - \Delta_{it}(\gamma, \delta_t) - \tilde{y_{it}}(\gamma)) \cdot y_{i,t-2}\ \ t = 3, ...T \end{pmatrix} \tag{6.20}$$

6.6.6 The Benefits of GMM Estimation Over MLE

The GMM estimator suggested above has multiple advantages over weighted maximum likelihood estimation that justify its use. First, it makes no distributional assumptions on the composite error term. In the static Tobit RE panel model, μ_i and ϵ_i are randomly distributed composite error terms. The error term, ϵ_{it}, is assumed to be IID(0,$\hat{\sigma}_\epsilon$) across i and over t as well as the unobserved effect, $\mu_i \sim IID(0, \sigma_\mu^2)$. The GMM estimator suggested makes no such assumptions. Also, by construction it handles problems of endogeneity while eliminating the potential for heteroscedasticity and autocorrelation in parameter estimates.

CHAPTER 7

EMPIRICAL RESULTS

7.1 Parameter Estimates

Tables 7.1 and 7.2 display the parameter estimates for the random-effects Tobit panel model for organic purchases of raw fresh bananas, peppers, and tomatoes from ACNielsen. Tables 7.3 and 7.4 display marginal effect estimates based on McDonald and Moffitt (1980) and marginal effect estimates for the probability of being uncensored. Table 7.5 displays own- and cross-price elasticity estimates for organic purchases of different fresh produce types. Table 7.6 and 7.7 presents parameter estimates and marginal effects for the random-effects Poisson model where quarterly organic purchase counts are regressed on given variables and average quarterly organic and conventional prices.

7.2 Price Effects

For all models examined, price effects of organic and conventional items have a statistically significant effect on organic purchase quantities as well as on the number of organic purchases made by households. In terms of own-price, all models suggest that organic price has a strongly significant effect on organic purchase quantity and likelihood of purchase. Own-price effects are negative as would be expected. Conventional cross-prices have a statistically significant effect on organic purchase activity also. As conventional produce can be considered a substitute item to organic produce, positive cross-price effects are to be expected. These results show that relative prices between organic and conventional items clearly affect organic purchase decisions and activity.

When examining the relative magnitudes of the price effects, noteworthy is the fact that the strength of the effect seems to depend on whether organic quantity or number of purchases is being modeled. In the Tobit models where organic purchase quantity is the dependent variable, organic own-price generally has a larger magnitude effect than conventional cross-prices. In the Poisson model where number of purchases is the dependent variable, conventional prices have a stronger effect in terms of magnitude.

7.3 Income Effects

Although the income bracket dummies for increasing levels of income are largely time-invariant across household, the results suggest that they have a strong significant negative effect on organic purchase activity. In examining the marginal effects of a discrete change into each incremental income bracket, every effect is statistically significant for every sample and empirical model with the exception of one coefficient. This result suggests that household income is clearly an important element in determining household organic purchase activity.

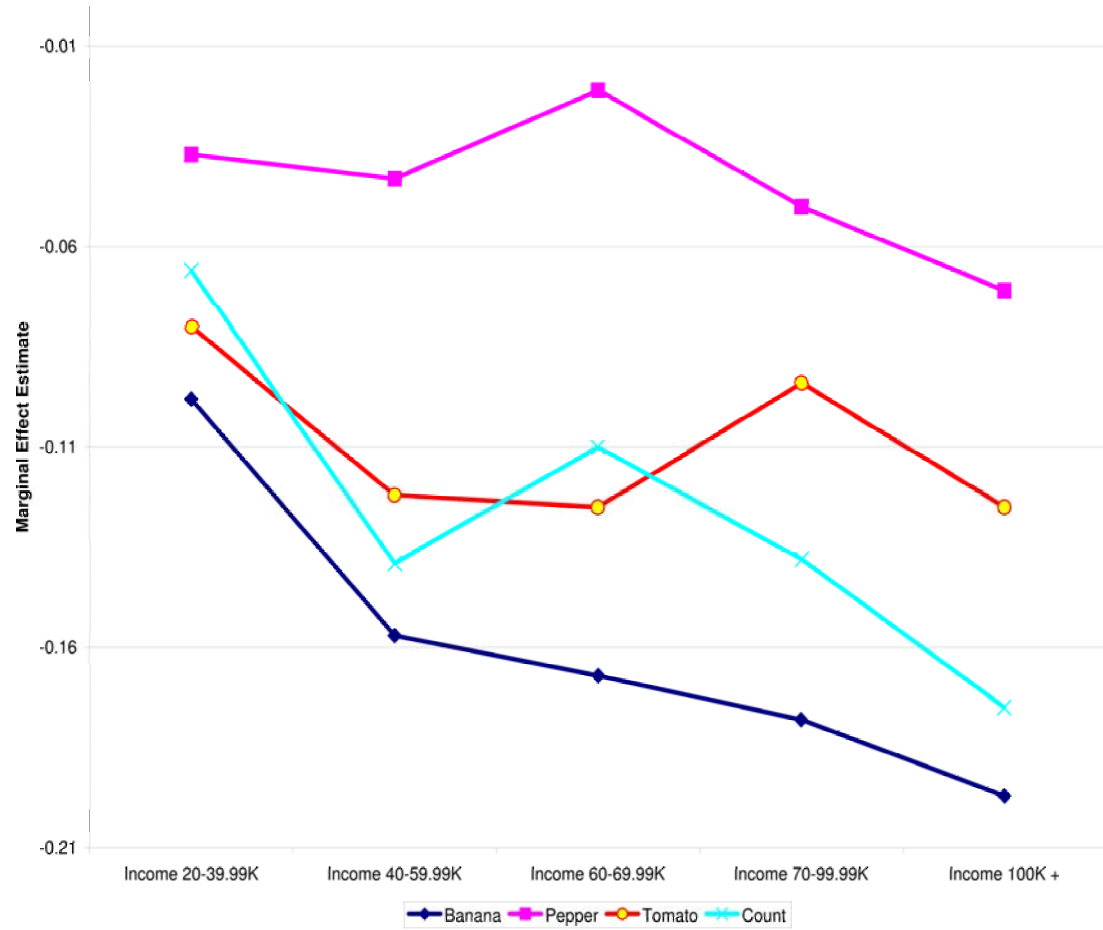
The results also suggest that higher levels of household income have an increasingly negative affect in terms of magnitude. Figure 7.1 displays graphically the marginal effect estimates for the models estimated. Results suggest that the income brackets for higher levels of income have a negative and significant effect on both quantity and number of organic purchases.

7.4 Demographic Effects

7.4.1 Household Composition

When examining the effects of presence of children, the results suggest that the number of adolescent children between the age of 13 and 17 years in the household has a positive and strongly significant effect in every model examined. Also, the number of younger children in the household, more specifically children under 6, also had a positive effect on organic produce purchases with statistical significance in two of the random effect Tobit models. The number of children under 6 was not a significant variable in the count model for organic purchases, however.

Figure 7.1: Marginal Effect Parameter Estimates of Increasing Household Income Categories



Notably, the number of adults living in a household, which excludes the male and female household heads, also had a positive effect. This variable was significant for 2 Tobit models and the Poisson count model. This result is interesting and is ambiguous as to the specific composition of these households (ie. younger households with roommates, etc.).

7.4.2 Age Effects

The models suggest that age of female head of household shows a statistically significant effect on organic purchase quantities for the banana and pepper models and on number of organic purchases in the Poisson count model. Every incremental 5 year age bracket in which the female head of the household falls into has a negative effect on organic purchases. Male age effects are primarily insignificant across all models.

7.4.3 Employment and Occupation Effects

Female head of household with full-time working status has a positive, significant effect on the likelihood of making an organic purchase based on both the banana and pepper Tobit and Poisson random effects model results. Male head with a "professional" occupation has a negative, significant effect on organic purchases for the household for the pepper and tomato Tobit and the Poisson model.

7.5 Seasonality Trends

As the number of produce purchases occur in higher frequency during the summer, the Poisson RE model suggests that seasonality has a strong positive effect on the number of quarterly organic household purchases. Results from the Tobit RE model suggest that seasonality effects do not explain monthly household consumption volumes of organic pepper and tomatoes as insignificant seasonality dummy coefficients are found for both models. This result could be expected for perishable produce as quantity purchased in a given month often does not vary extensively. When examining quarterly number of purchases on the other hand, seasonality trends become more pronounced.

7.6 Individual Unobserved Effects

Confirming other empirical studies with household scanner data, the presence of heterogeneity among households was found to be present within both the Tobit and Poisson random effects models that were estimated. Both models suggest the presence of unobserved heterogeneity among produce purchasing households. Using an amended likelihood ratio test similar to Lejeune (1996) when testing individual random effects, the null hypothesis of no heterogeneous effect is rejected at the 1% significance level in all instances for both types of models. Intuitively, this result was to be expected. Especially with a diverse "representative" US panel such as ACNielsen's Homescan,^{TM} the presence of unobserved effects is to be expected and was appropriately modeled.

Table 7.1: Parameter Estimates for Random Effects Tobit Model

Dependent Variable	Organic Produce Quantity Purchased (lbs)		
Independent Variable	Banana	Pepper	Tomato
Organic Price	-.571***	-.173***	-.235***
	(.072)	(.031)	(.034)
Conventional Price	.149*	.230***	.071**
	(.090)	(.042)	(.034)
# Children Under 6	.257***	.140**	-.009
	(.066)	(.062)	(.073)
# Children 6-12	.313***	.08	-.007
	(.055)	(.056)	(.06)
# Children 13-17	.406***	.137***	.123**
	(.056)	(.05)	(.057)
Extra Members	.485***	.287***	.050
	(.061)	(.058)	(.064)
ageF	-.091***	-.152***	-.047
	(.031)	(.03)	(.032)
ageM	-.012	.064*	-.013
	(.033)	(.033)	(.037)
noHeadF	-.015	-1.098***	-.306
	(.311)	(.3)	(.315)
noHeadM	-.653**	.060	-.213
	(.321)	(.308)	(.347)
dadFT	-.002	-.005	.106
	(.117)	(.106)	(.118)
momFT	.360***	.074	.258***
	(.093)	(.082)	(.095)
dadPro	-.116	-.264***	-.274***
	(.103)	(.094)	(.102)
momPro	.031	.068	-.097
	(.108)	(.098)	(.112)
dadColl	.039	-.120	.081
	(.104)	(.099)	(.108)
momColl	-.210**	-.145	.033
	(.105)	(.098)	(.107)
dadGrad	-.407***	.085	.093
	(.138)	(.127)	(.138)
momGrad	.137	.358***	.149
	(.148)	(.125)	(.144)

Standard error values in parentheses.
Asterisks indicate statistical significance at the 10%(*), 5%(**) and 1%(***) levels.

Table 7.2: Parameter Estimates for Random Effects Tobit Model-Continued

Dependent Variable	Organic Produce Quantity Purchased (lbs)		
Independent Variable	Banana	Pepper	Tomato
married	-.072	-.302	-.253
	(.206)	(.183)	(.2)
divorced	-.332**	.009	-.087
	(.157)	(.152)	(.149)
Inc20to40	-.747***	-.272	-.335*
	(.176)	(.173)	(.173)
Inc40to60	-1.248***	-.312*	-.519***
	(.185)	(.177)	(.18)
Inc60to70	-1.559***	-.150	-.564***
	(.21)	(.195)	(.2)
Inc70to100	-1.541***	-.385**	-.400**
	(.203)	(.192)	(.195)
HighInc	-1.946***	-.599***	-.560***
	(.228)	(.211)	(.217)
Q1	.218**	-.110	-.073
	(.09)	(.091)	(.089)
Q2	.298***	-.156*	-.014
	(.089)	(.093)	(.082)
Q3	.134	-.036	-.004
	(.092)	(.092)	(.083)
Constant	-1.288***	-.228	-.312
	(.347)	(.315)	(.352)
$\hat{\sigma}_\mu$	1.951***	.836***	1.696***
	(.031)	(.028)	(.041)
$\hat{\sigma}_\epsilon$	2.257***	1.044***	1.365***
	(.036)	(.038)	(.029)
Households	4285	1546	2527
nT	23556	3987	7990
Log Likelihood	-9580.06	-1884.00	-5419.18
Wald chi-square	394.92***	145.46***	104.87***

Standard error values in parentheses.
Asterisks indicate statistical significance at the 10%(*), 5%(**) and 1%(***) levels.

Table 7.3: Marginal Effect Estimates for Random Effects Tobit Model

	Banana		Pepper		Tomato	
	ME[a]	$P(y_{it}>0)^b$	ME[a]	$P(y_{it}>0)^b$	ME[a]	$P(y_{it}>0)^b$
Organic Price	-.083***	-.044***	-.026***	-.030***	-.059***	-.034***
	(.011)	(.006)	(.005)	(.005)	(.008)	(.005)
Conventional Price	.024*	.012*	.040***	.040***	.018**	.010**
	(.013)	(.006)	(.006)	(.006)	(.009)	(.005)
# Children Under 6	.037***	.020***	.021**	.024**	-.002	-.001
	(.010)	(.005)	(.009)	(.011)	(.019)	(.011)
# Children 6-12	.045***	.024***	.012	.014	-.002	-.001
	(.008)	(.004)	(.008)	(.010)	(.015)	(.009)
# Children 13-17	.059***	.031***	.020***	.024***	.031**	.018**
	(.008)	(.004)	(.007)	(.009)	(.014)	(.008)
Extra Members	.070***	.037***	.042***	.050***	.012	.007
	(.009)	(.005)	(.009)	(.010)	(.016)	(.009)
ageF	-.013***	-.007***	-.023***	-.026***	-.012	-.007
	(.005)	(.002)	(.004)	(.005)	(.008)	(.005)
ageM	-.002	-.001	.009*	.011*	-.003	-.002
	(.005)	(.003)	(.005)	(.006)	(.009)	(.005)
noHeadF	-.002	.001	-.093***	-.124***	-.071	-.043
	(.045)	(.024)	(.013)	(.019)	(.068)	(.042)
noHeadM	-.085**	-.046**	.009	.011	-.051	-.030
	(.037)	(.021)	(.048)	(.055)	(.081)	(.049)
dadFT	-.0003	.0001	.001	-.001	.027	.016
	(.017)	(.009)	(.016)	(.018)	(.030)	(.017)
momFT	.053***	.028***	.011	.012	.066***	.038***
	(.014)	(.007)	(.012)	(.014)	(.025)	(.014)
dadColl	.006	.003	-.017	-.021	.020	.012
	(.015)	(.008)	(.014)	(.017)	(.027)	(.016)
momColl	-.030**	-.016**	-.021	-.025	.008	.005
	(.015)	(.008)	(.014)	(.016)	(.027)	(.016)
dadGrad	-.055***	-.030***	-.013	-.015	.024	.014
	(.017)	(.010)	(.020)	(.023)	(.036)	(.021)
momGrad	.020	-.011	-.061**	-.068***	.039	.022
	(.022)	(.012)	(.024)	(.026)	(.039)	(.022)
dadPro	-.017	-.009	-.037***	-.044***	-.066***	-.039***
	(.015)	(.008)	(.012)	(.015)	(.024)	(.014)
momPro	-.005	-.002	.010	.012	-.024	-.014
	(.016)	(.008)	(.015)	(.017)	(.027)	(.016)

Standard error values in parentheses.
Asterisks indicate statistical significance at the 10%(*), 5%(**) and 1%(***) levels.
[a] Denotes marginal effect from McDonald and Moffitt (1980).
[b] Denotes marginal effect for the probability of being uncensored.
Marginal effect for binary variables is for a discrete change from 0 to 1.

Table 7.4: Marginal Effect Estimates for Random Effects Tobit Model-Continued

	Banana		Pepper		Tomato	
	ME^a	$P(y_{it}>0)^b$	ME^a	$P(y_{it}>0)^b$	ME^a	$P(y_{it}>0)^b$
married	-.010	-.006	-.049	-.055	-.066	-.038
	(.030)	(.016)	(.033)	(.036)	(.054)	(.030)
divorced	-.045**	-.024**	.001	.002	-.022	-.013
	(.020)	(.011)	(.023)	(.026)	(.036)	(.021)
Inc20to40	-.098***	-.053***	-.037*	-.044*	-.080**	-.048**
	(.021)	(.012)	(.021)	(.026)	(.039)	(.024)
Inc40to60	-.157***	-.086***	-.043*	-.051*	-.122***	-.073***
	(.021)	(.011)	(.022)	(.027)	(.040)	(.024)
Inc60to70	-.167***	-.095***	-.021	-.025	-.125***	-.077***
	(.016)	(.010)	(.025)	(.031)	(.039)	(.025)
Inc70to100	-.178***	-.099***	-.050**	-.061**	-.094**	-.056**
	(.019)	(.011)	(.022)	(.027)	(.043)	(.026)
HighInc	-.197***	-.113***	-.071***	-.088***	-.125***	-.077***
	(.016)	(.010)	(.020)	(.026)	(.043)	(.028)
Q1	.033**	.017**	-.016	-.019	-.018	-.011
	(.014)	(.007)	(.013)	(.015)	(.022)	(.013)
Q2	.045***	.023***	-.022	-.026	-.003	-.002
	(.014)	(.007)	(.013)	(.015)	(.021)	(.012)
Q3	.020	.010	.005	-.006	-.001	-.001
	(.014)	(.007)	(.014)	(.016)	(.021)	(.012)

Standard error values in parentheses.
Asterisks indicate statistical significance at the 10%(*), 5%(**) and 1%(***) levels.
[a] Denotes marginal effect from McDonald and Moffitt (1980).
[b] Denotes marginal effect for the probability of being uncensored.
Marginal effect for binary variables is for a discrete change from 0 to 1.

Table 7.5: Own and Cross-Price Elasticity Estimates (From Tobit RE)

Item	$\xi_d[E(y_{it}\|x_{it})]^a$ Est.	95% CI		$\xi_d[Pr(y_{it}>0)]^b$ Est.	95% CI	
Organic Raw Fresh Bananas						
Organic Price	-.260	-.325	-.195	-.210	-.262	-.158
Conventional Price	.059	-.010	.131	.048	-.008	.105
Organic Raw Fresh SW Peppers						
Organic Price	-.469	-.629	-.310	-.383	-.511	-.254
Conventional Price	.640	.482	.799	.522	.395	.649
Organic Raw Fresh Tomatoes						
Organic Price	-.296	-.380	-.212	-.224	-.288	-.160
Conventional Price	.083	.003	.163	.063	.002	.124

Elasticities evaluated at sample price means.

[a] Elasticity determined by $\frac{\partial E[y_{it}|x_{it},\mu_i]}{\partial \bar{x_{it}}} \cdot \frac{x_{it}}{E[y_{it}|x_{it},\mu_i]}$ (McDonald and Moffitt, 1980).

[b] Elasticity determined by $\frac{\partial Pr(y_{it}>0)}{\partial \bar{x_{it}}} \cdot \frac{x_{it}}{Pr[y_{it}>0]}$

Table 7.6: Parameter Estimates for Random Effects Poisson Panel Model

Dependent Variable	Quarterly Organic Produce Purchases$^{\delta}$ (Count)	
Independent Variable	Coefficient	Marginal Effect
Organic Price$^{\alpha}$	-.044*	-.015*
	(.023)	(.008)
Conventional Price$^{\alpha}$	.327***	.113***
	(.027)	(.010)
# Children Under 6	.025	.009
	(.035)	(.012)
# Children 6-12	.013	.004
	(.031)	(.010)
# Children 13-17	.106***	.037***
	(.032)	(.011)
Extra Members	.101***	.035***
	(.037)	(.013)
ageF	-.078***	-.027***
	(.02)	(.007)
ageM	.027	.009
	(.022)	(.008)
noheadF	-.36*	-.107**
	(.191)	(.049)
noheadM	.153	.056
	(.202)	(.077)
dadFT	-.004	-.001
	(.073)	(.025)
momFT	.118**	.041**
	(.058)	(.020)
dadPro	-.173***	-.057***
	(.063)	(.020)
momPro	-.068	-.023
	(.067)	(.022)
dadColl	-.015	-.005
	(.072)	(.025)
momColl	-.066	-.023
	(.07)	(.023)
dadGrad	.094	.034
	(.088)	(.033)
momGrad	.174*	.064
	(.1)	(.040)
Married	.079	.027
	(.123)	(.041)
Divorced	-.082	-.027
	(.098)	(.032)

Standard error values in parentheses.
Asterisks indicate statistical significance at the 10%(*), 5%(**) and 1%(***) levels.
$^{\alpha}$Mean Quarterly Price
$^{\delta}$ Banana, Pepper, and Tomato Organic Fresh Items

Table 7.7: Parameter Estimates for Random Effects Poisson Panel Model

Dependent Variable	Quarterly Organic Produce Purchases$^{\delta}$ (Count)	
Independent Variable	Coefficient	Marginal Effect
Inc20to40	-.201**	-.066**
	(.098)	(.030)
Inc40to60	-.441***	-.139***
	(.103)	(.030)
Inc60to70	-.362***	-.110***
	(.119)	(.031)
Inc70to100	-.454***	-.138***
	(.117)	(.031)
HighInc	-.631***	-.175***
	(.13)	(.029)
Q1	.132***	.047***
	(.041)	(.015)
Q2	.206***	.074***
	(.04)	(.015)
Q3	.208***	.076***
	(.04)	(.016)
Constant	-.857***	
	(.207)	
lnalpha	1.134***	
	(.035)	
alpha	3.10***	
	(.109)	
Households	5181	
nT	20020	
Log Likelihood	-11377.88	
Wald chi-square	281.18***	

Standard error values in parentheses.
Asterisks indicate statistical significance at the 10%(*), 5%(**) and 1%(***) levels.

CHAPTER 8

SUMMARY AND CONCLUSIONS

The primary purpose of this thesis is to provide increased understanding of organic produce demand at the household level. Given the few studies that include direct purchase and price information, this study aims to utilize the panel nature of the ACNielsen Homescan^{TM} data to generate empirical results in the current literature. In estimating price elasticities, demographic effects, household composition influences, and seasonality effects while controlling for unobserved heterogeneity, the empirical study has yielded some interesting results. This chapter provides summary and conclusions to these empirical findings and offers suggestions for further research.

8.1 Elasticities and Price Effects

Of particular interest to agricultural economists and organic marketers are price elasticity measures for organic produce. Some organic producers believe that demand for organic produce is rather unresponsive to price changes. The results of this study corroborate the belief that organic produce purchases are inelastic in response to both organic and conventional prices. Shown in table 7.5 are own and cross-price elasticity estimates for fresh bananas, peppers, and tomatoes.

Own-price elasticity results suggest that reductions in retail prices for organic produce due to increased competition between producers, scale economies in production, lower costs of production, or other factors would have a less than equal positive influence on organic produce market shares. Also, cross-price elasticity results imply that organic quantity demanded is found to have a less than proportionate responsiveness to percent changes in conventional produce prices. Park and Lohr (1996) found similar elasticity estimates when examining organic carrot, broccoli, and lettuce wholesale purchase elasticities.

In comparison to market-level elasticities measured for frozen organic vegetables by Glaser and Thompson (1998), own-price elasticity measures were found to be significantly higher for organic frozen vegetables at the aggregate level as opposed to those at the household level in this study. This result was to be expected as price responses are more pronounced on the aggregate. Aggregate cross-price elasticity values estimated by Glaser and Thompson were relatively similar in most cases to those found on the household level in this study which was an interesting result.

8.2 Income Effects

Income effects on the likelihood of an organic purchase versus a conventional purchase are quite interesting. Despite price premiums for organic foods, higher household incomes do not indicate a higher likelihood of organic purchases. For all models, the likelihood of an organic purchase generally diminishes at a statistically significant level across increasing income brackets. As households move to higher income categories, the models predict lower numbers of organic purchases and quantity purchased.

These income results contradict some related studies' findings, but corroborate others. These results suggest that income related lifestyles, values, and concerns have key roles to play in organic food purchasing decisions. As mentioned by Thompson (1998), "some consumer segments with relatively lower incomes seem to be more entrenched buyers of organic products and tend to have shopped for organics at retail outlets" (p 1115). Multiple studies such as Ott (1990), Jolly (1991), and Buzby and Skees (1994) found that household income was not a significant explanatory variable for organic purchases or reductions in pesticide exposure.

8.3 Household Composition Effects

A number of household composition factors are revealed to be important determinants of fresh organic produce demand. Many interesting aspects such as household size, the presence of children, and marital status are found to have significant effects on the likelihood of an organic purchase.

All empirical results suggest that an increase in the number of individuals in the household is found to have a positive effect on organic purchases with statistical significance. These results corroborate Huang (1996) where willingness to buy organic produce increased with household size in Georgia.

The presence of children, especially older children, was generally positive and significant in influencing organic purchase activity. This result corroborates Thompson and Kidwell (1998) where the number of persons in the consumer's household under of the age of eighteen had a positive significant probability of purchasing organic produce.

The above result could be explained simply by the fact that the caloric intake of teenagers is larger than that of younger children; also, other household members who are adults consume more calories per day than children generally. Further research regarding caloric intake volumes by household member's age could corroborate the results of this study.

In this study, the incidence of an existing marriage or a divorce had almost no statistically significant effect on the likelihood of an organic purchase. Marriage status has not been examined in many previous studies. This study provides no evidence for potential effects of martial status on the propensity to purchase organics.

8.4 The Role of the Female Head

Shown empirically are many interesting household dynamics that suggest that living and employment situations of household male and female heads have differing effects on organic purchase decisions. The results of the model suggest that attributes related to the female head of household have significant effects on the likelihood of organic produce purchases while male head of household attributes are far less significant.

Intuitively, these results are to be expected as many studies such as Jolly (1991) and Buzby and Skees (1994) verify that females are more willing to pay for organic produce or decreased pesticide exposure than men. If the female head of household does the majority of the grocery shopping for the entire household (the majority of two-partner US households), these effects logically have important influence on household organic consumption patterns.

8.4.1 Age

The age bracket in which the female head of the household falls into shows a statistically significant effect on purchase activity while male age effects are primarily insignificant across all three types of organic produce items examined. Every incremental 5 year age bracket in which the female head of the household falls into has a negative and significant effect on organic purchases. Similar age effects shown by Jolly (1991), Misra et al. (1991), and Buzby and Skees (1994) suggest that younger individuals certainly have a higher propensity to consume organic over conventional produce. The model in this study, however, suggests that the age of female heads of household as compared to male age effects has a much stronger influence statistically on household organic purchases.

8.4.2 Education

In terms of educational effects on propensity to purchase organic over conventional produce, the variables that designated male and female heads with college degrees and the variables that designated a graduate degree were mainly insignificant. In the few models where the results were statistically different from zero, they had a negative effect on organic produce purchases. This result was similar to the general findings from Misra et al. (1991) where college-educated consumers were more price elastic to organic produce consumption.

8.4.3 Occupational Effects

When examining full-time working status, interesting household elements suggest that non-traditional household structure encourages organic over conventional purchases. Female head of household with full-time working status has a positive, significant effect on the likelihood of making an organic purchase based on both the Tobit and Poisson random effects model results. These empirical findings suggest that households with "non-homemaker" females who find work outside the home have a higher likelihood to purchase organic items either for themselves or for their family.

Only 43% of women were in the labor force in 1970 compared to 59% in 1994 (Kinsey and Senauer, 1996) indicating labor force participation of women continues to increase in the US. The results from this study suggest that the continuing trend of labor force participation of women has a strong influence on organic purchase decisions and could be a source of sustained organic produce demand in the future.

Male head of household full-time employment did not have a significant effect on organic purchases. For every organic item examined, male full-time employment has no statistically significant influence on organic purchases. Interestingly, on the other hand, households with a male head with a "professional" occupation have a negative, significant effect on organic purchases for the household while female head occupation type is not significant. While a "professional" occupational description is quite vague, the result is interesting nonetheless.

8.4.4 Race Effects

Racial effects on propensity to purchase organic produce shown in Tables A.2 and A.4 in the appendix spawned interesting results. Though households that are designated as "White" consume the majority of fresh produce purchases, minority races consume organic produce versus conventional items at a much higher overall percentage than Whites. Based on the empirical results of this study, Whites showed a lower likelihood to purchase organic items over conventional purchases. For all items analyzed, a household that was designated as White had a negative statistically significant effect on organic purchases while households designed as "Black" and "Hispanic" showed a positive propensity to choose organic produce over conventional.

As mentioned in Chapter 4, the ACNielsen Random Weight panel is not representative of the racial composition of the US population. Therefore, results in regards to racial effects of organic purchases should be interpreted with caution as race had very little significance in explaining number of purchases of organic items.

8.5 Suggestions for Further Research

As mentioned, estimating a dynamic model that allows for unobserved individual effects in a censored context remains at the frontier of the econometric literature on the subject. Clearly such a model would bear some benefit to the empirical results but at too high a cost for this study at the present time. Perhaps as the literature on the subject expands, estimation of this model will be possible at one point and all the benefits of a GMM estimator can be realized.

As high quality household scanner data steadily become available, the potential for more interesting empirical analysis is possible. As data resources such as ACNielsen continue to increase their national exposure and household participation, more representative results could be estimated. As mentioned in Chapter 4, the Homescan^{TM} "random-weight" panel differs slightly in composition from the representative US population. Analysis that incorporates weighting methods could be explored in future research to avoid over or underemphasizing effects that certain demographical groups have on results.

It is important to remember that this analysis involved only a small proportion of the purchasing activity that the Homescan panel contains. Examining only 3 distinct types of conventional and organic fresh produce facilitated straightforward comparison of price and purchasing activity. There are, however, many other different types and forms of organic and conventional food items described in the data.

Prepackaged, prepared, and juice items with a “organic” or “non-organic” label still exist among others in the panel and could be further studied. Comparison to ensure an accurate association between items is quite time intensive and requires a fair deal of due diligence. Nevertheless, ambitious researchers with moderate data programming ability would have much empirical study to contribute not just to the directly related literature, but to other household demand studies as well.

APPENDIX A

Preliminary Regression Results

Table A.1: Preliminary Random Effects Tobit Model

Dependent Variable	Organic Produce Quantity Purchased (lbs)		
Independent Variable	Banana	Pepper	Tomato
Constant	-2.14***	-1.13***	-.192
	(.464)	(.435)	(.434)
Organic Price	-.599***	-.225***	-.232***
	(.075)	(.039)	(.034)
Conventional Price	.161*	.398***	.070**
	(.095)	(.051)	(.034)
Household Size	.350***	.201***	.082*
	(.049)	(.046)	(.046)
ageF	-.079**	-.219***	-.052*
	(.035)	(.036)	(.031)
ageM	-.004	.106***	-.015
	(.037)	(.038)	(.036)
noKids	.161	.413***	.211
	(.141)	(.136)	(.131)
noHeadF	.373	-1.37***	-.314
	(.353)	(.348)	(.319)
noHeadM	-.242	.624*	-.196
	(.358)	(.349)	(.341)
dadFT	.082	.011	.121
	(.128)	(.120)	(.115)
momFT	.379***	.170*	.254***
	(.103)	(.094)	(.093)
dadColl[a]	.207*	-.198**	.024
	(.110)	(.101)	(.097)
momColl[a]	-.242**	-.296***	-.032
	(.109)	(.101)	(.098)

Standard error values in parentheses.
Asterisks indicate statistical significance at the 10%(*), 5%(**) and 1%(***) levels.
[a] Variable coding error occurred which could affect results.

Table A.2: Preliminary Random Effects Tobit Model-Continued

Dependent Variable	Organic Produce Quantity Purchased (lbs)		
Independent Variable	Banana	Pepper	Tomato
dadPro	-.121	-.176*	-.230**
	(.110)	(.11)	(.097)
momPro	-.041	-.088	-.034
	(.115)	(.106)	(.104)
married	-.064	-.088	-.238
	(.231)	(.213)	(.200)
divorced	-.393**	.037	-.072
	(.180)	(.176)	(.150)
Hispanic	.316*	.371**	-.140
	(.175)	(.148)	(.157)
White	-.710***	-.511**	-.405*
	(.216)	(.199)	(.214)
Black	.585**	.305	-.349
	(.247)	(.233)	(.252)
Oriental	-.214	-.174	-1.13***
	(.341)	(.338)	(.361)
Inc20to40	-.758***	-.307	-.332*
	(.186)	(.201)	(.174)
Inc40to60	-1.25***	-.390*	-.500***
	(.194)	(.203)	(.180)
Inc60to70	-1.49***	-.130	-.539***
	(.222)	(.226)	(.199)
Inc70to100	-1.60***	-.432**	-.351*
	(.213)	(.220)	(.194)
HighInc	-1.96***	-.435*	-.449**
	(.237)	(.232)	(.212)
Q1	.222**	-.164	-.071
	(.095)	(.119)	(.089)
Q2	.304***	-.171	-.014
	(.095)	(.122)	(.082)
Q3	.130	.030	-.001
	(.097)	(.120)	(.083)
$\hat{\sigma}_\mu$	2.34***	.026***	1.69***
	(.040)	(.0006)	(.041)
$\hat{\sigma}_\epsilon$	2.27***	1.65***	1.37***
	(.036)	(.057)	(.029)
Households	4285	1546	2527
nT	23556	3987	7990
Log Likelihood	-9256.41	-1879.89	-5416.35
Wald chi-square	436.71***	219.95***	108.89***

Table A.3: Preliminary ME Estimates for Random Effects Tobit Model

	Banana		Pepper		Tomato	
	ME[a]	$P(y_{it} > 0)^b$	ME[a]	$P(y_{it} > 0)^b$	ME[a]	$P(y_{it} > 0)^b$
Organic Price	-.087***	-.042***	-.029***	-.029***	-.058***	-.034***
	(.011)	(.005)	(.005)	(.005)	(.009)	(.005)
Conventional Price	.024*	.012*	.051***	.051***	.017**	.010**
	(.014)	(.006)	(.007)	(.006)	(.008)	(.005)
Household Size	.051***	.024***	.025***	.026***	.021*	.020*
	(.007)	(.003)	(.006)	(.006)	(.012)	(.007)
ageF	-.011**	-.006**	-.028***	-.028***	-.014*	-.007*
	(.005)	(.002)	(.005)	(.004)	(.008)	(.003)
ageM	-.001	-.0003	.014***	.014**	-.004	-.002
	(.005)	(.003)	(.005)	(.005)	(.009)	(.005)
noKids	.023	.011	.049***	.050***	.052*	.030
	(.020)	(.009)	(.015)	(.015)	(.031)	(.019)
noHeadF	.058	.027	-.099***	-.111***	-.073	-.044
	(.060)	(.027)	(.012)	(.016)	(.068)	(.043)
noHeadM	-.034	-.017	.100	.092	-.047	-.028
	(.048)	(.024)	(.068)	(.058)	(.079)	(.048)
dadFT	.012	.006	.001	.001	.030	.018
	(.018)	(.009)	(.016)	(.015)	(.029)	(.017)
momFT	.056***	.027***	.022*	.022*	.065***	.037***
	(.015)	(.007)	(.012)	(.012)	(.024)	(.014)
dadColl[c]	.031*	.015*	-.024**	-.025**	.006	.003
	(.017)	(.008)	(.012)	(.012)	(.025)	(.014)
momColl[c]	-.034**	-.017**	-.036***	-.036***	-.008	-.005
	(.015)	(.007)	(.011)	(.012)	(.024)	(.014)
dadPro	-.017	-.008	-.022*	-.022*	-.056**	-.033**
	(.016)	(.008)	(.012)	(.012)	(.023)	(.014)
momPro	-.006	-.003	.011	.011	-.008	-.005
	(.017)	(.008)	(.013)	(.013)	(.026)	(.015)
married	-.009	-.005	-.012	-.011	-.062	-.035
	(.034)	(.016)	(.028)	(.028)	(.054)	(.030)
divorced	-.053**	-.026**	.005	.005	-.018	-.010
	(.022)	(.011)	(.023)	(.023)	(.036)	(.022)

Standard error values in parentheses.
Asterisks indicate statistical significance at the 10%(*), 5%(**) and 1%(***) levels.
[a] Denotes marginal effect from McDonald and Moffitt (1980).
[b] Denotes marginal effect for the probability of being uncensored.
Marginal effect for binary variables is for a discrete change from 0 to 1.
[c] Variable coding error occurred which could affect results.

Table A.4: Preliminary ME Estimates for Random Effects Tobit Model-Continued

	Banana		Pepper		Tomato	
	ME[a]	$P(y_{it} > 0)^b$	ME[a]	$P(y_{it} > 0)^b$	ME[a]	$P(y_{it} > 0)^b$
Hispanic	.049*	.023*	.056**	.052**	-.034	-.020
	(.029)	(.013)	(.026)	(.023)	(.037)	(.022)
White	-.116***	-.054***	-.079**	-.074**	-.111*	-.062*
	(.039)	(.018)	(.036)	(.035)	(.064)	(.034)
Black	.095**	.044**	.045	.042	-.081	-.049
	(.045)	(.020)	(.038)	(.035)	(.053)	(.033)
Oriental	-.030	-.014	-.020	-.021	-.206***	-.136***
	(.045)	(.022)	(.037)	(.038)	(.044)	(.033)
Inc20to40	-.100***	-.050***	-.036*	-.037*	-.079**	-.047**
	(.022)	(.011)	(.021)	(.022)	(.039)	(.024)
Inc40to60	-.159***	-.079***	-.046**	-.047**	-.118***	-.071***
	(.022)	(.011)	(.022)	(.023)	(.040)	(.024)
Inc60to70	-.166***	-.086***	-.016	-.016	-.120***	-.074***
	(.018)	(.010)	(.026)	(.027)	(.039)	(.025)
Inc70to100	-.186***	-.095***	-.049**	-.050**	-.083*	-.050*
	(.020)	(.010)	(.022)	(.023)	(.043)	(.027)
HighInc	-.203***	-.107***	-.049**	-.050**	-.103**	-.062**
	(.017)	(.009)	(.023)	(.024)	(.044)	(.028)
Q1	.033**	.016**	-.020	-.020	-.018	-.010
	(.014)	(.007)	(.014)	(.014)	(.022)	(.013)
Q2	.046***	.022***	-.021	-.021	-.004	-.002
	(.015)	(.007)	(.014)	(.015)	(.021)	(.012)
Q3	.019	.009	.004	-.004	-.0004	-.0002
	(.015)	(.007)	(.015)	(.015)	(.021)	(.012)

Standard error values in parentheses.
Asterisks indicate statistical significance at the 10%(*), 5%(**) and 1%(***) levels.
[a] Denotes marginal effect from McDonald and Moffitt (1980).
[b] Denotes marginal effect for the probability of being uncensored.
Marginal effect for binary variables is for a discrete change from 0 to 1.

Table A.5: Preliminary Random Effects Poisson Panel Model

Dependent Variable	Quarterly Organic Produce Purchases$^{\delta}$ (Count)		
Independent Variable	Coefficient	Independent Variable	Coefficient
Constant	-.973***	*Hispanic*	.089
	(.245)		(.094)
Organic Price$^{\alpha}$	-.042*	*White*	-.203*
	(.023)		(.107)
Conventional Price$^{\alpha}$	.330***	*Black*	.009
	(.026)		(.137)
Household Size	.068**	*Oriental*	-.212
	(.027)		(.202)
ageF	-.073***	*Inc20to40*	-.192**
	(.020)		(.097)
ageM	.028	*Inc40to60*	-.412***
	(.022)		(.103)
noKids	.109	*Inc60to70*	-.329***
	(.079)		(.119)
noHeadF	-.278	*Inc70to100*	-.398***
	(.194)		(.115)
noHeadM	.215	*HighInc*	-.575***
	(.202)		(.127)
dadFT	.001	*Q1*	.135***
	(.073)		(.041)
momFT	.119**	*Q2*	.206***
	(.058)		(.040)
dadColl	-.038	*Q3*	.213***
	(.063)		(.040)
momColl	-.094	α	3.09***
	(.065)		(.108)
dadPro	-.157**		
	(.062)		
momPro	-.045		
	(.065)		
married	.010		
	(.124)		
divorced	-.081		
	(.099)		
Households	5181		
nT	20020		
Log Likelihood	-11377.12		
Wald chi-square	282.37***		

Standard error values in parentheses.
Asterisks indicate statistical significance at the 10%(*), 5%(**) and 1%(***) levels.
$^{\alpha}$Mean Quarterly Price
$^{\delta}$ Banana, Pepper, and Tomato Organic Fresh Items

REFERENCES

Baker, G. A. and P. Crosbie (1993). Measuring Food Safety Preferences: Identifying Consumer Segments. *Journal of Agricultural and Resource Economics*, **18**(2), pp. 277–287.

Baltagi, B., G. Bresson, and A. Pirotte (2005). *Joint LM Test for Heteroscedasticity in a One-way Error Component Model.* Working Paper, Texas A&M University, College Station, TX.

Baltagi, B. H. (2005). *Econometric Analysis of Panel Data (Third Edition).* John Wiley & Sons, Inc., West Sussex, England.

Baron, S. and A. Lock (1995). The Challenges of Scanner Data. *The Journal of the Operational Research Society*, **46**(1), pp. 50–61.

Benson, J. T., F. Breidth, and J. R. Schroeter (2002). Television and Beef Demand: Bayesian Inference in a Random Effects Tobit Model. *Canadian Journal of Agricultural Economics*, (50), pp. 201–219.

Brewerton, T. D., D. D. Krahn, T. A. Hardin, T. A. Wehr, and N. E. Rosenthal (1994). Findings from the Seasonal Pattern Assessment Questionnaire in Patients with Eating Disorders and Control Subjects: Effect of Diagnosis and Location. *Psychiatry Res*, (52), pp. 71–84.

Bucklin, R. E. (1999). Commercial Use of UPC Scanner Data: Industry and Academic Perspectives. *Marketing Science*, **18**(3), pp. 247–273.

Butler, J. S. and R. Moffitt (1982). A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model. *Econometrica*, **50**(3), pp. 761–64.

Buzby, J. C. and J. R. Skees (1994). Consumers Want Reduced Exposure to Pesticides on Food. *Food Review*, **17**(2), pp. 19–26.

Glaser, L. K. and G. D. Thompson (1998). Demand for Organic and Conventional Frozen Vegetables. In *Paper Presented at Western Agricultural Economics Association Annual Meetings*, pp. 1–13.

Greene, W. H. (2003). *Econometric Analysis (Fifth Edition).* Prentice-Hall, Inc., Upper Saddle River NJ.

Harless, D. W. and G. E. Hoffer (2002). Do Women Pay More for New Vehicles? Evidence from Transaction Price Data. *The American Economic Review*, **92**(1), pp. 270–279.

Hausman, J., B. H. Hall, and Z. Griliches (1984). Econometric Models for Count Data with an Application to the Patents-R & D Relationship. *Econometrica*, **52**(4), pp. 909–938.

Holly, A. and L. Gardiol (2000). A Score Test for Individual Heteroscedasticity in a One-Way Components Model, Chapter 10. *Panel Data Econometrics: Future Directions*, pp. 199–211.

Honoré, B. E. (1993). Orthogonality Conditions for Tobit Models with Fixed Effects and Lagged Dependent Variables. *Journal of Econometrics*, **59**, pp. 35–61.

Hsiao, C. (2003). *Analysis of Panel Data (Second Edition)*. Cambridge University Press, Cambridge, UK.

Hsiao, C. (2005). Why Panel Data? *IEPR WORKING PAPER 05.33*, pp. 1–19.

Hu, L. (2002). Estimation of a Censored Dynamic Panel Data Model. *Econometrica*, **70**(6), pp. 2499–2517.

Huang, C. L. (1996). Consumer Preferences and Attitudes Towards Organically Grown Produce. *Journal of Agricultural and Resource Economics*, **23**, pp. 331–342.

Jensen, H. H. and J. R. Schroeter (1992). Television Advertising and Beef Demand: An Econometric Analysis of "Split-Cable" Household Panel Scanner Data. *Canadian Journal of Agricultural Economics*, (40), pp. 271–294.

Jolly, D. A. (1991). Determinants of Organic Horticultural Products Consumption Based on a Sample of California Consumers. *Acta Horticulturae*, **295**, pp. 141–148.

Kinsey, J. and B. Senauer (1996). Consumer Trends and Changing Food Retailing Formats. *American Journal of Agricultural Economics*, **78**(5), pp. 1187–1191.

Krissoff, B. (1998). Emergence of U.S. Organic Agriculture: Can We Compete? Discussion. *American Journal of Agricultural Economics*, **80**(5), pp. 1130–1133.

Kuhn, H. W. and A. W. Tucker (1951). Nonlinear programming. In Neyman, J. (ed.) *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 481–492. University of California Press, Berkeley, California.

Lejeune, B. (1996). A full heteroscedastic one-way error components model allowing for unbalanced panel: Pseudo-maximum likelihood estimation and specification testing. In *CORE Discussion Paper 9606, Universite Catholique de Louvain*, pp. 1–28.

McDonald, J. and R. Moffitt (1980). The Uses of Tobit Analysis. *Review of Economics and Statisitics*, **82**, pp. 318–321.

Misra, S. K., C. L. Huang, and S. L. Ott (1991). Consumer Willingness to Pay for Pesticide-Free Fresh Produce. *Western Journal of Agricultural Economics*, **16**(2), pp. 218–227.

National Research Council of the National Academies (2005). *Improving Data to Analyze Food and Nutrition Policies.* The National Academies Press, Washington, DC.

Nerlove, M. (2002). *Essays in Panel Data Econometrics.* Cambridge University Press, Cambridge, UK.

Ott, S. L. (1990). Supermarket Shoppers Pesticide Concerns and Willingness to Purchase Certified Pesticide-Residue-Free Fresh Produce. *Agribusiness*, **6**(6), pp. 593–602.

Park, T. A. and L. Lohr (1996). Supply and Demand Factors for Organic Produce. *American Journal of Agricultural Economics,*, **78**(3), pp. 647–655.

The Packer (2007). *Fresh Trends 2007.* Vance publishing Corp., Overland Park KS.

Thompson, G. D. (1998). Consumer Demand for Organic Foods: What We Know and What We Need to Know. *American Journal of Agricultural Economics*, **80**(5), pp. 1113–1118.

Thompson, G. D. and J. Kidwell (1998). Explaining the Choice of Organic Produce: Cosmetic Defects, Prices, and Consumer Preferences. *American Journal of Agricultural Economics*, **80**(2), pp. 277–287.

Thompson, G. D. and P. N. Wilson (1999). Market Demands for Bagged, Refrigerated Salads. *Journal of Agricultural and Resource Economics*, **24**(2), pp. 463–481.

Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica*, **26**(1), pp. 24–36.

US Census Bureau (2001). Census 2000 Briefs and Special Reports. Available from http://www.census.gov/population/www/cen2000/briefs.html.

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

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, MA.