THE PEANUTS AND TREE NUTS MARKET IN THE UNITED STATES: DEMAND ANALYSIS, VARIETY SEEKING, HOUSEHOLD CHOICES, AND CONSUMER PROFILING

A Dissertation

by

GUO CHENG

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

DOCTOR OF PHILOSOPHY

Chair of Committee,	Oral Capps, Jr.
Co-Chair of Committee,	Senarath Dharmasena
Committee Members,	Ximing Wu
	Yan Liu
Head of Department,	Mark L. Waller

August 2019

Major Subject: Agribusiness and Managerial Economics

ABSTRACT

Nuts industry is a critical component of the U.S. agricultural economy. Given the health benefits, peanuts and tree nuts are purchased by consumers on a daily basis in the United States. Our study pertains to eight types of nuts products, peanuts, pecans, almonds, cashews, walnuts, macadamia nuts, pistachios, and mixed nuts.

First, we looked at the market level demand for these products using demand system model by giving attention to price and expenditure elasticities. Meanwhile, mixed nut is a unique category that enables us to test consumer's variety-seeking behavior, empirically. Pre-commitment levels are also included to examine the level of consumption of nuts by U.S. households regardless of price and income.

Second, we examined the household-level choices of respective nuts using pooled probit model and profiled U.S. nuts consumers using socio-demographic factors. Age of household head, education, presence of children, race, and ethnicity were found to be driving forces of household-level demand for various nuts. We provided insight for stakeholders in terms of market segmentation strategy, targeting consumers.

Third, due to the data-censoring issues resting in our dataset, a dynamic unobserved effects Tobit panel was used to estimate the demand for various nuts, to investigate the effects of socio-demographic factors, to address sample selection bias, and to control for unobserved heterogeneity.

The main findings speak to the nuts market in terms of pricing strategies, packaging strategies, and market segmentation. Most peanuts and tree nuts was

estimated to be inelastic demand, in which increasing price would be appropriate for purveyors and producers to generate more revenues. Peanuts should be packaged and mixed with other products in order to meet consumers' variety-seeking behaviors and almonds are better to be sold individually. Meanwhile, these companies are advised to target households with older head, higher educations, more members, larger income level, and without children. In detailed category, the producers of walnuts and cashews is better to spend efforts on targeting households locating in central area while macadamia nuts have a better market in Pacific area. As such, we provide more up-to-date and thorough analysis of the demand for peanuts and tree nuts, presently lacking in the extant literature.

DEDICATION

This dissertation is dedicated to my dear parents.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Oral Capps, Jr. and co-chair, Dr. Senarath Dharmasena for their tireless support and help not only on my dissertation but also the knowledge and life experience provided. Without their guidance and mentoring, I would not be able to accomplish this much. I also thank Dr. Capps and Dr. Dharmasena for their efforts of assisting me with the job hunting. I look forward to strengthening the connections with them even as I am stepping into the next stage of my career.

Additionally, I would like to thank my committee members, Dr. Ximing Wu, and Dr. Yan Liu, for their guidance and support throughout the course of this research and my graduate career here at Texas A&M University.

I would like to extend my sincere thanks to Dr. David J. Leatham and Dr. Frederick O. Boadu for providing the assistantship and chances of being an undergrad course instructor. The invaluable two years teaching experiences speak the volume.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to my mother and father for their encouragement, support, and love.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Dr. Oral Capps, Jr. [advisor – chair], Dr. Senarath Dharmasena [advisor-co-chair], Dr. Ximing Wu of the Department of Agricultural Economics, and Dr. Yan Liu of the Department of Marketing.

The data analyzed was provided by Professor Oral Capps Jr. and The Nielsen Company (US), LLC. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of Nielsen.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a graduate assistantship from the Department of Agricultural Economics at Texas A&M University.

NOMENCLATURE

AMEs	Average Marginal Effects
AIDS	Almost Ideal Demand System
GAIDS	Generalized Almost Ideal Demand System
GQUAIDS	Generalized Quadratic Almost Ideal Demand System
LA/AIDS	Linear Approximated Almost Ideal Demand System
LA/QUAIDS	Linear Approximated Quadratic Almost Ideal Demand System
QUAIDS	Quadratic Almost Ideal Demand System

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Page

1. INTRODUCTION

In this dissertation, three studies are undertaken. They are independent on one hand and are also related on the other hand via common thread, the data. We looked at the purchases of peanuts and tree nuts in the United States through the angles of market-level demand, household-level choices, and household-level purchases. In order to meet the needs of these studies, consumer level scanner data was used and constructed into three different structures, time-series, cross-sectional, and panel. These studies are summarized here.

In Study I, entitled, “Incorporating Variety and Pre-commitments in Demand Interrelationships for Nut Products in the United States”, we looked at the market level demand for these products using demand system model by giving attention to price and expenditure elasticities. The data was derived from Nielsen Homescan for calendar year 2004 to 2015. We aggregated the household scanner data to market level with a time-series data structure of 144 observations for nine nut categories. Meanwhile, mixed nut is a unique category that enables us to test consumer’s variety-seeking behavior empirically. A variety-seeking index was included to control for such behavior. Pre-commitment levels were also included to examine the level of consumption that U.S. households do not respond to either price or income. Main findings indicate variety-seeking behavior only for peanuts and negative pre-commitment for various nuts. Most peanuts and tree nuts was estimated to be inelastic demand.

In Study II, entitled, "Household Choice to Purchase Peanuts and Tree Nuts in the United States: Evidence from the Nielsen Homescan Panel", we examined the household-level choices of respective nuts using pooled probit model and profiled U.S. consumers using socio-demographic factors. This study focused on the data from calendar year 2015. Around sixty thousand households reported their purchases regarding peanuts and tree nuts purchases on a quarterly basis. A binary choice model was used to model households' decision to purchase nuts. Demographics factors were also included. Age of household head, education, presence of children, race, and ethnicity were found to be driving forces of household-level demand for various nuts. We provided insight for stakeholders in terms of market segmentation strategy targeting consumers.

In Study III, entitled, "Household Demand for Peanuts and Tree Nuts in the United States: A Dynamic Panel Tobit Analysis", due to the data-censoring issues resting in our dataset, a dynamic unobserved effects Tobit panel was used to estimate the demand for various nuts, to investigate the effects of socio-demographic factors, to address sample selection bias, and to control for unobserved heterogeneity. The data used in this study were derived from Nielsen Homescan Panel for calendar years 2009 to 2015. Common households who stay in the sample throughout the seven years were examined. We also test the habitual patterns of households' purchase patterns by including their purchases from previous time periods. The demand for various nuts was found to be inelastic while numerous substitute and complimentary relationships were revealed. Age, education, and race were found to drive the demand of various nuts at different time of a year and at different locations.

2. STUDY I: INCORPORATING VARIETY AND PRE-COMMITMENTS IN DEMAND INTERRELATIONSHIPS FOR NUT PRODUCTS IN THE UNITED STATES

2.1. Introduction

About 40% of U.S. adults consume nuts on a regular basis (Nielsen, Kit, and Ogden, 2014). Nuts provide high energy and contain more dietary fiber, vitamins, minerals and unsaturated fat compared to other salted snacks. Various studies have confirmed the association between the consumption of nuts and health benefits. King et al. (2008) revealed an inverse association between frequency of nut consumption and body mass index (BMI). Kris-Etherton (2008) confirmed the benefits of tree nuts and peanuts in preventing coronary heart disease. Additionally, O'Neil et al. (2010) found that tree nut consumption improved nutrient intake and diet quality of U.S. adults. Further, tree nuts and peanuts have been recommended to be part of daily intakes of children and adults, replacing other snack foods (Rehm and Drewnowski, 2017). In the latest Dietary Guidelines for Americans 2015-2020, nuts are included in the spectrum of healthy foods.

Annual per capita consumption of peanuts and tree nuts in the United States has been on the rise (Figure 2-1). In 2016, the per capita consumption of tree nuts was 4.7 pounds, up from 1.8 pounds in 1970, and the per capita consumption of peanuts was 7.2 pounds, up from 5.7 pounds in 1970. Tree nuts include almonds, Brazil nuts, cashews, chestnuts, filberts/hazelnuts, macadamia nuts, pecans, pistachios, and walnuts.

Per capita consumption of specific tree nuts over the period 1970 to 2016 is exhibited in Figure 2-2. Over this period, on average the per capita consumption of tree nuts was as follows: (1) almonds—0.83 pounds; (2) filberts/hazelnuts—0.06 pounds; (3) macadamia nuts—0.09 pounds; (4) pecans—0.45 pounds; (5) pistachios—0.13 pounds; (6) walnuts—0.45 pounds; and (7) other tree nuts (defined as Brazil nuts, cashews, chestnuts, and pine nuts)—0.64 pounds. The dominant tree nuts in terms of per capita consumption are almonds, pecans, walnuts, and other tree nuts. In 2016, the total crop value of nuts was as follows: almonds $5.16 billion, hazelnuts $118.8 million, pecans $696.8 million, walnuts $1.24 billion, macadamia nuts $42.0 million, and pistachios $1.51 billion.

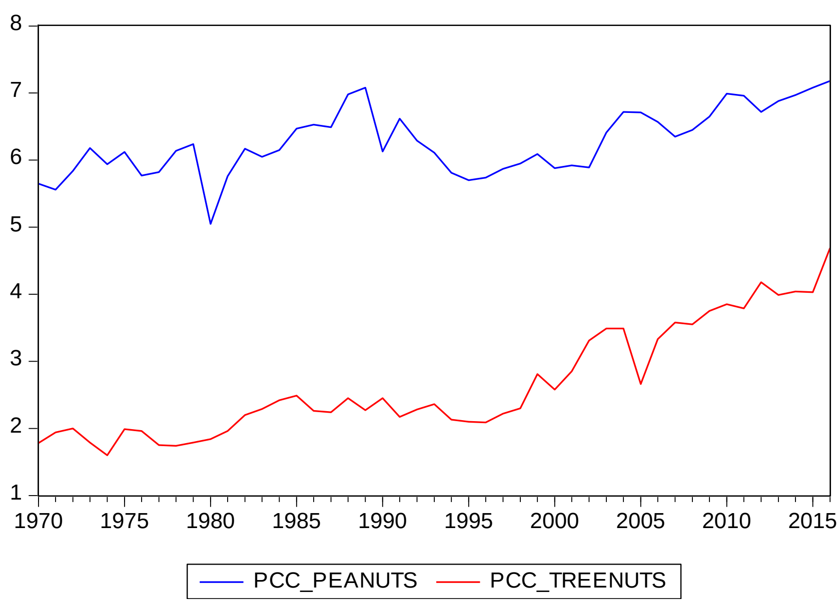


Figure 2-1 Annual U.S. per Capita Consumption (PCC) of Peanuts and Tree Nuts: 1970-2016

Source: Economic Research Service per capita availability data for peanuts and various issues of the *Fruit and Nuts Yearbook* for tree nuts.

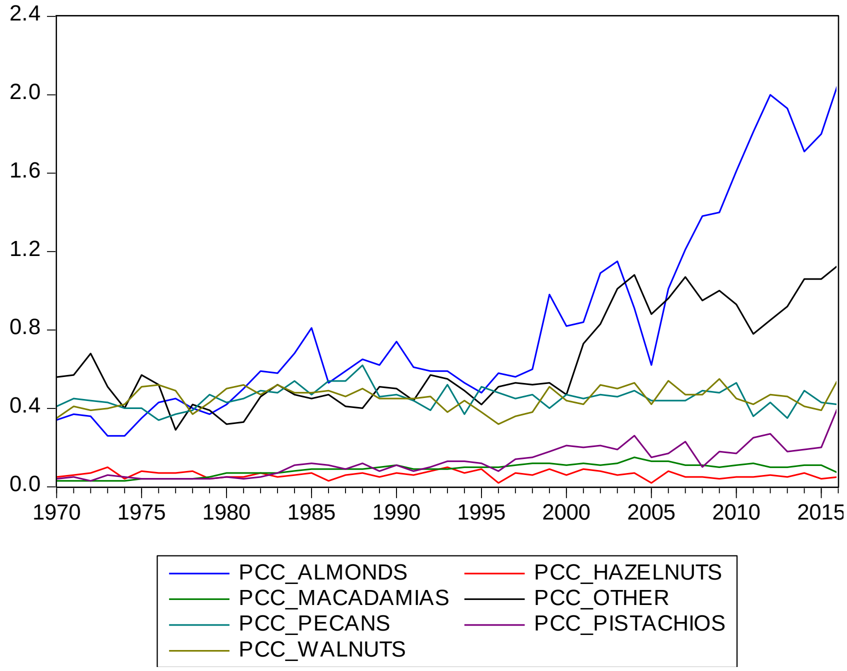


Figure 2-2 Annual U.S. Per Capita Consumption (PCC) of Specific Tree Nuts: 1970-2016
Source: USDA *Fruit and Nuts Yearbook*, various issues.

As shown in Figure 2-2, peanuts and tree nuts are an important part of U.S. consumers' diet in terms of healthy snacks. But few studies on the demand interrelationships among nut products exist in the economic literature. Moreover, as pointed out by Jevons (1871), consumers desire variety in the purchases of different products and brands. The nuts market provides an opportunity for empirically testing variety-seeking behavior. The presence of mixed nuts points to this variety issue in the nuts market. At present, to the best of our knowledge, no studies exist in the literature dealing with variety-seeking behavior in the nuts market. Additionally, we examine the level of pre-commitment of nuts purchased by households. That is to say that households might commit to purchase nuts in order to satisfy their needs of fiber intakes and seasonal consumption. In this way, we address components of demand not sensitive to the influences of income and prices (Bollino, 1987; Tonsor and Marsh, 2007; Hovhannisyan and Gould 2011; Hovhannisyan and Gould 2014) also known as non-discretionary consumption.

The objectives of this research are: (1) to estimate demand interrelationships for peanuts and tree nuts using the Linearized and Generalized Almost Ideal Demand System (LA/GAIDS) model; (2) to calculate own-price, cross-price, expenditure and income elasticities for peanuts and tree nuts; (3) to determine the presence/absence of variety-seeking behavior in the nuts market; and (4) to determine the level of pre-commitment for peanuts and tree nuts. The nut categories in this analysis correspond to peanuts, pecans, almonds, cashews, walnuts, macadamia nuts, pistachios, mixed nuts, and other nuts. To address these objectives, monthly household purchase and expenditure data for the aforementioned nut categories are derived from the Nielsen Homescan Panel for the years 2004 through 2015. As such, we provide more up-to-date and thorough analysis of the demand for peanuts and tree nuts, presently lacking in the extant literature.

The results of this article are as follows. Variety-seeking behavior was evident for peanuts but not for tree nuts. The level of pre-commitment was negative for the respective nut products. The demand for pecans and macadamia nuts was unitary elastic; the demand for peanuts, almonds, cashews, walnuts, pistachios, mixed nuts, and other nuts was inelastic. Predominantly, the various nut products were net substitutes. The income elasticities of nut products showed them to be necessities.

This article is organized as follows. First, we review previous studies related to tree nuts and peanuts. Second, we develop the LA/GAIDS model with variety-seeking behavior. Third, we discuss the data associated with this demand system analysis. Fourth, we present a discussion of the estimation of the LA/GAIDS model with attention given to endogeneity and autocorrelation. Fifth, we provide the empirical results centering attention on own-price, cross-price, and expenditure elasticities, variety-seeking behavior, and the level of pre-commitment of the various nut products. Finally, we make concluding remarks and discuss limitations of the research.

2.2. Literature Review

2.2.1. Demand for Nuts

Lee (1950) estimated season's average returns to almond growers for the period 1924-25 through 1948-49 by using linear regression and including the domestic volume of almonds, volume imported, and prices of competing products as explanatory factors. The estimated own-price elasticities of demand for almonds varied from -0.46 to -5.03. Wells, Miller, and Thompson (1986) estimated farm level demand for pecans using annual data from 1970-1982 based on a price-dependent demand function. The own-price flexibility of pecans at the farm level was estimated to be -0.97.

Lerner (1959) made the first attempt to investigate demand interrelationships of various tree nut products examining improved pecans, seedling pecans, general pecans, walnuts, filberts, and almonds using annual time-series data from 1922-1955 and seemingly unrelated regression. The estimated own-price elasticities were -2.73 for seedling pecans, -3.44 for improved pecans, -1.19 for general pecans, -1.80 for walnuts, -23.04 for filberts, and -0.86 for almonds. Pecans and walnuts were found to be gross complements. Pecans and filberts, pecans and almonds, and walnuts and almonds were found to be gross substitutes.

Dhaliwal (1972) examined demand interrelationships among eight tree nuts, including almonds, filbert, pecans, walnuts, pistachios, Brazil nuts, and cashews using annual time-series data from 1922-1955. The own-price elasticities were estimated to be -0.91 for pecans, -0.29 for walnuts, -1.93 for filberts, and -0.55 for almonds. Pecans and walnuts, pecans and Brazil nuts, and Brazil nuts and cashews were found to be gross substitutes; almonds and filberts as well as pecans and pistachios were found to be gross complements.

Russo, Green, and Howitt (2008) estimated price and income elasticities of almonds and walnuts by utilizing seemingly unrelated regression. Own-price elasticities for almonds were estimated to range from -0.35 to -0.48 and own-price elasticities for walnuts were estimated to range from -0.25 to -0.28. No substitution between almond and walnuts was evident.

Using causality structures identified through machine learning methods, Kim and Dharmasena (2017) examined prices received by growers of pecans from 2005-2016 to investigate market integration patterns in Texas, Oklahoma, Georgia, and Louisiana. They found that current pecan prices received by growers in Texas were directly caused by pecan prices received in Oklahoma, Georgia and Louisiana. Past period pecan prices in Georgia were found to influence current prices in other states.

To the best of our knowledge, no other studies have examined demand interrelationships of tree nuts and peanuts directly as well as the level of detail in tree nuts previously.

2.2.2. Variety-Seeking Behavior

The concept of variety-seeking behavior was introduced by Jevons (1871) who addressed the insatiability of consumer tastes. Subsequent works focused on theoretical developments in the attempt to explain the existence of variety-seeking behavior from the standpoint of economic theory (Anderson et al 1992; De Groot and Nahuis, 1998; Seegmuller 2008). Using data from the 1972-73 Consumer Expenditure Survey of the Bureau of Labor Statistics, Jackson (1984) found that the number of commodities purchased increased with total expenditure. Consequently, variety-seeking behavior of consumers was directly related to the level of their total expenditure.

Lancaster (1990) provided evidence that the level of product variety increased with the level of market competitiveness. Benassy (1996) revisited the link between market structure and the taste for variety and separated the taste for variety from the degree of market power.

Dubé (2004) provided a method to evaluate and test for variety seeking. He used a structural approach to build a multiple discreteness model to address multi-unit and multi-brand shopping behavior by using ACNielsen scanner data in the Denver area between January 1993 and March 1995 for carbonated soft drinks.

Drescher, Thiele, and Weiss (2008) tested the existence of variety-seeking behavior empirically utilizing a panel-data analysis of household expenditures on 182 different soft drinks in Germany. A hedonic method was used by regressing unit prices paid by households for these products on a variety index and product attributes. The variety index was constructed as the sum of the logarithms of quantity shares weighted by the product of quantity shares. The results supported the presence of variety-seeking behavior for the respective soft drink products, *ceteris paribus*.

2.2.3. Pre-commitment

Bollino (1987) introduced the generalized AIDS (GAIDS) model by explicitly including pre-commitment levels defined as those levels that are not sensitive to price and income. They were estimated in addition to parameters in the AIDS model following the method adopted by Gorman (1976) and popularized by Pollak and Wales (1981). This study revealed the superiority of GAIDS model over the AIDS model along with its advantages in assessing demographic effects using Italian household data from the period 1973 to 1983.

Tonsor and Marsh (2007) addressed the demand for meat and fish products in the United States and Japan using the GAIDS model. They found levels of pre-commitment for beef and pork in the United States as well as for beef and fish in Japan.

Hovhannisyan and Gould (2011) utilized the generalized quadratic almost ideal demand system (GUQAIDS) model to examine a system of eleven food commodities from the urban area of China using household-level expenditure survey data during 1995 and 2003. They estimated the model using full information maximum likelihood procedure and identified the structure change of food consumption in China that households tended to include more food from western countries as complements for traditional ones. Based on their previous work, Hovhannisyan and Gould (2014) developed the GQUAIDS model by incorporating a time transition function to allow budget share error terms to be autocorrelated overtime. Using provincial-level data through 2002 to 2010, Chinese food preferences were found to continuously evolve.

Rowland, Mjelde, and Dharmasena (2017) compared the ability of the AIDS and GAIDS models in addressing the demand for energy products in the United States, including oil, natural gas, and coal. The inclusion of pre-commitment levels led to a better understanding of demand for these respective energy categories.

Our research adds to the literature by providing a more up-to-date analysis of demand interrelationships for tree nuts and peanuts, while incorporating taste for variety and levels of pre-commitment.

2.3. Methodology

2.3.1. The Demand System Model

To the best of our knowledge, no previous studies in the extant literature have looked at demand interrelationships among peanuts and tree nuts by incorporating variety-seeking behavior and pre-commitment. We developed the Linearized and Generalized Almost

Ideal Demand System augmented with Entropy Index to analyze demand interrelationships for tree nuts and peanuts. The model builds on previous work by Deaton and Muellbauer (1980). Suppose an indirect utility function given by equation (1),

$$(1) \quad \ln V(\boldsymbol{p}, m) = [\left\{\frac{\ln m - \ln a(\boldsymbol{p})}{b(\boldsymbol{p})}\right\}^{-1} + \lambda(\boldsymbol{p})]^{-1},$$

where

$$(2) \quad \ln a(\boldsymbol{p}) = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln p_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} \ln p_i \ln p_j$$

and

$$(3) \quad b(\boldsymbol{p}) = \prod_{i=1}^{n} p_i^{\beta_i}.$$

p_i is the price for the ith product, m is total expenditure of the respective separable group of nut products in the demand system, n is the number of nut categories, $\ln a(\boldsymbol{p})$ is the translog price aggregator, and $b(\boldsymbol{p})$ is the Cobb-Douglas price aggregator. Then the budget share w_i for the ith product is given by equation (4),

$$(4) \quad w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln[\frac{m}{a(\boldsymbol{p})}].$$

2.3.2. Pre-commitment, Variety, and Seasonality

However, the conventional AIDS model neither considers pre-committed consumption nor preference for variety. Although nuts products have been recommended as part of daily eating habits of consumers, little is known about the demand for nuts along with pre-commitment levels and variety-seeking behavior. Following Bollino (1987) equation (5) can be derived,

$$(5)\quad w_i = \frac{c_i p_i}{m} + \left(1 - \frac{\sum_i c_i p_i}{m}\right)\left(\alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{m - \sum_i c_i p_i}{a(\boldsymbol{p})}\right]\right),$$

where c_i is the pre-committed quantities and $c_i p_i$ is pre-committed expenditure of product i defined as the component of demand not sensitive to adjustments of income and prices, also known as non-discretionary consumption. Also, we allow pre-committed quantities to vary across time by incorporating yearly dummy variables[1],

$$(6)\quad c_i = c_0 + \sum_{y=1}^{11} \tau_y c_y\ ,$$

where c_0 is the base year 2015, and c_y is the yearly dummies for 2004 through 2014. In addition, we include the Entropy Index into equation (5) in order to capture consumers taste for variety (Drescher, Thiele, and Weiss 2008). Unlike Dube (2004), we used an index to capture purchasing behavior of multiple products.

The index enters into the model as a part of the constant term, α_i, where

$$(7)\quad \alpha_i = \rho_{i0} + \rho_i E,$$

E is the Entropy Index[2], $E = -\sum_{i=1}^{n} s_i \ln s_i$, where $s_i = q_i/Q$, q_i is the quantity of product i and Q is the total quantity for all categories. The Entropy Index measures the level of variety-seeking behavior with larger values indicating higher levels of the taste

[1] We estimated several alternative models by incorporating trend and quadratic trend variables. According to model selection criteria, the model with only yearly dummies performed the best.
[2] We followed the method of Drescher, Thiele, and Weiss (2008) forcing the index to be positive by putting a negative sign in front since the quantity share is between 0 and 1.

for variety. Statistically significant and positive coefficients associated with this index (ρ_i) indicate the presence of variety seeking behavior.

After incorporating the Entropy Index as well as quarterly dummies D_s to account for seasonality, the budget share equations are given by equation (8),

$$(8)\quad w_i = \frac{c_i p_i}{m} + \sum_{s=1}^{3} d_{is} D_s$$

$$+ \left(1 - \frac{\sum_i c_i p_i}{m}\right)\left(\rho_{i0} - \rho_i \sum\nolimits_{i=1}^{n} s_i \ln s_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j\right.$$

$$\left. + \beta_i \ln\left[\frac{m - \sum_i c_i p_i}{a(\boldsymbol{p})}\right]\right),$$

Finally, theoretical restrictions are applied to conform to demand system conditions. They are adding up, homogeneity, and Slutsky symmetry as presented in equation (9).

$$(9)\quad \text{Adding} - \text{up: } \sum_{i=1}^{n} \alpha_i = 1\,, \qquad \sum_{i=1}^{n} \beta_i = 0\,, \quad \sum_{i=1}^{n} \gamma_{ij} = 0\,;$$

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

$$\text{Slutsky symmetry: } \gamma_{ij} = \gamma_{ji}.$$

2.3.3. Price and Expenditure Elasticities

In order to calculate elasticities from this model, we differentiate equation (8) with respect to the price and the expenditure terms. The expenditure elasticities,(ε_i) are given as follows.

$$(10)\quad \varepsilon_i = 1 - \frac{c_i p_i}{m w_i} + \frac{\sum_i c_i p_i}{m w_i}\left(\rho_{i0} - \rho_i \sum\nolimits_{i=1}^{N} s_i \ln s_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{m - \sum_i c_i p_i}{a(\boldsymbol{p})}\right]\right) + \left(\frac{1}{w_i} - \frac{\sum_i c_i p_i}{m w_i}\right)\left(\frac{m\beta_i}{m - \sum_i c_i p_i}\right)$$

The uncompensated price elasticities, (e_{ij}^{u}) are given by,

$$(11)\quad e_{ij}^{u} = -\frac{c_i p_i}{m w_i}\left(\rho_{i0} - \rho_i \sum\nolimits_{i=1}^{N} s_i \ln s_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{m - \sum_i c_i p_i}{a(\boldsymbol{p})}\right]\right) + \left(\frac{1}{w_i} - \frac{\sum_i c_i p_i}{m w_i}\right)\left(\gamma_{ij} + \beta_i\left(\frac{-c_j p_j}{m - \sum_i c_i p_i} - w_{jt-1}\right)\right) - \delta_{ij}$$

where δ_{ij} is the Kronecker delta where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. The compensated price elasticities, e_{ij}^{c}, are derived using Slutsky's equation, $e_{ij}^{c} = e_{ij}^{u} + \varepsilon_i w_j$. Note that the respective elasticity calculations depend not only on the estimated parameters of the LA/GAIDS model but also on prices, budget shares, the Entropy Index, pre-commitment levels, and total expenditure of the various nut products.

2.4. Data

The data used are monthly observations over the period 2004 through 2015 derived based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. We categorized peanuts and tree nuts based on product module codes and product descriptions provided by Nielsen. The nine categories are: (1)

peanuts, (2) pecans, (3) almonds, (4) cashews, (5) walnuts, (6) macadamia nuts, (7) pistachios, (8) mixed nuts, (9) other nuts. Other nuts consist of Brazil nuts, nuts toppings, pumpkin nuts, filberts, and sunflower seeds.

In the Nielsen Homescan Panel, purchases of nuts are reported for each household over time, including the amount paid in dollars, the coupon value in dollars, and the amount purchased in ounces. Initially, we generated monthly purchases and expenditures of peanuts and tree nuts made by each household over the period from 2004 through 2015. Next, the monetary values and net of coupon values paid by all households were summed to derive household expenditure for the respective nuts per month. As well, the amount purchased was summed up over all households for each month for the respective nuts. Then we divided the previously mentioned household expenditure and quantity data by the corresponding number of households who purchased the corresponding nut category to arrive at monthly purchases and expenditures of peanuts and tree nuts on a per household basis. Because not all households purchase nuts in any given time period, we further adjusted per household purchases and expenditures of peanuts and tree nuts by the annual market penetration for each respective nut category. The annual market penetration (see Table 2.1) was calculated by dividing the number of households who purchased each respective nut by the total number of households in Nielsen Homescan Panel for each year. The expenditure and quantity data subsequently are expressed in terms of dollars and ounces purchased per household per month.

Table 2.1 Annual Market Penetration

Year	Peanuts	Pecans	Almonds	Cashews	Walnuts
2004	53%	24%	27%	36%	28%
2005	54%	23%	28%	36%	29%
2006	53%	24%	30%	36%	30%
2007	50%	24%	31%	35%	28%
2008	48%	24%	34%	34%	27%
2009	48%	24%	35%	32%	30%
2010	48%	22%	35%	32%	28%
2011	48%	21%	38%	29%	27%
2012	46%	21%	38%	27%	26%
2013	46%	23%	36%	28%	25%
2014	46%	22%	33%	28%	23%
2015	43%	20%	26%	24%	22%

Year	Macadamia Nuts	Pistachios	Mixed Nuts	Other Nuts	Total Nuts
2004	3%	15%	34%	24%	86%
2005	3%	13%	33%	23%	86%
2006	4%	12%	33%	24%	86%
Table 2.1 Continued					
2007	3%	15%	32%	25%	85%
2008	3%	16%	31%	24%	85%
2009	3%	14%	31%	23%	85%
2010	3%	16%	30%	22%	85%
2011	2%	23%	30%	23%	85%
2012	2%	26%	29%	23%	84%
2013	2%	23%	28%	23%	84%
2014	2%	20%	27%	23%	83%
2015	2%	17%	23%	20%	80%

Note: Market penetration is defined as the number of households purchasing the nut category divided by the total number of households in any given year.

Source: Nielsen Homescan Panel 2004-2015, and calculations by the author.

Further, we calculated the monthly unit values (used as a proxy for price) for each nut category by dividing monthly expenditures by monthly quantities purchased.

Subsequently, we summed all expenditures of peanuts and tree nuts to derive total expenditure per month on a household basis. By dividing expenditure of each nut type by total expenditure, we obtained the respective budget shares for peanuts and tree nuts per month. Bottom line, we developed monthly household purchases (ounces), prices ($/ounce) and expenditures (dollars) over the period of 2004 through 2015, a total of 144 observations for nuts products in the United States.

2.4.1. Descriptive Statistics

Descriptive statistics for prices ($/ounce), budget shares, quantities (ounces/month) corresponding to each of the nut categories, total expenditure ($/month), income ($), and Entropy Index are presented in Table 2.2. Macadamia nuts were the most expensive nut product at 71 cents per ounce, followed by pecans, almonds, walnuts, pistachios, cashews and mixed nuts. Peanuts were the least expensive product purchased at 13 cents per ounce.

Monthly purchases were highest for peanuts at 17.28 ounces per household on average, followed by that of mixed nuts, cashews, and almonds at 7.95, 7.06, and 6.51 ounces on average. Monthly purchases of walnuts, pistachios, other nuts, pecans, and macadamia nuts per household were 5.23, 4.44, 3.82, 3.45, and 0.29 ounces on average. Over the 144 months, the averages of the budget shares were as follows: peanuts 14.92%; pecans 10.66%; almonds 14.75%; cashews 14.74%; walnuts 11.73%; macadamia nuts 1.31%; pistachios 9.87%; mixed nuts 16.28%; other nuts 5.75%. Budget shares were highest for mixed nuts, peanuts, almonds, and cashews, followed by walnuts, pecans, and pistachios. Budget shares were lowest for other nuts and macadamia nuts. Per household total expenditure for peanuts and tree nuts purchased at home was on average $15.36 per month over the period 2004 through 2015. Monthly income per capita over the period 2004 through 2015 was $36,635.40 on average. The Entropy Index, on average, was 1.95.

Table 2.2 Descriptive Statistics of prices, Budget Shares, Quantities, Total Expenditure, Income, and Entropy Index

		Mean	Std. Dev.	Min	Max
Prices ($/ounce)	Peanuts	0.13	0.02	0.11	0.17
	Pecans	0.48	0.07	0.36	0.62
	Almonds	0.35	0.05	0.26	0.50
	Cashews	0.33	0.07	0.20	0.43
	Walnuts	0.35	0.08	0.22	0.50
	Macadamia Nuts	0.71	0.14	0.51	0.99
	Pistachios	0.35	0.11	0.20	0.58
	Mixed Nuts	0.32	0.06	0.22	0.43
	Other Nuts	0.23	0.04	0.17	0.31
Budget Shares	Peanuts	0.15	0.01	0.13	0.18
	Pecans	0.11	0.01	0.10	0.13
	Almonds	0.15	0.01	0.11	0.17
	Cashews	0.15	0.01	0.12	0.17
	Walnuts	0.12	0.01	0.10	0.14
	Macadamia Nuts	0.01	0.00	0.01	0.02
	Pistachios	0.10	0.02	0.06	0.14
	Mixed Nuts	0.16	0.01	0.14	0.19
	Other Nuts	0.06	0.00	0.05	0.07
Quantities Purchased (ounces/month)	Peanuts	17.28	1.13	13.98	20.27
	Pecans	3.45	0.37	2.78	4.56
	Almonds	6.51	1.25	3.79	9.19
	Cashews	7.06	1.41	4.90	10.35
	Walnuts	5.23	0.61	3.96	6.63
	Macadamia Nuts	0.29	0.09	0.11	0.46
	Pistachios	4.44	0.85	2.70	6.36

Table 2.2 Continued

		Mean	Std. Dev.	Min	Max
	Mixed Nuts	7.98	0.87	5.56	10.18
	Other Nuts	3.82	0.34	2.59	4.94
Expenditures ($/month)	Total Expenditure	15.36	1.82	12.17	18.88
Variety Seeking	Entropy Index	1.95	0.02	1.90	1.99
Income ($/month)	Per Capita Income[a]	36,635.40	3,397.56	30,216.00	42,847.00

[a] Source: Federal Reserve Data Base St. Louis.

Source: Nielsen Homescan Panel 2004-2015, and calculations by the author.

2.5. Estimation

Several issues need to be addressed in estimation. First, because we employ time-series data in the analysis, the presence of serial correlation needs to be examined through the examination of autocorrelation and partial autocorrelation functions of the disturbance terms,

$$(12)\quad w_{it} = \sum_{k} \rho_k \, w_{it-k} + f(x_{it}, \beta) - \sum_{k} \rho_k f(x_{it-k}, \beta) + \sum_{s=1}^{3} d_{is} \, D_s + \epsilon_{it}$$

where k is the number of lag terms and $f(x_{it}, \beta)$ is the function form from equation (8) (Berndt and Savin 1975; Dharmasena and Capps 2012; Hovhannisyan and Gould 2014). The optimal lag of ρ_k is based on the autocorrelation and partial autocorrelation function of the error terms. Examination of these functions revealed an *AR(1)* process of the disturbance terms. Owing to adding-up, estimation of a single ρ across the system was necessary.

The second issue is that total expenditure in the GAIDS model suffers from endogeneity. To mitigate this problem, the endogeneity issue is addressed through predictions of total expenditure, m_t obtained through the use of linear regression (see equation (13)), wherein the natural log of expenditure m_t is regressed on the natural log of the income, $\ln \text{Income}_t$. This instrument-variable method is similar to the works of Attfield (1985), Capps et al. (1994), and Dharmasena and Capps (2012).

The specification, estimation, and choice of optimal lags were executed in the statistical software EViews 9 (Global 2016). All of coefficients estimated were significant at the 1% level. The goodness-of-fit (R^2) is 0.9421, the adjusted R^2 is 0.9404, and Durbin-Watson statistics is 2.0751.

$$(13) \quad (1 - 0.4753L)(1 - 0.4193L^2) \ln m_t = -4.1449 + 0.6527 \ln \text{Income}_t + (1 - 0.8432L^{12})\phi_t$$

where L is the lag operator, and Income_t is income at time period t, ϕ_t is the disturbance term.

Predicted values of m_t based on equation (13) were subsequently plugged back in the demand system model. We derive the income elasticity from equation (14) as follows,

$$(14) \quad IE_i = \frac{\%\Delta \text{ Total Expenditure}}{\%\Delta \text{ Income}} \times \frac{\%\Delta \text{ Quantity Demanded}_i}{\%\Delta \text{ Total Expenditure}} = 0.6527 \times \varepsilon_i,$$

where IE_i is the income elasticity for nut i, ε_i is expenditure elasticity derived from equation (10), and 0.6527 is the estimated coefficient from equation (13), and %Δ depicts the percentage change of each variable.

Third, the Entropy Index also suffers from endogeneity. To circumvent this issue, we lag the Entropy Index one period.

Finally, because the GAIDS model is highly non-linear[3], in order to reduce the complexity and hence difficulty of estimation, we linearized the two price indexes $\ln a(\boldsymbol{p})$ and $b(\boldsymbol{p})$ as follows,

$$(15)\quad \ln a(\boldsymbol{p}) = \sum_i w_{it-1} \ln p_{it}$$

and

$$(16)\quad b(\boldsymbol{p}) = \sum_i (w_{it-1} - w_{it-2}) (\ln p_{it-1} - \ln p_{it-2})$$

We used Stone's index to replace $\ln a(\boldsymbol{p})$. In order to avoid any contemporaneous correlation among the budget shares in Stone's price index and the budget shares as the dependent variable in the GAIDS[4] model, we modified the Stone index by lagging the budget shares by one period in equation (15). The Cobb-Douglas price aggregator in the GAIDS model was linearized via equation (16), following the method suggested by Diewert (1987) and Matsuda (2006) and as applied by Dharmasena and Capps (2012). Hence, the Linearized and Generalized Almost Ideal Demand System[5] (LA/GAIDS) with Entropy Index is defined by incorporating equations (8), (9), (12), (15), and (16).

---

[3] The generalized version of the QAIDS presented in this paper with Entropy index and pre-commitments adds nonlinearity to the system as well as the two price indexes used in the original AIDS model. We used the linearized version of the Generalized Almost Ideal Demand System (LA/GAIDS) following suggestion by Matsuda (2006) in order to overcome estimation complexities and generate sensible results that are theoretically acceptable and empirically valid.
[4] We also considered nonlinear Engel curves that is usually tackled with the use of quadratic AIDS (QUAIDS) model. However, after estimating the QUAIDS, all of the coefficients for quadratic terms (lambdas) were not significant neither individually nor jointly. The other coefficients were very similar as AIDS model used in this study.
[5] Another potential source of endogeneity come from prices variables, determined by the forces of supply and demand simultaneously. Using micro household level data and unit values reduces the significance of

2.6. Empirical Results

The LA/GAIDS model was estimated using SAS© software, Version 9.4 (SAS Institute, 2012) through the use of an iterative seemingly unrelated regression procedure (ITSUR) available in the PROC MODEL procedure. In order to accommodate adding-up restrictions imposed in equation (9), we dropped one equation from the estimation process. Arbitrarily, the omitted equation was associated with mixed nuts. We recovered all coefficients for mixed nuts using equation (9). Goodness-of-fit statistics R^2 and adjusted R^2 as well as Durbin-Watson statistics are provided in Table 2.3. The R^2 metrics for the LA/GAIDS model ranged from 0.7152 to 0.9779. The model was corrected for serial correlation using an *AR(1)* process in the disturbance terms as previously stated; as exhibited in Table 2.3, *r1* is the estimated parameter associated with this *AR(1)* process. The Durbin-Watson statistics corresponding to this demand system ranged from 1.6655 to 2.1236, thereby providing evidence of the absence of autocorrelation in the error terms.

2.6.1. Parameters and Seasonality

All of the estimated parameters and associated *p*-values are shown in Table 2.3. Bold numbers indicate that estimated coefficients are considered to be significantly different from zero provided that the corresponding *p*-values are less than 0.05, the level of significance chosen for our analysis. We recovered the parameters for the omitted category, namely mixed nuts, using adding-up restrictions as shown in equation (9).

---

this issue since household purchase decisions typically do not influence the market and the equilibrium price (Zhen et al. 2013).

Twenty-eight out of the forty-five estimated gamma parameters, γ_{ij}, were statistically different from zero, while four out of the nine estimated alpha parameters, and five out of the nine estimated beta parameters were statistically different from zero. The estimated autocorrelation coefficient (*r1*) was statistically different from zero.

Table 2.3 Nonlinear ISTUR Parameter Estimates and Associated p-Values

Parameter	Estimate	p-value		Estimate	p-value
Gammas (γ)			**Gammas (γ)**		
gpeanutpeanut	**0.0806**	0.00	gmacamaca	-0.0026	0.56
gpeanutpecan	-0.0009	0.84	gmacapistachio	0.0006	0.58
gpeanutalmond	**-0.0383**	0.00	gmacaother	**0.0024**	0.01
gpeanutcashew	-0.0002	0.96	gmacamix	0.0021	0.13
gpeanutwalnut	**-0.0089**	0.02	gpistachiopistachio	**0.0735**	0.00
gpeanutmaca	0.0010	0.53	gpistachioother	**-0.0069**	0.01
gpeanutpistachio	**-0.0164**	0.00	gpistachiomix	**-0.0111**	0.00
gpeanutother	**-0.0112**	0.00	gotherother	**0.0677**	0.00
gpeanutmix	-0.0056	0.25	gothermix	**-0.0149**	0.00
gpecanpecan	0.0291	0.08	gmixmix	**0.0700**	0.00
gpecanalmond	**-0.0209**	0.00			
gpecancashew	0.0031	0.49	**AR(1)**		
gpecanwalnut	0.0006	0.87	r1	**0.1835**	0.00
gpecanmaca	**0.0024**	0.05			
gpecanpistachio	-0.0050	0.14	Alphas (α)		
gpecanother	-0.0039	0.06	apeanut	0.1334	0.32
gpecanmix	-0.0046	0.26	apecan	**-0.2195**	0.04
galmondalmond	**0.1908**	0.00	aalmond	**0.9165**	0.00
galmondcashew	**-0.0340**	0.00	acashew	-0.1106	0.23
galmondwalnut	**-0.0252**	0.00	awalnut	**-0.1776**	0.02
galmondmaca	**-0.0051**	0.00	amaca	-0.0365	0.14
galmondpistachio	**-0.0191**	0.00	apistachio	0.1378	0.23
galmondother	**-0.0193**	0.00	aother	**0.2268**	0.00
galmondmix	**-0.0290**	0.00	amix	0.1297	0.16
gcashewcashew	**0.0630**	0.00			

Table 2.3 Continued

Parameter	Estimate	p-value		Estimate	p-value
gcashewwalnut	-0.0049	0.16	Betas (β)		
gcashewmaca	0.0015	0.25	bpeanut	0.0045	0.81
gcashewpistachio	**-0.0115**	0.00	bpecan	**0.0540**	0.00
gcashewother	**-0.0080**	0.00	balmond	**-0.1209**	0.00
gcashewmix	**-0.0088**	0.03	bcashew	**0.0383**	0.01
gwalnutwalnut	**0.0487**	0.00	bwalnut	**0.0521**	0.00
gwalnutmaca	**-0.0022**	0.04	bmaca	0.0074	0.07
gwalnutpistachio	-0.0042	0.18	bpistachio	-0.0061	0.72
gwalnutother	**-0.0059**	0.00	bother	**-0.0269**	0.02
gwalnutmix	0.0020	0.63	bmix	-0.0023	0.87

Goodness-of-Fit

Equation	R^2	Adjusted R^2	Durbin Watson
Peanuts	0.9383	0.9259	1.9899
Pecans	0.8685	0.8420	2.1236
Almonds	0.9174	0.9008	1.7286
Cashews	0.9545	0.9453	1.6655
Walnuts	0.9193	0.9030	2.1208
Macadamia Nuts	0.9203	0.9043	2.0650
Pistachios	0.9779	0.9735	1.7597
Other Nuts	0.7152	0.6579	1.9703
Mixed Nuts	0.9080	0.8872	2.0293

[a]After we recovered the parameters for mixed nuts using adding-up restrictions (equation (9)), we calculated the predicted values of budget shares of mixed nuts, and then derived the goodness-of-fit, R^2, adjusted R^2, and Durbin Watson statistic for mixed nuts.

Note: Bold numbers indicate significance at the 5% level. Each equation has 144 observations.

Source: Nielsen Homescan Panel 2004-2015, and calculations by the author.

Quarterly dummies were included in each equation to capture quarterly seasonality. As shown in Table 2.4, seasonality was evident for pecans, almonds, cashews, macadamia nuts, pistachios, and mixed nuts. No seasonal pattern was found for peanuts, walnuts, and other nuts.

Table 2.4 Parameters Associated with Seasonality and Joint Tests of Seasonality

	Estimate	p-value		Estimate	p-value		Estimate	p-value
dpeanut1	0.0000	1.00	dcashew1	**-0.0024**	0.01	dpistachio1	0.0020	0.08
dpeanut2	-0.0010	0.31	dcashew2	-0.0006	0.56	**dpistachio2**	0.0038	0.00
dpeanut3	-0.0016	0.08	dcashew3	0.0008	0.41	**dpistachio3**	0.0025	0.04
dpecan1	**-0.0056**	0.00	dwalnut1	0.0002	0.75	**dother1**	-0.0013	0.05
dpecan2	**-0.0067**	0.00	dwalnut2	-0.0004	0.67	**dother2**	-0.0021	0.01
dpecan3	**-0.0063**	0.00	dwalnut3	-0.0005	0.54	dother3	-0.0013	0.07
dalmond1	**0.0084**	0.00	dmaca1	0.0000	0.92	dmix1	-0.0014	0.13
dalmond2	**0.0052**	0.00	dmaca2	0.0005	0.23	dmix2	0.0012	0.21
dalmond3	**0.0067**	0.00	dmaca3	-0.0003	0.38	dmix3	-0.0002	0.87

Joint Tests of Seasonal Dummies

Null Hypotheses	Chi-squared Statistic	p-value
dpeanut1=dpeanut2=dpeanut3=0	5.67	0.13
dpecan1=dpecan2=dpecan3=0	**63.03**	0.00
dalmond1=dalmond2=dalmond3=0	**63.51**	0.00
dcashew1=dcashew2=dcashew3=0	**17.57**	0.00
dwalnut1=dwalnut2=dwalnut3=0	1.38	0.71
dmaca1=dmaca2=dmaca3=0	**8.04**	0.05
dpistachio1=dpistachio2=dpistachio3=0	**9.31**	0.03
dother1=dother2=dother3=0	7.12	0.07
dmix1=dmix2=dmix3=0	**12.39**	0.01

Note: Bold numbers indicate significance at the 5% level.

Source: Nielsen Homescan Panel 2004-2015, and calculations by the author.

2.6.2. Pre-commitment

The level of pre-commitment is defined as the component of demand not sensitive to the influences of income and price, also known as non-discretionary consumption. The consumption of nuts has been on the rise and most nuts are consumed as snacks. Taking pre-commitment into consideration delineates demand in terms of autonomous consumption or subsistence level of consumption. Additionally, we allow the level of pre-commitment (quantities) to vary by year for each nut category. The estimated parameters associated with pre-commitment levels (see equation (6)) over the period 2004 through 2015 shown in Table 2.5; joint tests of pre-commitment are exhibited in Table 2.6. Our results show that pre-commitment was evident for pecans, almonds, walnuts, macadamia nuts, pistachios, and other nuts based on the joint tests; as well as pre-commitment levels for nuts varied over time. However, pre-commitment levels were negative over the period 2004 through 2015. The empirical finding of significant negative pre-commitment levels is counterintuitive and reflects household preferences for a shift out of subsistence expenditures into supernumerary expenditures (Tonsor and Marsh 2007).

Table 2.5 Estimates of Parameters Associated with Pre-committed Quantities

	Estimate	p-value		Estimate	p-value		Estimate	p-value
cpeanut0	**-28.0805**	0.00	ccashew0	**-7.7260**	0.02	cpistachio0	**-9.6782**	0.00
cpeanut1	-8.6456	0.18	ccashew1	-4.8084	0.13	cpistachio1	**-4.7804**	0.03
cpeanut2	-3.1914	0.39	ccashew2	-1.2291	0.43	cpistachio2	**-3.3662**	0.00
cpeanut3	**-13.5741**	0.02	ccashew3	**-5.7676**	0.03	cpistachio3	**-6.2396**	0.00
cpeanut4	2.0264	0.62	ccashew4	1.2852	0.44	cpistachio4	-0.5767	0.62
cpeanut5	6.4955	0.11	ccashew5	3.0824	0.07	cpistachio5	1.4356	0.23
cpeanut6	-2.4918	0.51	ccashew6	-1.0369	0.51	cpistachio6	-1.4378	0.20
cpeanut7	5.2873	0.22	ccashew7	1.9975	0.26	cpistachio7	1.3306	0.27
cpeanut8	8.1438	0.10	ccashew8	2.9855	0.13	cpistachio8	**3.7105**	0.01
cpeanut9	**13.5474**	0.02	ccashew9	**4.7105**	0.03	cpistachio9	**6.0162**	0.00
cpeanut10	5.3089	0.16	ccashew10	1.9810	0.19	cpistachio10	**2.9125**	0.01
cpeanut11	**15.6737**	0.00	ccashew11	**6.1019**	0.00	cpistachio11	**4.6541**	0.00
cpecan0	-2.8835	0.20	cwalnut0	-4.1090	0.07	cother0	**-11.8066**	0.00
cpecan1	-2.8534	0.09	cwalnut1	**-4.9328**	0.05	cother1	-2.1810	0.15
cpecan2	-0.7739	0.40	cwalnut2	-2.0028	0.14	cother2	-0.9818	0.26
cpecan3	**-3.3042**	0.03	cwalnut3	**-5.6867**	0.01	cother3	-2.5634	0.06
cpecan4	0.7044	0.47	cwalnut4	0.5565	0.68	cother4	1.0223	0.26
cpecan5	1.9205	0.06	cwalnut5	**2.4938**	0.05	cother5	**2.1415**	0.03
cpecan6	-0.5919	0.55	cwalnut6	0.0267	0.98	cother6	-0.1623	0.85
cpecan7	1.2450	0.24	cwalnut7	2.1626	0.13	cother7	1.3675	0.18
cpecan8	2.0029	0.07	cwalnut8	**2.9685**	0.05	cother8	2.1578	0.07
cpecan9	**3.0584**	0.01	cwalnut9	**4.5443**	0.01	cother9	**3.8405**	0.01
cpecan10	1.6421	0.08	cwalnut10	1.8970	0.11	cother10	1.6964	0.06
cpecan11	**3.7632**	0.00	cwalnut11	**4.2666**	0.01	cother11	**4.0891**	0.00
calmond0	**-28.3356**	0.00	cmaca0	0.0361	0.90	cmix0	**-11.7226**	0.00
calmond1	-4.5407	0.06	cmaca1	0.0603	0.59	cmix1	-3.3358	0.22
calmond2	-1.9145	0.21	cmaca2	0.1264	0.07	cmix2	-0.8306	0.59
calmond3	**-5.3100**	0.02	cmaca3	0.0160	0.89	cmix3	**-4.7853**	0.04
calmond4	1.5205	0.38	cmaca4	**0.2015**	0.00	cmix4	1.4565	0.36
calmond5	**4.3714**	0.02	cmaca5	**0.2375**	0.00	cmix5	3.0860	0.07
calmond6	0.4608	0.76	cmaca6	0.0288	0.68	cmix6	-0.3523	0.82
calmond7	3.3734	0.07	cmaca7	0.1062	0.09	cmix7	2.6327	0.13
calmond8	**5.7759**	0.01	cmaca8	0.0769	0.28	cmix8	**3.9293**	0.05
calmond9	**9.0559**	0.00	cmaca9	0.1417	0.13	cmix9	**5.9896**	0.01
calmond10	**4.3805**	0.01	cmaca10	**0.1085**	0.04	cmix10	2.7471	0.07
calmond11	**8.2194**	0.00	cmaca11	**0.2097**	0.02	cmix11	**6.6288**	0.00

Note: Bold numbers indicate significance at the 5% level. The base year is 2015, and 1-11 represents the yearly dummies for 2004 to 2014.

Source: Nielsen Homescan Panel 2004-2015, and calculations by the author.

Table 2.6 Joint Tests of Pre-commitment

Null Hypotheses	**Chi-squared Statistic**	**p-value**
cpeanut0=cpeanut1=cpeanut2=...=cpeanut10=cpeanut11=0	14.73	0.26
cpecan0=cpecan1=cpecan2=...=cpecan10=cpecan11=0	**26.7**	0.01
calmond0=calmond1=calmond2=...=calmond10=calmond11=0	**41.12**	0.00
ccashew0=ccashew1=ccashew2=...=ccashew10=ccashew11=0	13.43	0.34
cwalnut0=cwalnut1=cwalnut2=...=cwalnut10=cwalnut11=0	**23.22**	0.03
cmaca0=cmaca1=cmaca2=...=cmaca10=cmaca11=0	**28.93**	0.00
cpistachio0=cpistachio1=cpistachio2=...=cpistachio10=cpistachio11=0	**34.77**	0.00
cother0=cother1=cother2=...=cother10=cother11=0	**28.4**	0.00
cmix0=cmix1=cmix2=...=cmix10=cmix11=0	14.86	0.25

Note: Bold numbers indicate significance at the 5% level. The base year is 2015, and 1-11 represents the yearly dummies for 2004 to 2014.

Source: Nielsen Homescan Panel 2004-2015, and calculations by the author.

The level of pre-commitment of peanuts and tree nuts over the period 2004 to 2015 are exhibited pictorially in Figure 2.3. Although most of estimated pre-committed quantities were negative, the pre-commitment levels of peanuts and tree nuts generally exhibited an upward trend (that is, less negative over time) over the period from 2004 through 2015.

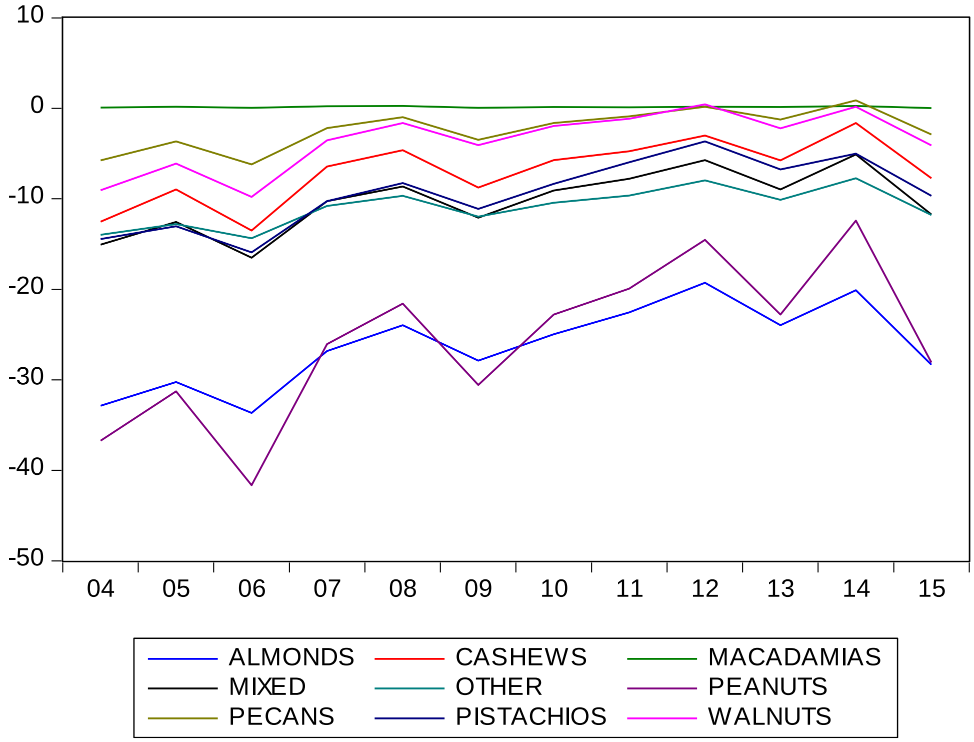


Figure 2-3 Annual Pre-Committed Quantities of Peanuts and Tree Nuts in Ounces from 2004-2015

Source: Nielsen Homescan Panel 2004-2015, calculations by the author.

2.6.3. Variety-Seeking Behavior

Statistically significant and positive coefficients indicate the preference for variety-seeking behavior. As shown in Table 2.7, variety-seeking behavior was evident for peanuts but not for almonds. Estimated coefficients associated with the Entropy Index are positive although not statistically significant for pecans, cashews, macadamia nuts, other nuts, and mixed nuts.

Table 2.7 Estimated Parameters Associated with Variety-Seeking Behavior

	Estimate	p-value
Peanuts	**0.0283**	0.03
Pecans	0.0037	0.76
Almonds	**-0.0436**	0.01
Cashews	0.0143	0.28
Walnuts	-0.0089	0.44
Macadamia Nuts	0.0029	0.55
Pistachios	-0.0067	0.71
Other Nuts	0.0020	0.84
Mixed Nuts	0.0080	0.53

Note: Bold numbers indicate significance at the 5% level.

Source: Nielsen Homescan Panel 2004-2015, and calculations by the author.

2.6.4. Uncompensated and Compensated Price Elasticities

We calculated uncompensated and compensated own-price and cross-price elasticities for each data point (144 months from January 2004 through December 2015) for each nut category. The respective elasticities are dependent not only on the estimated parameters but also on prices, total expenditure, Entropy Index, pre-commitment levels, and budget shares.

Uncompensated own-price and cross-price elasticities are exhibited in Table 2.8. These respective elasticities correspond to the means of the respective 144 monthly observations over the sample period from January 2004 to December 2015. All of the nine own-price elasticities were statistically significant, ranging from -0.4503 to -1.0560. Unitary elastic demands were noted for pecans and macadamia nuts; inelastic demands were evident for peanuts, almonds, cashews, walnuts, pistachios, other nuts, and mixed nuts. Compensated own-price and cross-price elasticities are exhibited for each nut category in Table 2.9.

As expected, most nut types were revealed to be net substitutes for each other. Pecans, almonds, cashews, walnuts, macadamia nuts, other nuts and mixed nuts were net substitutes for peanuts. Peanuts, almonds, cashews, walnuts, macadamia nut, pistachios, other nuts, and mixed nuts were net substitutes for pecans. Peanuts, pecans, cashews, walnuts, pistachios, other nuts and mixed nuts were net substitutes for almonds. Peanuts, pecans, almonds, walnuts, macadamia nuts, pistachios, other nuts, and mixed nuts were net substitutes for cashews. Pecans, almonds, cashews, pistachios, other nuts, and mixed nuts were net substitutes for walnuts. Peanuts, pecans, cashews, pistachios, other nuts, and mixed nuts were net substitutes for macadamia nuts. Pecans, almonds, cashews, walnuts, macadamia nuts, other nuts, and mixed nuts were net substitutes for pistachios. Peanuts, pecans, cashews walnuts, macadamia nuts, and pistachios were net substitutes for other nuts. Peanuts, pecans, almonds, cashews, walnuts, macadamia nuts, and pistachios were net substitutes for mixed nuts.

Table 2.8 Uncompensated Own-Price and Cross-Price Elasticities

	Peanuts	Pecans	Almonds	Cashews	Walnuts	Macadamia Nuts	Pistachios	Other Nuts	Mixed Nuts
Peanuts	**-0.8180**	**0.0389**	**-0.1186**	**0.1172**	**-0.1017**	**0.0097**	**-0.1134**	**-0.0497**	**0.0824**
	(0.0123)	(0.0044)	(0.0042)	(0.0043)	(0.0025)	(0.0005)	(0.0012)	(0.0019)	(0.0029)
Pecans	**0.0251**	**-0.9620**	**-0.0270**	**0.0382**	**-0.0589**	**0.0341**	**-0.0278**	**0.0102**	**-0.0992**
	(0.0042)	(0.0334)	(0.0095)	(0.0042)	(0.0048)	(0.0009)	(0.0055)	(0.0015)	(0.0019)
Almonds	**-0.1267**	**-0.0676**	**-0.7035**	**-0.1381**	**-0.1228**	**-0.0706**	**0.0785**	**-0.0485**	**0.0610**
	(0.0028)	(0.0042)	(0.0192)	(0.0024)	(0.0020)	(0.0011)	(0.0055)	(0.0031)	(0.0047)
Cashews	**0.1077**	**0.0479**	**-0.1176**	**-0.7685**	**-0.1003**	**0.0109**	**-0.0800**	**-0.0228**	**-0.0683**
	(0.0038)	(0.0046)	(0.0043)	(0.0145)	(0.0031)	(0.0005)	(0.0029)	(0.0012)	(0.0015)
Walnuts	**-0.1509**	**-0.0379**	**-0.1035**	**-0.1407**	**-0.5713**	**-0.0743**	0.0028	**-0.0331**	**0.0689**
	(0.0022)	(0.0038)	(0.0055)	(0.0028)	(0.0204)	(0.0016)	(0.0049)	(0.0013)	(0.0037)
Macadamia Nuts	**0.1382**	**0.3967**	**-0.8707**	**0.1930**	**-0.6190**	**-1.0560**	**0.1318**	**0.5609**	**0.3532**
	(0.0046)	(0.0120)	(0.0252)	(0.0068)	(0.0177)	(0.0343)	(0.0073)	(0.0158)	(0.0111)
Pistachios	**-0.2116**	**-0.0446**	**0.1552**	**-0.1543**	**-0.0154**	**0.0142**	**-0.7997**	**-0.0108**	**-0.0644**
	(0.0076)	(0.0030)	(0.0085)	(0.0055)	(0.0034)	(0.0009)	(0.0218)	(0.0037)	-(0.1683)
Other Nuts	**-0.1239**	**0.0149**	**-0.1265**	**-0.0522**	**-0.0656**	**0.1236**	**0.0916**	**-0.4503**	**-0.2903**
	(0.0038)	(0.0041)	(0.0097)	(0.0020)	(0.0020)	(0.0031)	(0.0093)	(0.0263)	(0.0062)
Mixed Nuts	**0.0949**	**-0.0158**	**0.0373**	**-0.0222**	**0.0948**	**0.0290**	**-0.0130**	**-0.1052**	**-0.9416**
	(0.0030)	(0.0031)	(0.0035)	(0.0020)	(0.0044)	(0.0007)	(0.0022)	(0.0024)	(0.0098)

Note: The estimated elasticities correspond to the means of 144 monthly values for the period 2004 to 2015. Bold numbers indicate significance at the 5% level; the numbers below the estimated elasticities correspond to the standard errors over the time period 2004 to 2015.

Source: Nielsen Homescan Panel 2004-2015, and calculations by the authors.

Table 2.9 Compensated Own-Price and Cross-Price Elasticities

	Peanuts	**Pecans**	**Almonds**	**Cashews**	**Walnuts**	**Macadamia Nuts**	**Pistachios**	**Other Nuts**	**Mixed Nuts**
Peanuts	**-0.6764**	**0.1403**	**0.0228**	**0.2569**	**0.0104**	**0.0257**	**-0.0182**	**0.0051**	**0.2370**
	(0.0118)	(0.0045)	(0.0039)	(0.0046)	(0.0021)	(0.0007)	(0.0024)	(0.0022)	(0.0030)
Pecans	**0.1844**	**-0.8481**	**0.1305**	**0.1957**	**0.0664**	**0.0257**	**0.0769**	**0.0714**	**0.0746**
	(0.0058)	(0.0325)	(0.0093)	(0.0056)	(0.0049)	(0.0007)	(0.0042)	(0.0018)	(0.0033)
Almonds	**0.0434**	**0.0540**	**-0.5359**	**0.0304**	**0.0104**	**-0.0556**	**0.1894**	**0.0169**	**0.2468**
	(0.0045)	(0.0052)	(0.0182)	(0.0043)	(0.0025)	(0.0010)	(0.0069)	(0.0033)	(0.0068)
Cashews	**0.2547**	**0.1532**	**0.0297**	**-0.6236**	**0.0164**	**0.0236**	**0.0197**	**0.0342**	**0.0923**
	(0.0042)	(0.0046)	(0.0031)	(0.0140)	(0.0025)	(0.0005)	(0.0021)	(0.0015)	(0.0020)
Walnuts	0.0048	**0.0734**	**0.0495**	**0.0133**	**-0.4497**	**-0.0605**	**0.1042**	**0.0267**	**0.2388**
	(0.0027)	(0.0050)	(0.0056)	(0.0026)	(0.0198)	(0.0013)	(0.0040)	(0.0012)	(0.0054)
Macadamia Nuts	**0.2523**	**0.4786**	**-0.7558**	**0.3055**	**-0.5274**	**-1.0464**	**0.2107**	**0.6052**	**0.4780**
	(0.0062)	(0.0132)	(0.0232)	(0.0086)	(0.0159)	(0.0343)	(0.0075)	(0.0165)	(0.0132)
Pistachios	**-0.0420**	**0.0767**	**0.3214**	**0.0140**	**0.1171**	**0.0294**	**-0.6915**	**0.0542**	**0.1206**
	(0.0051)	(0.0045)	(0.0095)	(0.0034)	(0.0045)	(0.0013)	(0.0212)	(0.0034)	(0.0035)
Other Nuts	**0.0211**	**0.1187**	0.0194	**0.0904**	**0.0498**	**0.1361**	**0.1913**	**-0.3939**	**-0.1319**
	(0.0049)	(0.0035)	(0.0117)	(0.0021)	(0.0022)	(0.0031)	(0.0106)	(0.0255)	(0.0074)
Mixed Nuts	**0.2198**	**0.0736**	0.1625	**0.1009**	**0.1939**	**0.0399**	**0.0716**	**-0.0569**	**-0.8052**
	(0.0031)	(0.0030)	(0.0031)	(0.0021)	(0.0039)	(0.0008)	(0.0029)	(0.0027)	(0.0099)

Note: The estimated elasticities correspond to the means of 144 monthly values for the period 2004 to 2015. Bold numbers indicate significance at the 5% level; the numbers below the estimated elasticities correspond to the standard errors over the time period 2004 to 2015.

Source: Nielsen Homescan Panel 2004-2015, and calculations by the authors.

Meanwhile several net complementarity relationships were revealed, such as peanuts and pistachios; almonds and macadamia nuts; walnuts and macadamia nuts; and mixed nuts and other nuts. We summarize these findings concerning the substitutability and complementarity of peanuts and tree nuts in Table 2.10.

Table 2.10 Evidence of Net Substitutability and Net Complementarity among Nut Categories

	Substitutes	Complements
Peanuts	Pecans, Almonds, Cashews, Walnuts, Macadamia Nuts, Other Nuts and Mixed Nuts	Pistachios
Pecans	Peanuts, Almonds, Cashews, Walnuts, Macadamia Nuts, Pistachios, Other Nuts, and Mixed Nuts	
Almonds	Peanuts, Pecans, Cashews, Walnuts, Pistachios, Other Nuts and Mixed Nuts	Macadamia Nuts
Cashews	Peanuts, Pecans, Almonds, Walnuts, Macadamia Nuts, Pistachios, Other Nuts, and Mixed Nuts	
Walnuts	Pecans, Almonds, Cashews, Pistachios, Other Nuts, and Mixed Nuts	Macadamia Nuts
Macadamia Nuts	Peanuts, Pecans, Cashews, Pistachios, Other Nuts, and Mixed Nuts	Almonds and Walnuts
Pistachios	Pecans, Almonds, Cashews, Walnuts, Macadamia Nuts, Other Nuts, and Mixed Nuts	Peanuts
Other Nuts	Peanuts, Pecans, Cashews, Walnuts, Macadamia Nuts and Pistachios	Mixed Nuts
Mixed Nuts	Peanuts, Pecans, Almonds, Cashews, Walnuts, Macadamia Nuts, and Pistachios	Other Nuts

Source: Evidence from Table 2.9.

2.6.5. Expenditure and Income Elasticities

All expenditure elasticities were positive and statistically different from zero. As shown in Table 2.11, almonds had the highest expenditure elasticity, 1.1382, while macadamia nuts had lowest expenditure elasticity, 0.7725. Pecans, almonds, walnuts, and pistachios were more sensitive to changes in total expenditure of peanuts and tree nuts than peanuts, cashews, macadamia nuts, mixed nuts, and other nuts. Using equation (14), calculated income elasticities varied from 0.5042 to 0.7429, indicating that peanuts and tree nuts were necessities.

Table 2.11 Expenditure and Income Elasticities

	Expenditure Elasticities	*Income Elasticities*[a]
	Estimate	**Estimate**
Peanuts	**0.9534**	**0.6223**
Pecans	**1.0672**	**0.6966**
Almonds	**1.1382**	**0.7429**
Cashews	**0.9911**	**0.6469**
Walnuts	**1.0403**	**0.6790**
Macadamia Nuts	**0.7725**	**0.5042**
Pistachios	**1.1313**	**0.7384**
Other Nuts	**0.9796**	**0.6394**
Mixed Nuts	**0.8419**	**0.5495**

Note: The estimated expenditure elasticities correspond to the means of 144 monthly values for the period 2004 to 2015. Bold numbers indicate significance at the 5% level.

[a] Income elasticities are calculated using equation (14) by multiplying estimates of expenditure elasticities by the coefficient, 0.6527. The significance of income elasticities is based on the statistical significance of the expenditure elasticities.

Source: Nielsen Homescan Panel 2004-2015, and calculations by the author.

In Table 2.12, we provide a comparison of the results from our article with those from the extant literature. Only three previous studies dealt with demand interrelationships of peanuts and tree nuts. As stated previously these studies are not up-to-date regarding data used and models employed. Our work is current to these studies in terms of the use of more recent data and advanced generalized demand systems model. Comparisons were only possible for pecans, walnuts, and almonds. Except for the Lerner (1959), our own-price elasticities for pecans, walnuts, and almonds were larger in magnitude than those previously estimated due in part to the fact that we included a richer delineation of nut products as well as incorporation of pre-commitment levels and variety-seeking behavior. As such, we allow for a more thorough specification of the demand interrelationships of peanuts and tree nuts coupled with the used of more current data than previous studies.

2.7. Concluding Remarks

We utilized the LA/GAIDS model incorporating variety-seeking behavior and pre-commitment levels to address the demand for various nut products in the United States[6]. Variety-seeking behavior was supported for peanuts but not supported for almonds. The finding offers insights for nuts processors and food companies regarding packaging strategies. Processor and purveyors of nut products are advised to combine peanuts with other products to enhance variety. Strategically, almonds should be sold individually without any combination with other nut products.

Pre-commitment levels were evident for pecans, almonds, walnuts, macadamia nuts, pistachios, and other nuts. Although most of pre-committed quantities were

---

[6] The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

negative, almost all of them generally exhibited an upward trend over the period from 2004 through 2015. As such, households were relatively more committed to purchasing peanuts and tree nuts over this period.

Due to different types of data, time periods, and modeling techniques, our results regarding own-price and cross-price elasticities differed from previous studies. Since the demand for peanuts, almonds, cashews, walnuts, pistachios, other nuts and mixed nuts was found to be inelastic, appropriate price strategies to increase sales would be to raise prices for these nut products. In general, the various nut products were net substitutes. Although substitution among nut products was far more common, some complementarity among nut products was also evident. Information concerning the substitution and complementary patterns among peanuts and tree nuts is important strategically for stakeholders in the nuts market. As well, the nut products were found to be necessities. This finding relates well in the recommendation of peanuts and tree nuts as healthy snacks.

Table 2.12 Comparison of Models, Data, Nut Products, Own-Price Elasticities with Other Studies in the Literature

Study	Model	Data	Nut Products	Own-Price Elasticity	
This article	GQUAIDS with Entropy Index	Monthly time-series 2004-2015 from Nielsen HomeScan Panel	Peanuts	-0.8180	
			Pecans	-0.9620	
			Almonds	-0.7035	
			Cashews	-0.7685	
			Walnuts	-0.5713	
			Macadamia Nuts	-1.0560	
			Pistachios	-0.7997	
			Mixed Nuts	-0.9416	
			Other Nuts	-0.4503	
Lerner (1959)	Single Equation Seemingly unrelated regression	Annually time – Series, 1922-1955	Seedling pecans	-2.729	
			Improved pecans	-3.442	
			Pecans	-1.188	
			Walnuts	-1.803	
			Filberts	-23.042	
			Almonds	-0.863	
				(Linear Function Form)	(Double Logarithmic)
Dhaliwal (1972)	Single Equation	Annually time-series 1922-1955	Pecans	-0.909	-0.856
			Walnuts	-0.286	-0.420
			Filberts	-1.926	-0.891
			Almonds	-0.548	-0.888
				(Single Equation)	(Seemingly Unrelated Regression)
Russo, Green, and Howitt (2008)	Single Equation Seemingly unrelated regression	Annually time-series 1970-2001	Almonds	–0.480 to -0.350[a]	-0.140
			Walnuts	–0.266 to -0.284	-0.200

[a] The range of estimates is due to different function forms, including linear, double log, and Box-Cox specifications.

Source: Compilations from the author.

Declining market penetration of peanuts and tree nuts over the period 2004 through 2015 signals a decreasing number of households who purchased peanuts and tree nuts for home use. As such, a noteworthy implication is that nut processors and nut purveyors need to expend effort in increasing the market penetration of their product for home use.

This article addresses demand interrelationships for peanuts and tree nuts along with variety-seeking behavior and pre-commitment using monthly time-series data across households in the United States over the period 2004 to 2015. Future work could focus on individual households in order to get a micro-perspective viewpoint as to how variety-seeking behavior and pre-commitments affect the demand for peanuts and tree nuts. In this way, censored demand models could be developed to analyze individual household purchases of peanuts and tree nuts incorporating socio-demographic factors such as age, household size, region, education, and ethnicity.

A limitation of our analysis concerns the implicit assumption of separability of peanuts and tree nuts from other snack products. Additionally, although the LA/GAIDS model accounts for price, total expenditure, seasonality, variety-seeking behavior and pre-commitment, other explanatory factors were excluded from the analysis. Branded and generic advertising expenditures were not included in this analysis due to unavailability of monthly data over the period 2004 to 2015. Perhaps other macroeconomic factors could have been included in the LA/GAIDS model. Despite the omission of these additional explanatory variables, these factors were part of the error terms. Because of the use of the iterative SUR estimation procedure, these omitted variables were implicitly accounted for in the analysis. Moreover, the household regional heterogeneity of the nuts market could be addressed in future studies. Despite these limitations, our article adds measurably to the extant literature for the nuts market by way of a more up-to-date and more thorough demand system analysis incorporating seasonality, pre-commitment and variety-seeking behavior.

2.8. References

Anderson, S. P., De Palma, A., Thisse, J. F., 1992. Discrete choice theory of product differentiation. MIT press.

Attfield, C. L., 1985. Homogeneity and endogeneity in systems of demand equations. J. Econ. 27(2), 197–209.

Banks, J., Blundell, R., Lewbel, A., 1997. Quadratic engel curves and consumer demand. Rev. Econ. Stat. 79(4), 527–539.

Benassy, J.-P., 1996. Taste for variety and optimum production patterns in monopolistic competition. Econ. Lett. 52(1), 41–47.

Berndt, E. R., Savin, N. E., 1975. Estimation and hypothesis testing in singular equation systems with autoregressive disturbances. Econometrica: Journal of the Econometric Society pp. 937–957.

Bollino, C. A., 1987. Gaids: a generalised version of the almost ideal demand system. Econ. Lett. 23(2), 199–202.

Capps Jr, O., Tsai, R., Kirby, R., Williams, G. W., 1994. A comparison of demands for meat products in the pacific rim region. J. Agric. Res. Econ. pp. 210–224.

De Groot, H. L., Nahuis, R., 1998. Taste for diversity and the optimality of economic growth. Econ. Lett. 58(3), 291–295.

Deaton, A., Muellbauer, J., 1980. An almost ideal demand system. Am. Econ. Rev. 70(3), 312–326.

Dhaliwal, H. S., 1972. An econometric investigation of demand interrelationships among tree nuts and peanuts. Ph.D. thesis, Oregon State University.

Dharmasena, S., Capps, O., 2012. Intended and unintended consequences of a proposed national tax on sugar-sweetened beverages to combat the us obesity problem. Health Econ. 21(6), 669–694.

Diewert, W., 1987. (1987): Index numbers. The new Palgrave dictionary of economics. Vol. 2 p. 767.

Drescher, L., Thiele, S., Weiss, C. R., 2008. The taste for variety: A hedonic analysis. Econ. Lett. 101(1), 66–68.

Dubé, J.-P., 2004. Multiple discreteness and product differentiation: Demand for carbonated soft drinks. Marketing Science 23(1), 66–81.

Global, I., 2016. Eviews 9. Irvine: Quantitative Micro Software.

Gorman, W., 1976. 11 tricks with utility functions. Essays in Economic Analysis: The Proceedings of the Association of University Teachers of Economics Sheffield 1975 p. 211.

US Department of Health and Human Services., 2015. 2015–2020 dietary guidelines for americans. Washington (DC): USDA.

Hovhannisyan, V., Gould, B. W., 2011. Quantifying the structure of food demand in china: An econometric approach. Agric. Econ. 42, 1–18.

Hovhannisyan, V., Gould, B. W., 2014. Structural change in urban Chinese food preferences. Agric. Econ. 45(2), 159–166.

Insititute, S., 2012. Sas 9.4 for windows. SAS Institue Inc., Cary, NC, USA .

Jackson, L. F., 1984. Hierarchic demand and the engel curve for variety. Rev. Econ. Stat. pp. 8–15.

Jevons, W. S., 1871. The Theory of Polical Economy.

Kim. Y., and Dharmasena, S., 2017. Forthcoming. “Price Discovery and Integration in the United States Pecan Markets.” Journal of Food Distribution Research, March 2018 proceedings edition.

King, J. C., Blumberg, J., Ingwersen, L., Jenab, M., Tucker, K. L., 2008. Tree nuts and peanuts as components of a healthy diet. The Journal of nutrition 138(9), 1736S–1740S.

Kris-Etherton, P. M., Hu, F. B., Ros, E., Sabaté, J., 2008. The role of tree nuts and peanuts in the prevention of coronary heart disease: multiple potential mechanisms. The Journal of nutrition 138(9), 1746S–1751S.

Lancaster, K., 1990. The economics of product variety: A survey. Marketing science 9(3), 189–206.

Lee, I. M., 1950. Statistical Analysis of Annual Average Returns to Growers of Almonds, 1924/25 to 1948/49. 103, University of California, College of Agriculture, Agriculture Experiment Station.

Lerner, E. B., 1959. An econometric analysis of the demand for pecans with special reference to the demand interrelationships among domestic tree nuts. Ph.D. thesis, Oklahoma State University.

Matsuda, T., 2006. Linear approximations to the quadratic almost ideal demand system. Empirical Econ. 31(3), 663–675.

Nielsen, S. J., Kit, B. K., Ogden, C. L., 2014. Nut consumption among US adults, 2009-2010. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

O'Neil, C. E., Keast, D. R., Fulgoni, V. L., Nicklas, T. A., 2010. Tree nut consumption improves nutrient intake and diet quality in us adults: an analysis of national health and nutrition examination survey (nhanes) 1999-2004. Asia Pacific journal of clinical nutrition 19(1), 142.

Pollak, R. A., Wales, T. J., 1981. Demographic variables in demand analysis. Econometrica: Journal of the Econometric Society pp. 1533–1551.

Rehm, C. D., Drewnowski, A., 2017. Replacing american snacks with tree nuts increases consumption of key nutrients among us children and adults: Results of an nhanes modeling study. Nutrition journal 16(1), 17.

Rowland, C. S., Mjelde, J. W., Dharmasena, S., 2017. Policy implications of considering pre-commitments in us aggregate energy demand system. Energy Policy 102, 406–413.

Russo, C., Green, R., Howitt, R. E., 2008. Estimation of supply and demand elasticities of california commodities .

Seegmuller, T., 2008. Taste for variety and endogenous fluctuations in a monopolistic competition model. Macroeconomic Dynamics 12(4), 561–577.

Tonsor, G. T., Marsh, T. L., 2007. Comparing heterogeneous consumption in us and japanese meat and fish demand. Agric. Econ. 37(1), 81–91.

Wells, G. J., Miller, S. E., Thompson, C. S., 1986. Farm level demand for pecans reconsidered. J. Agric. Appl. Econ. 18(1), 157–160.

Zhen, C., Finkelstein, E. A., Nonnemaker, J. M., Karns, S. A., Todd, J. E., 2013. Predicting the effects of sugar-sweetened beverage taxes on food and beverage demand in a large demand system. Am. J. Agric. Econ. 96(1), 1–25.

3. STUDY II: HOUSEHOLD CHOICE TO PURCHASE PEANUTS AND TREE NUTS IN THE UNITED STATES: EVIDENCE FROM THE NIELSEN HOMESCAN PANEL

3.1. Introduction

The nut industry is a notable component of U.S. agricultural economy. According to the United States Department of Agriculture (USDA) Economic Research Service (ERS), the total crop values of peanuts and tree nuts were $11.74 billion and $9.02 respectively for 2018. Meanwhile, health benefits of nuts products have been widely documented in the literature. The consumption of nuts was found to reduce incidence of coronary heart disease, gallstones, diabetes, hypertension, cancer, and inflammation (Fraser et al. 1992; Kris-Etherton et al. 2008; Ros 2010); to decrease body mass index (BMI) (Sabaté 2003; King et al. 2008; Mattes, Kris-Etherton, and Foster 2008). In the latest *Dietary Guidelines for Americans 2015-2020*, nuts are included in the spectrum of healthy foods. However, relatively little is known not only the demand for peanuts and tree nuts but also the factors affecting the decision to purchase these products. The latter is the focus of this study.

There are 127.59 million households in the United States reported by the Census Bureau. About 50% of them purchased peanuts and 80% of them consumed tree nuts products in the year of 2015[7]. However, little is known about their demands as well as the drivers of purchases.

Aiming at examining household choice to purchase peanuts and tree nuts in the United States, the principle objectives of this study are to determine the impacts of socio-demographic factors on the purchase decision to nuts and to develop profile of target households to strategically assist stakeholders in the nuts market. In order to meet these objectives, a cross-sectional pooled probit analysis was conducted using the Nielsen Homescan Panel for calendar year 2015.

This paper is organized as follows. First, we review previous economic studies related to tree nuts and peanuts. Second, we introduce the binary choice probit model and discuss the data associated with this analysis. Third, we present a discussion of the estimation of the respective probit models with attention given to marginal effects. Finally, we make concluding remarks and discuss limitations of this study.

3.2. Related Literature

The focus of this study is to examine the household-level choices in the decision to purchase peanuts and tree nuts in the United States. To the best of our knowledge, research in this area is quite limited. Only several studies exist which deal with national and regional demand for nuts products, demand interrelationships, and consumer choices of snack peanuts. No studies in the extant literature exists concerning household choice to purchase peanuts and tree nuts in the United States. This study fills this research void.

---

[7] We calculated the market penetration using the Nielsen Homescan Panel 2015 data. The households in the sample are used as a projection of the entire nation.

The literature related to the demand for peanuts and tree nuts is summarized below. Lee (1950), Lerner (1959), Dhaliwal (1972), Wells, Miller, and Thompson (1986), and Russo, Green, and Howitt (2008), examined the demand for several nuts products. All of these studies used annual time-series data at the market level and single-equation model. None of these studies looked at the driving forces of demand at the household-level due to limitations of data availability and methods.

Rimal and Fletcher (2002) investigated the effects of socioeconomic characteristics on market participation and frequency of purchase of snack peanuts. Using a double-hurdle and Cragg model, they found that household's nutritional considerations affected the purchase frequency but not the market participation. Similarly, He, Fletcher, and Rimal (2005) utilized survey and multinomial logit model and explored the effects of age, education, race, and nutrition intakes on consumer preferences of six types of snack peanuts.

Kim and Dharmasena (2018) examined prices received by growers of pecans from 2005 to 2016 to investigate market integration patterns in Texas, Oklahoma, Georgia, and Louisiana using causality structures identified through machine learning methods. Green (1999) estimated the demand for almonds in California by considering prices of tree nuts, per capita income, and socio-demographic factors. The own-price elasticity of demand for U.S. almonds was found to be -0.83. No studies ever looked through a micro-perspective viewpoint as to how socio-demographic factors affect purchase decisions.

Overall, our study differs from those in the extant literature by focusing on socio-demographic factors in affecting the decision to purchase peanuts and tree nuts. As well, we provide a look at specific tree nuts, a more granular depiction. The nuts considered in this analysis are: (1) almonds; (2) cashews; (3) macadamia nuts; (4) mixed nuts; (5) peanuts; (6) pecans; (7) pistachios; (8) walnuts.

3.3. Methodology

Due to the household-level data used and to serve the purpose of this study, we employ a binary choice model to capture household choices regrading peanuts and tree nuts. In this section, we discuss the probit model, initially introduced by Bliss (1934a; 1934b), to estimate the decision to purchase peanuts and tree nuts.

The data used in this study are from Nielsen Homescan Panel for calendar year 2015. Household recorded information concerning expenditures, quantities purchased, and coupon value. Socio-demographics factors pertaining to household size, household income, education, race, ethnicity, age of household head, and region/location. We consider 61,830 households in the sample across four quarters of calendar year 2015. As such, the data comprise a panel, and the total number of observations is 245,520. The data for calendar year 2015 were the most current available to us. However, not every household purchased all the nuts products at each time point (quarter)[8], meaning our data are censored where zero purchases were observed.

---

[8] We calculated the market penetration for each tree nuts and peanuts products using Nielsen Homescan Panel 2015, that is peanuts 23%, pecans 7%, almonds 13%, cashews 12%, walnuts 9%, macadamia nuts 1%,

3.3.1. Pooled Probit Model

We employ a pooled probit model to capture the purchase decision of households regarding nuts products. Let Z_{hit} denote the purchase decision of household h on product i at time point t, that takes the value of one if this household purchases a nut product, and the value of zero if not. The pooled probit model is given as follows,

(1) $\Pr[Z_{hit} = 1] = \Phi(\boldsymbol{W}_{ht}\beta_i)$

and

(2) $\Pr[Z_{hit} = 0] = 1 - \Phi(\boldsymbol{W}_{ht}\beta_i), \quad h = 1, \dots, 61{,}830 \quad i$

$= \text{peanuts}, \dots, \text{mixed nuts} \quad t = 1, 2, 3, 4$

where $\Pr(\cdot)$ indicates probability, $\Phi(\cdot)$ is cumulative standard normal distribution function, and $\boldsymbol{W}$ consists of the set of explanatory variables used in the probit model at time point t in order to capture household "participation". The parameter β_i is the set of coefficients associated with the respective explanatory variables. We include quarterly dummies, own-price, household size, annual income, age of household head, education level of household head, region, race, and presence of children.

Quarterly dummies are used to capture seasonal patterns regarding the decision to purchase nuts. Household size delineates the number of in the household. All socio-demographic factors provide insights for nut purveyors about market segmentation and

---

pistachios 7%, mixed nuts 11%, tree nuts 46%. The market penetration is defined as the number of purchases corresponding category divided by the total number of observations (61,380× 4) in the panel year of 2015 (see Table 2b).

targeting strategies. Households with larger size, higher income level, and without children are expected to have a higher propensity purchasing nuts that associated coefficients are anticipated to be positive. As consumer age increases, they are expected to purchase more nuts products to meet healthy requirement and concerns. Meanwhile, household with higher education tend to understand nutrition information and take over new lifestyles that coefficient associated with this factor is expected to be increase as education level goes up. The model takes the following form,

(3) $$\Phi(\boldsymbol{W}_{ht}\beta_i)$$

$$= \Pr\Bigg(\mu_i + \sum_{k=1}^{3} \alpha_{ki} q_{khit} + \beta_i P_{hit} + \beta_{1i} \ln HZ_h + \beta_{2i}\, ln\, INC_h + \beta_{3i} Age_{1h}$$

$$+ \beta_{4i} Age_{2h} + \beta_{5i} Age_{3h} + \beta_{6i} Edu_{1h} + \beta_{7i} Edu_{2h} + \beta_{8i} Edu_{3h} + \beta_{9i} NE_h$$

$$+ \beta_{10i} MA_h + \beta_{11i} ENC_h + \beta_{12i} WNC_h + \beta_{13i} SA_h + \beta_{14i} ESC_h$$

$$+ \beta_{15i} WSC_h + \beta_{16i} Mou_h + \beta_{17i} WH_h + \beta_{18i} BL_h + \beta_{19i} AS_h + \beta_{20i} \mathrm{NC}_h$$

$$+ \beta_{21i} NC_h + e_{hit} > 0\Bigg),$$

where

i	–	Nine peanuts and tree nuts products;
h	–	61,830 households;
t	–	Four quarters;
μ_i	–	Constant terms for product i
q_{khit}	–	Seasonal dummy of quarter k for household h and product i at time period t;
P_{hit}	–	Unit price paid by household h for product i at time period t;
HZ_h	–	Household size for household h;

INC_h	–	Household h annual income at time period t;
Age_{1h}	–	Age of household h under the age of 35;
Age_{2h}	–	Age of household h between the age of 35 and 49;
Age_{3h}	–	Age of household h between the age of 50 and 64;
Edu_{1h}	–	Less than high school education of household h;
Edu_{2h}	–	High school education of household h;
Edu_{3h}	–	Some college education of household h;
NE_h	–	Household h located in the New England region;
MA_h	–	Household h located in the Middle Atlantic region;
ENC_h	–	Household h located in the East North Central region;
WNC_h	–	Household h located in the West North Central region;
SA_h	–	Household h located in the South Atlantic region;
ESC_h	–	Household h located in the East South Central region;
WSC_h	–	Household h located in the West South Central region;
Mou_h	–	Household h located in the Mountain region;
WH_h	–	Household h white/Caucasian;
BL_h	–	Household h Black;
AS_h	–	Household h Asian;
HR_h	–	Household h Hispanic origins;
NC_h	–	Household h without children under the age of 18;
e_{hit}	–	Disturbance terms.

The base group for age category is age 65 and above; for education is graduate level; for region is Pacific area; for race is others. All these base groups were excluded from the analysis due to collinearity in the covariance-variance matrix. We also labelled each factor with the hypotheses that positive and negative relationships are anticipated.

The probit model is estimated using maximum-likelihood method. The log-likelihood function for probit is

(4) $\ln \mathcal{L}(\gamma_i; Z_{hit}, \boldsymbol{W}_{ht}) = Z_{hit} \ln \Phi(\boldsymbol{W}_{ht}\beta_i) + (1 - Z_{hit}) \ln(1 - \Phi(\boldsymbol{W}_{ht}\beta_i))$

Assuming that the observations are independent and identically distributed.

3.3.2. Marginal Effects

We derive the marginal probability effects in this section as the partial effects of each explanatory variable on the probability that the observed dependent variable $\Pr(Z_{hit} = 1)$. There are two cases regarding the marginal effects for continuous explanatory variables and binary explanatory variables (such as quarterly dummies, regional dummies, etc.). For subsets of $\boldsymbol{W}$ that is continuous W_j, the marginal effects are as follows,

(5) $$\text{Marginal effects of } W_j = \frac{\partial \Pr(Z_{hit} = 1)}{\partial W_j} = \frac{\partial \Phi(\boldsymbol{W}_{ht}\beta_i)}{\partial W_j} = \phi(W_{ht}\gamma_i)\beta_i$$

where $\phi(\cdot)$ is the standard normal probability density function. Meanwhile, for the subset of $\boldsymbol{W}$ that correspond to binary variable W_j, the marginal effects are as follows,

(6) $$\text{Marginal effects of } W_j = \Phi(\boldsymbol{W}_{ht1}\beta_i) - \Phi(\boldsymbol{W}_{ht0}\beta_i)$$

The two different vectors of regressor values, (1) $\boldsymbol{W}_{ht1}$ is the vector of explanatory variable when W_j equals to 1; (2) $\boldsymbol{W}_{ht1}$ is the same vector of explanatory variable but when W_j takes value of 0. The marginal effects of the binary variables represent the incremental probability of the included category relative to the base reference category. The standard error of the marginal effect in equation (5) and (6) is calculated using the delta method.

3.4. Data

The data used for this study is derived based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business[9]. As previously stated, we used the data for calendar year 2015, where 61,380 households are included, and we subsequently aggregated household purchases on a quarterly basis. These data were the most current available data at the time of this investigation. Nine categories of nuts products are examined. We categorized peanuts and tree nuts based on product module codes and product descriptions provided by Nielsen. The eight categories are: (1) almonds, (2) cashews, (3) macadamia nuts, (4) mixed nuts, (5) peanuts, (6) pecans, (7) pistachios, (8) walnuts. We first aggregated nuts purchases for each household for each quarter using household identification number provided by Nielsen, including how much they paid and the total quantities (in ounces) purchased. Then we calculated the unit values by dividing expenditures by total quantities purchased. Due to the fact that not all of households make purchases at each time point (quarter), we delineate all of those purchases as zeroes. As such, the panel data consists of 61,380 households cross four quarters, which gives the total number of observations 245,520.

[9] Disclaimer: Researcher(s) own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

Table 3.1 exhibits the descriptive statistics for unit values/price ($/ounce) and quantities (ounce). Peanuts had the biggest market share with the average of 50.77 ounces purchased by households for one quarter. Among tree nuts products, mixed nuts were consumed the most, and followed by cashews, almonds, and pistachios. Macadamia nuts had the highest unit price, 96 cents per ounce, while mixed nuts yielded the lowest price among tree nuts products, on the average of 41 cents per ounce. The unit prices of other tree nuts products range from 45 to 96 cents, on the average, for cashews, almonds, walnuts, pistachios, pecans, and macadamia nuts. The average unit price of peanuts products are 48 cents. As exhibited in Figure 3-1, we calculated the percentage of purchases in the Nielsen data by dividing the number of actual non-zero purchases by the total number of observations, peanuts were 23%, and tree nuts varied from 1% to 13%.

Table 3.1 Descriptive Statistics of Quantities (oz) and Unit Prices ($/oz)

Categories	Obs	Mean	Std. Dev.	Min	Max
Quantity (Oz purchased) per household per year					
Peanuts	56,509	50.7700	91.9000	0.4200	3520.0000
Pecans	18,130	20.5000	31.5100	0.7500	2368.0000
Almonds	31,627	25.6200	50.8100	0.2000	4220.0000
Cashews	28,855	29.5700	35.1600	0.7500	1216.0000
Walnuts	21,662	25.0100	35.9500	1.2500	1792.0000
Macadamia Nuts	1,392	11.7300	15.2200	1.0000	187.0000
Pistachios	17,967	26.1200	33.5100	0.3500	615.5000
Mixed Nuts	27,699	36.1900	42.8400	1.2500	1428.0000
Unit Value ($/oz)[a] by nuts category					
Peanuts	245,520	0.1704	0.0813	0.0003	11.9900
Pecans	245,520	0.6276	0.0781	0.0050	3.3300

Table 3.1 Continued

Categories	Obs	Mean	Std. Dev.	Min	Max
Almonds	245,520	0.5299	0.1004	0.0003	8.9100
Cashews	245,520	0.4462	0.0585	0.0006	11.0000
Walnuts	245,520	0.5509	0.0786	0.0038	7.2400
Macadamia Nuts	245,520	0.9587	0.0595	0.0450	3.0391
Pistachios	245,520	0.5959	0.0804	0.0003	19.8580
Mixed Nuts	245,520	0.4148	0.0663	0.0003	4.3744

[a] We compute missing unit values using a regression algorithm, see the Appendix A for more details.

Note: Actual quantity purchased are not used in the analysis, they were transformed into binary response with value of zero and one.

Source: Nielsen Homescan Panel 2015; calculations by the author.

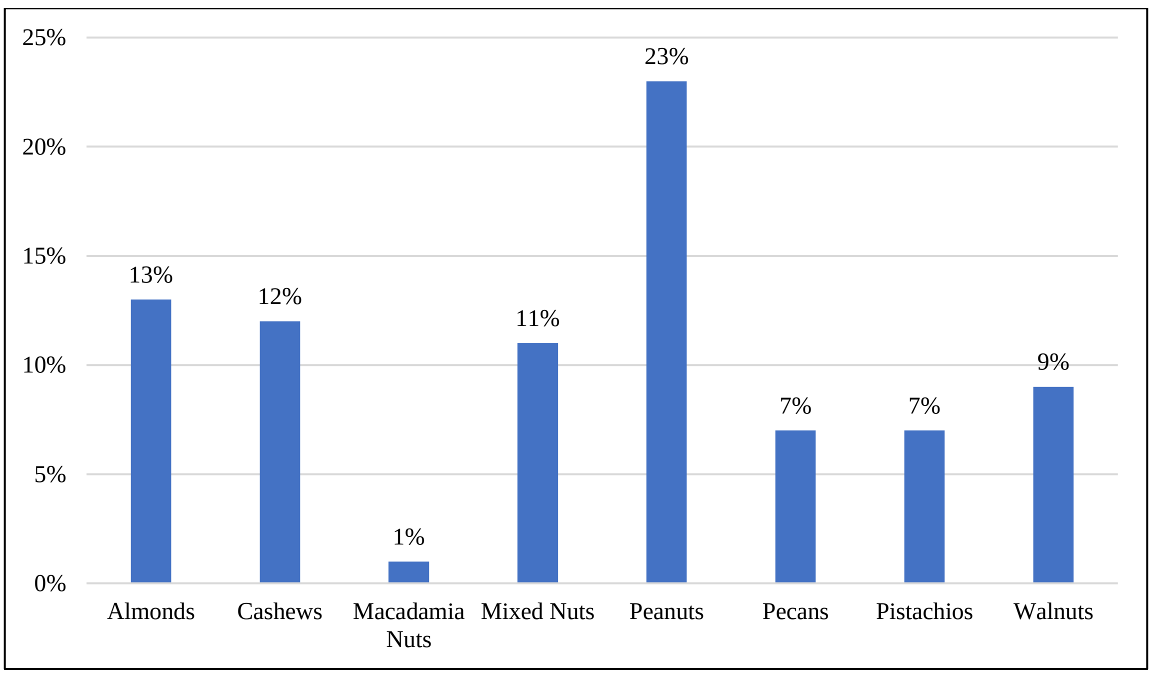


Figure 3-1 The Percentage (%) of Non-zero Purchases

Note: We calculate the percentage for respective nuts category by counting the number of observations who had actual purchase (obs in Table 1 for each quantity) and dividing it by the number of total observations (245,520) in the data.

Source: Nielsen Homescan Panel 2015; calculations by the author.

Nielsen also provided demographic information for each household in each calendar year. We selected eight dimensions. They are household size, annual income, age of household head, employment status of household head, highest education level of household head, region of country, race, ethnicity, and presence of children. In Figure 3-2, we plotted the household size in our data sample, roughly 67% households in Nielsen data have 1-2 members, and more than 90% households have under four members There are fifteen different household income categories as shown in Figure 3-3. On the average, annual income of households in the Nielsen data is around fifty-eight thousand dollars (see Table 1d).

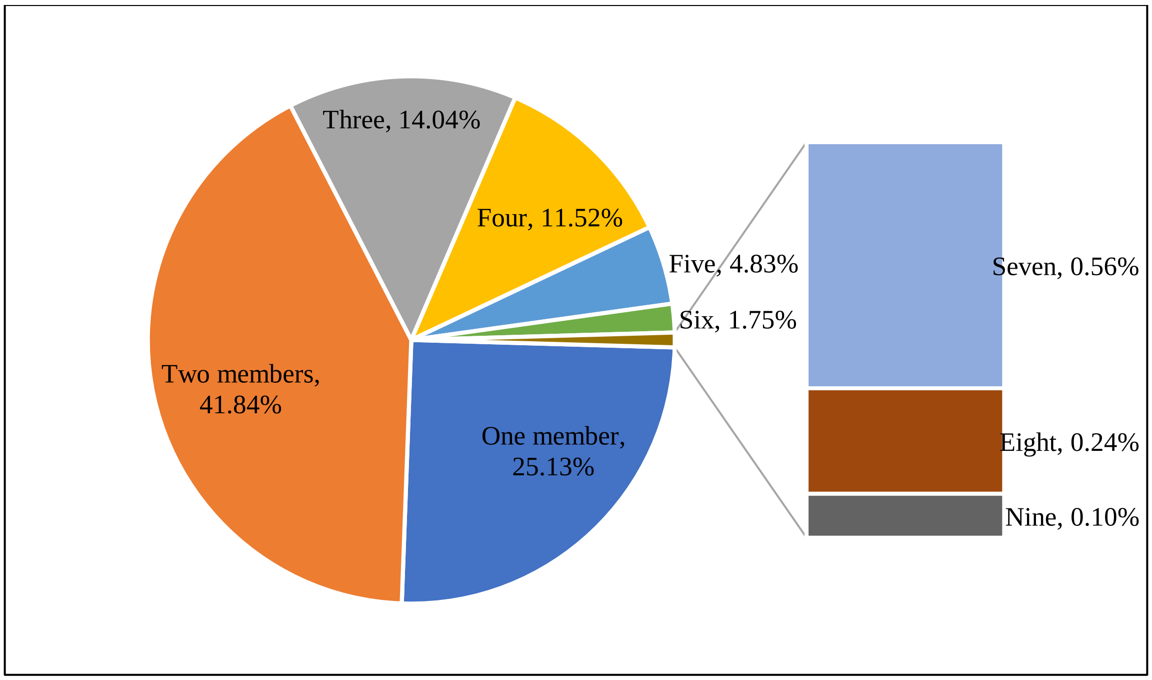


Figure 3-2 Household Size of the Panelists in the calendar year 2015

Note: Minimum is one household member and maximum is nine. Source: Nielsen Homescan Panel 2015.

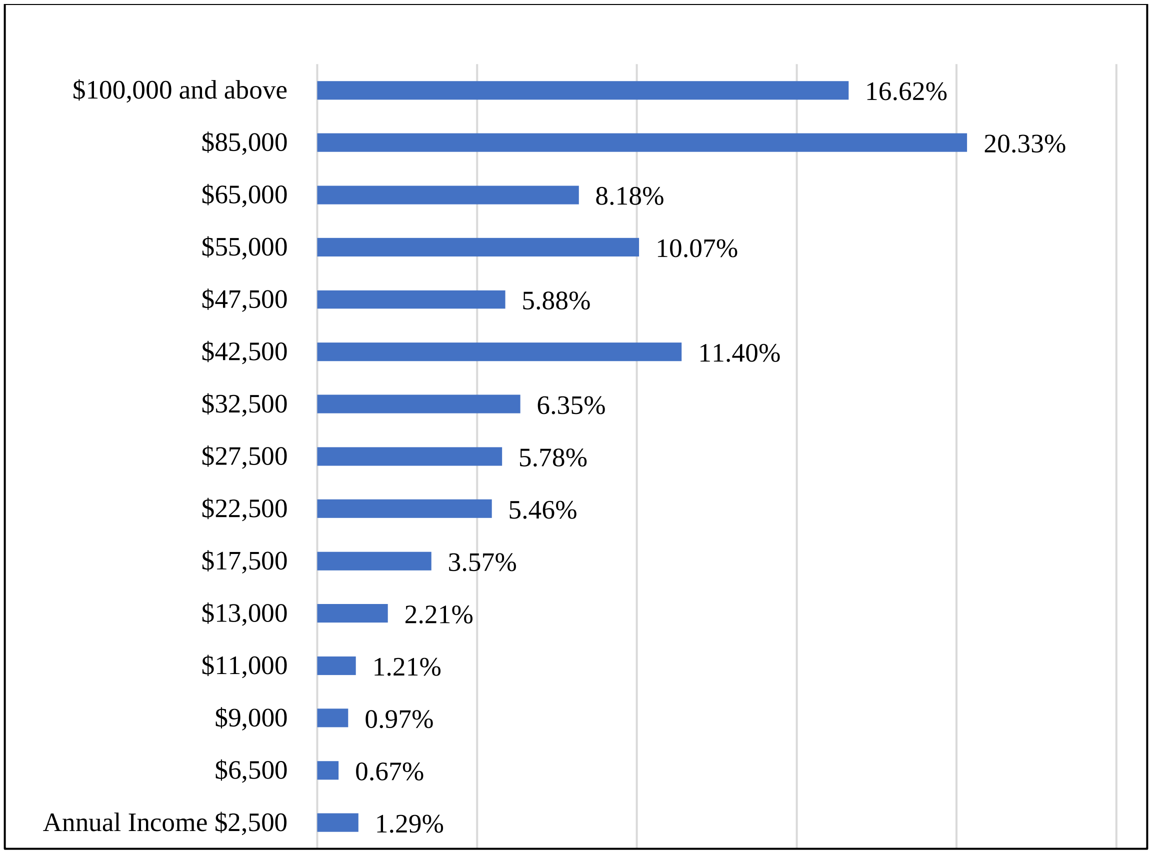


Figure 3-3 Income Levels of Households in the Nielsen Homescan Panel

Note: We used the income levels of Nielsen and counted the number of households in each category.

Source: Nielsen Homescan Panel 2015; calculations by the author.

Except for annual income ($), all these socio-demographic factors were coded into dummy variables used in the estimation process. Consequently, the mean is interpreted as the percentage and the base groups excluded in our analysis due to collinearity were labelled with asterisks in Table 3.2. Most of households are older people between age of 50 and 64. The samples were distributed unevenly nationwide according to the regional information, the largest share of households locates in the area of South Atlantic, around

21%. Least part of households located in New England region. We reported the list of regions according to U.S. Census Bureau in Figure 3-4. Under 20% households have education level less than college. More than four-fifths of households are white/Caucasian. Based on the information of presence of children, 78% households do not have any children under the age of 18.

Table 3.2 Descriptive Statistics for Panelist/Demographics

Variable/label		Description	Mean	Std. Dev.
Season				
	q1	The first quarter of calendar 2015	0.25	0.43
	q2	The second quarter of calendar 2015	0.25	0.43
	q3	The third quarter of calendar 2015	0.25	0.43
	*q4**	The fourth quarter of calendar 2015	0.25	0.43
Household Size				
	hsize	Number of household members	2.38	1.30
Income ($)				
	income	Household Annual Income	58,488	29,235
Age of household head				
	agehh_under35	Under the age of 35	0.08	0.26
	agehh_35to49	The age between 35 and 49	0.24	0.43
	agehh_50to64	The age between 50 and 64	0.43	0.50
	*agehh_65andabove**	The age of 65 and above*	0.25	0.43
Education level of household head				
	eduhh_lesshigh	Less than high school	0.01	0.11
	eduhh_highschool	High school	0.18	0.39
	eduhh_somecollege	Some college	0.29	0.45
	*eduhh_grad**	Graduate school*	0.52	0.50
Race				
	White	White	0.82	0.39
	Black	Black	0.11	0.31
	Asian	Asian	0.03	0.18
	*Other**	Other*	0.05	0.21
Ethnicity				
	hispanic_reg	Hispanic Origins	0.06	0.24
Region of country				
	NewEngland	New England	0.05	0.21
	MiddleAtlantic	Middle Atlantic	0.13	0.34

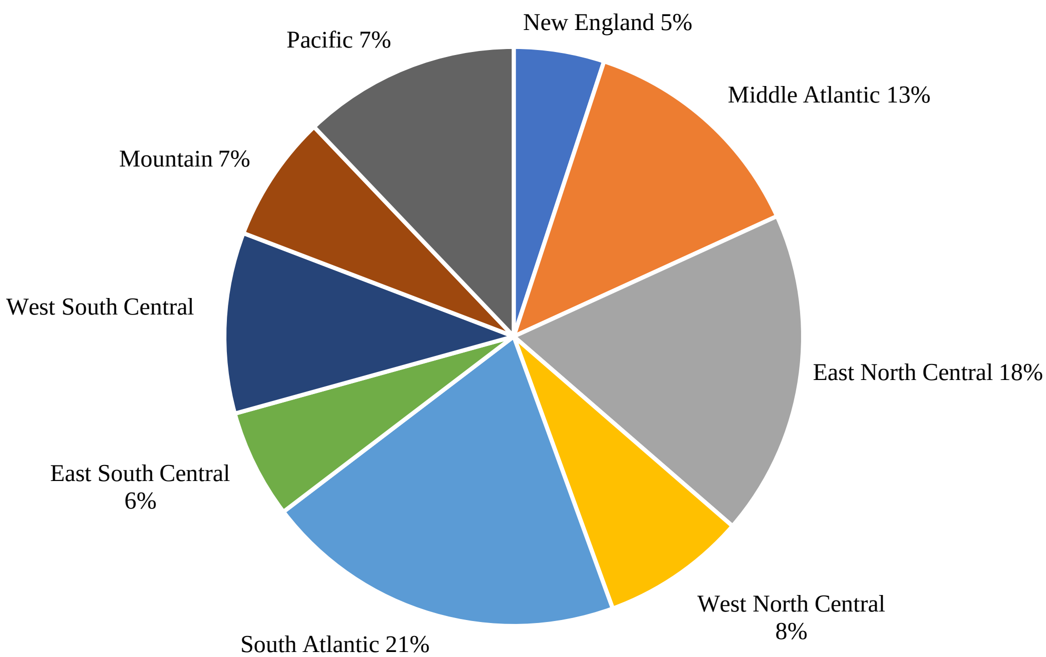
Table 3.2 Continued

Variable/label	Description	Mean	Std. Dev.
ENCentral	East North Central	0.18	0.38
WNCentral	West North Central	0.08	0.28
SouthAtlantic	South Atlantic	0.21	0.4
ESCentral	East South Central	0.06	0.24
WSCentral	West South Central	0.1	0.31
Mountain	Mountain	0.07	0.26
*Pacific**	Pacific*	0.12	0.33
Presence of children under 18			
no_child	No presence of children	0.78	0.42

Note: *The base group of dummies that are not included in the analysis. The list of regions is from U.S. Census Bureau, see Figure 3 for more details. Source: Nielsen Homescan Panel 2015; calculations by the author.

3.5. Empirical Results

The pooled probit model was estimated using Stata, Version 14 through the PROBIT and marginal probability effects were calculated using MARGINS commands respectively for each peanuts and tree nuts (Long 1997; Long and Freese 2014). We report the results of estimated coefficients and marginal probability effects in this section. Owing to the data-censoring issue, we need to compute the missing price variables that were included on the right-hand side. In short, we regressed the non-missing prices on selected socio-demographic factors and used the estimated coefficients to calculate missing prices for households along with the use of their corresponding socio-demographic factors. Then the missing prices were replaced with these predicted calculations. Please see the Appendix A as to how we imputed missing price variables.



- New England: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont
- Middle Atlantic: New Jersey, New York, and Pennsylvania
- East North Central: Indiana, Illinois, Michigan, Ohio, and Wisconsin
- West North Central: Iowa, Nebraska, Kansas, North Dakota, Minnesota, South Dakota, and Missouri
- South Atlantic: Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, and West Virginia
- East South Central: Alabama, Kentucky, Mississippi, and Tennessee
- West South Central: Arkansas, Louisiana, Oklahoma, and Texas
- Mountain: Arizona, Montana, Colorado, Utah, Idaho, Nevada, New Mexico, and Wyoming
- Pacific: Alaska, California, Hawaii, Oregon, and Washington

Figure 3-4 Locations of Households of Nielsen in the United States

Source: Nielsen Homescan Panel 2015; calculations by the author.

3.5.1. Coefficient Estimates and Joint Test

In order to investigate the driving forces of household-level demand for peanuts and tree nuts, we exploited a binary choice model. In Table 3.3, we provide the estimated results of Probit model, including coefficients, number of observations, model fitness, and significance levels. In each Probit model, we used own-price, seasonal dummies, household size, household annual income[10], age of household head, highest education of household head, regions of country, race, and presence of children. The nine categories of peanuts and tree nuts were estimated separately. Each equation has 245,520 observations. The model fitness, pseudo R_2, ranged from 0.026 to 0.043.

All own-price variables are statistically significant positive sign, meaning that household choices align with more expensive products of peanuts and tree nuts. All seasonal dummies are statistically significant at the 5% significance level for pecans, walnuts, macadamia nuts, pistachios, mixed nuts, and tree nuts. All household size and household annual income are also significant at the same level. Most of socio-demographic factors are statistically significant, indicative of driving forces of household-level choices in decision to purchase respective nuts.

---

[10] Household size (*hsize*) and household annual income (*income*) are transformed using logarithms form (*lhize* and *lincome*) in the binary choice model.

Table 3.3 Summary of Parameter Estimates of Probit Models for Peanuts and Tree Nuts

	(1) Peanuts	(2) Pecans	(3) Almonds	(4) Cashews	(5) Walnuts	(6) Macadamia Nuts	(7) Pistachios	(8) Mixed Nuts
up_peanuts	4.221***							
	(0.071)							
up_pecans		1.560***						
		(0.032)						
up_almonds			1.736***					
			(0.032)					
up_cashews				2.245***				
				(0.047)				
up_walnuts					1.756***			
					(0.036)			
up_macadamia						1.745***		
						(0.103)		
up_pistachios							1.598***	
							(0.037)	
up_mixed								2.315***
								(0.041)
q1	0.026***	-0.521***	0.179***	0.010	-0.478***	-0.184***	0.034***	-0.055***
	(0.008)	(0.011)	(0.010)	(0.009)	(0.010)	(0.027)	(0.011)	(0.009)
q2	-0.011	-0.577***	0.097***	-0.021**	-0.556***	-0.154***	-0.111***	-0.135***
	(0.008)	(0.011)	(0.009)	(0.009)	(0.011)	(0.026)	(0.011)	(0.010)
q3	-0.023***	-0.609***	0.002	-0.093***	-0.494***	-0.137***	-0.159***	-0.150***
	(0.008)	(0.011)	(0.009)	(0.009)	(0.010)	(0.026)	(0.011)	(0.010)
lhsize	0.253***	0.174***	0.080***	0.147***	0.147***	0.092***	0.107***	0.176***
	(0.008)	(0.011)	(0.009)	(0.009)	(0.010)	(0.026)	(0.010)	(0.009)
lincome	0.036***	0.130***	0.168***	0.060***	0.102***	0.070***	0.124***	0.060***
	(0.005)	(0.007)	(0.006)	(0.006)	(0.006)	(0.018)	(0.007)	(0.006)
agehh_under35	-0.489***	-0.346***	-0.092***	-0.428***	-0.534***	-0.064	-0.335***	-0.555***
	(0.014)	(0.019)	(0.015)	(0.017)	(0.019)	(0.045)	(0.019)	(0.018)
agehh_35to49	-0.268***	-0.276***	-0.021*	-0.249***	-0.415***	-0.016	-0.165***	-0.363***
	(0.009)	(0.013)	(0.011)	(0.011)	(0.012)	(0.032)	(0.013)	(0.011)

Table 3.3 Continued

agehh_50to64	-0.098***	-0.098***	0.030***	-0.070***	-0.188***	0.023	-0.028***	-0.112***
	(0.007)	(0.010)	(0.008)	(0.008)	(0.009)	(0.024)	(0.010)	(0.008)
eduhh_lesshigh	-0.111***	-0.190***	-0.317***	-0.139***	-0.164***	-0.311**	-0.204***	-0.147***
	(0.028)	(0.043)	(0.038)	(0.034)	(0.038)	(0.142)	(0.043)	(0.035)
eduhh_highschool	-0.029***	-0.068***	-0.194***	-0.067***	-0.070***	-0.104***	-0.054***	-0.091***
	(0.008)	(0.011)	(0.010)	(0.010)	(0.011)	(0.030)	(0.011)	(0.010)
eduhh_somecollege	-0.014**	-0.034***	-0.103***	-0.029***	-0.062***	-0.025	-0.005	-0.050***
	(0.007)	(0.009)	(0.008)	(0.008)	(0.009)	(0.023)	(0.009)	(0.008)
NewEngland	0.115***	-0.033	0.019	-0.019	0.172***	-0.112**	0.129***	-0.005
	(0.016)	(0.023)	(0.018)	(0.019)	(0.019)	(0.047)	(0.021)	(0.019)
MiddleAtlantic	0.033***	-0.203***	-0.031**	-0.018	0.124***	0.062*	0.084***	-0.061***
	(0.012)	(0.018)	(0.013)	(0.014)	(0.015)	(0.036)	(0.015)	(0.014)
ENCentral	0.165***	0.119***	0.023*	0.117***	0.186***	-0.294***	0.085***	0.032**
	(0.011)	(0.016)	(0.012)	(0.013)	(0.014)	(0.036)	(0.015)	(0.013)
WNCentral	0.144***	0.136***	0.034**	0.071***	-0.020	-0.318***	0.001	0.129***
	(0.013)	(0.018)	(0.015)	(0.015)	(0.017)	(0.044)	(0.018)	(0.015)
SouthAtlantic	0.140***	0.150***	0.044***	0.090***	0.066***	-0.197***	0.062***	0.151***
	(0.011)	(0.015)	(0.012)	(0.013)	(0.014)	(0.034)	(0.014)	(0.012)
ESCentral	0.164***	0.198***	-0.026	0.065***	0.002	-0.101**	-0.040**	0.196***
	(0.014)	(0.020)	(0.017)	(0.017)	(0.019)	(0.049)	(0.020)	(0.017)
WSCentral	0.091***	0.281***	0.007	0.001	0.040**	-0.255***	-0.047***	0.112***
	(0.012)	(0.017)	(0.014)	(0.015)	(0.016)	(0.040)	(0.017)	(0.014)
Mountain	0.035***	0.183***	-0.050***	0.089***	-0.002	-0.089**	0.076***	0.078***
	(0.014)	(0.019)	(0.015)	(0.016)	(0.018)	(0.040)	(0.018)	(0.016)
White	0.014	0.036*	0.054***	-0.051***	0.037*	-0.090**	0.000	-0.003
	(0.015)	(0.021)	(0.017)	(0.018)	(0.020)	(0.046)	(0.020)	(0.018)
Black	-0.018	0.107***	-0.106***	0.006	-0.037	-0.372***	0.039*	-0.012
	(0.017)	(0.024)	(0.020)	(0.020)	(0.023)	(0.061)	(0.023)	(0.021)
Asian	0.004	-0.208***	0.003	-0.010	0.005	-0.031	0.095***	0.065**
	(0.022)	(0.033)	(0.025)	(0.026)	(0.029)	(0.064)	(0.028)	(0.026)
hispanic_reg	-0.033**	-0.001	0.007	-0.082***	0.004	-0.072	0.058***	-0.012
	(0.013)	(0.018)	(0.015)	(0.016)	(0.018)	(0.044)	(0.017)	(0.016)
no_child	0.182***	0.127***	0.080***	0.150***	0.188***	0.107***	0.051***	0.191***
	(0.010)	(0.014)	(0.011)	(0.012)	(0.013)	(0.034)	(0.013)	(0.012)

Table 3.3 Continued

Constant	-2.134***	-3.673***	-4.050***	-2.921***	-3.196***	-4.727***	-3.790***	-2.906***
	(0.054)	(0.080)	(0.067)	(0.066)	(0.074)	(0.208)	(0.079)	(0.066)
McFadden R^2	0.029	0.071	0.034	0.026	0.053	0.034	0.027	0.043

Standard errors in parentheses
* p<0.1 ** p<0.05 *** p<0.01

Note: up_nut = unit value for each nut product (also used as proxy for price)
Source: Nielsen Homescan Panel 2015

As shown in Table 3.4, joint tests were conducted to test the socio-demographic factors, examining significant differences between these groups and their base one (see Table 1d). For almonds, cashews, mixed nuts, peanuts, pecans, pistachios, and walnuts, all factors are statistically significant at 5% level jointly. For macadamia nuts, age of household head is not a driving force and race of households is not a significant driver for overall tree nuts. As introduced in the previous section, marginal probability effects are calculated using coefficients estimated in each probit model.

Table 3.4 Joint Tests Associated with Socio-Demographic Factors in Probit Model

	Almonds		Cashews	
Joint Tests	Chi-sq Statistics	p-value	Chi-sq Statistics	p-value
Age of household head	**86.11**	0.00	**898.07**	0.00
Education of household head	**478.19**	0.00	**60.18**	0.00
Regions of country	**82.32**	0.00	**232.54**	0.00
Race	**206.35**	0.00	**36.75**	0.00
	Peanuts		Pecans	
	Chi-sq Statistics	p-value	Chi-sq Statistics	p-value
Age of household head	**1556.89**	0.00	**566.95**	0.00
Education of household head	**25.85**	0.00	**52.45**	0.00
Regions of country	**429.80**	0.00	**1053.08**	0.00
Race	**12.48**	0.01	**124.09**	0.00
	Macadamia Nuts		Mixed Nuts	
	Chi-sq Statistics	p-value	Chi-sq Statistics	p-value
Age of household head	5.04	0.17	**1512.05**	0.00
Education of household head	**16.12**	0.00	**102.33**	0.00
Regions of country	**150.32**	0.00	**513.64**	0.00
Race	**51.83**	0.00	**13.57**	0.00
	Pistachios		Walnuts	
	Chi-sq Statistics	p-value	Chi-sq Statistics	p-value
Age of household head	**400.33**	0.00	**1440.01**	0.00
Education of household head	**43.40**	0.00	**79.81**	0.00
Regions of country	**174.93**	0.00	**380.74**	0.00
Race	**27.79**	0.00	**38.19**	0.00

Table 3.4 Continued

Note: Joint tests are used to test socio-demographic coefficients jointly, indicative of the significant difference between the base group and others. Numbers in boldness are significant under 1% significance level. The null hypotheses are:

agehh_under3= *agehh_35to49*= *agehh_50to64*=0;

eduhh_lesshigh= *eduhh_highschool*= *eduhh_somecollege*=0;

NewEngland= *MiddleAtlantic*= *ENCentral*= *WNCentral*= *SouthAtlantic*= *ESCentral*= *WSCentral*= *Mountain*=0;

White= *Black*= *Asian*=0.

Source: Nielsen Homescan Panel 2015; calculations by the author.

3.5.2. Marginal Probability Effects

For a continuous covariate, including own-price, log of household size, and log of household annual income, we used MARGINS in Stata to compute the first derivative of the response with respect to the covariate (see equation (5)). For a discrete covariate, margins compute the effect of a discrete change of the covariate (see equation (6)), including seasonal dummies, age of household head, education of household head, regions of country, race, ethnicity, and presence of children. The average marginal probability effects are exhibited in Table 3.5. With increasing own-price of respective nuts, the increasing propensity of households purchases more peanuts and tree nuts varied from 0.025 to 1.163 points.

Household-level choices also revealed different seasonal patterns for various nuts. Household had a higher propensity to purchase peanuts, almonds, and pistachios in the first quarter, and pecans, cashews, walnuts, macadamia nuts, mixed nuts, and overall tree nuts in the last quarter. Meanwhile, as household size goes up by 1%, the propensity of households purchasing peanut and all tree nuts increase by 0.001 (macadamia nuts) to 0.075 (peanuts) points. As household annual income increases by one percent, household in our sample has a higher probability of purchases all nuts products, the higher propensity varied from 0.001 (macadamia nuts) to 0.034 (almonds).

We also calculated the actual marginal effects with respect to household size and household income since they were transformed using logarithms forms. Exhibited in Table 3.5, household size is only significant for peanuts, and income is significant for peanuts, pecans, almonds, and walnuts with small effects given the magnitude of income comparing to the dependent variables as probability.

Most of coefficients associated with age factors are significant, and the results are consistent as the age of household head gets older, the propensity of purchasing peanuts, pecans, cashews, walnuts, pistachios, mixed nuts, and overall tree nuts. Age factor is not a driving force of household-level demand for macadamia nuts. Similar findings with the highest education of household head, as household head has more education, the propensity get larger for all nuts.

Table 3.5 Average Marginal Effects (AMEs) of Price and Socio-Demographics for Peanuts and Tree Nuts

	(1) Peanuts	(2) Pecans	(3) Almonds	(4) Cashews	(5) Walnuts	(6) Macadamia Nuts	(7) Pistachios	(8) Mixed Nuts
up_peanuts	1.249***							
	(0.021)							
up_pecans		0.205***						
		(0.004)						
up_almonds			0.356***					
			(0.007)					
up_cashews				0.434***				
				(0.009)				
up_walnuts					0.269***			
					(0.006)			
up_macadamia						0.027***		
						(0.002)		
up_pistachios							0.218***	
							(0.005)	
up_mixed								0.429***
								(0.008)
q1	0.008***	-0.069***	0.037***	0.002	-0.073***	-0.003***	0.005***	-0.010***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.000)	(0.001)	(0.002)
q2	-0.003	-0.076***	0.020***	-0.004**	-0.085***	-0.002***	-0.015***	-0.025***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.000)	(0.001)	(0.002)
q3	-0.007***	-0.080***	0.000	-0.018***	-0.076***	-0.002***	-0.022***	-0.028***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.000)	(0.002)	(0.002)
lhsize	0.075***	0.023***	0.016***	0.028***	0.022***	0.001***	0.015***	0.032***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.000)	(0.001)	(0.002)
lincome	0.011***	0.017***	0.034***	0.012***	0.016***	0.001***	0.017***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
agehh_under35	-0.145***	-0.046***	-0.019***	-0.083***	-0.082***	-0.001	-0.046***	-0.103***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)
agehh_35to49	-0.079***	-0.036***	-0.004*	-0.048***	-0.064***	-0.000	-0.022***	-0.067***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.002)	(0.002)

Table 3.5 Continued

	(1) Peanuts	(2) Pecans	(3) Almonds	(4) Cashews	(5) Walnuts	(6) Macadamia Nuts	(7) Pistachios	(8) Mixed Nuts
agehh_50to64	-0.029***	-0.013***	0.006***	-0.014***	-0.029***	0.000	-0.004***	-0.021***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)	(0.001)	(0.002)
eduhh_lesshigh	-0.033***	-0.025***	-0.065***	-0.027***	-0.025***	-0.005**	-0.028***	-0.027***
	(0.008)	(0.006)	(0.008)	(0.007)	(0.006)	(0.002)	(0.006)	(0.007)
eduhh_highschool	-0.009***	-0.009***	-0.040***	-0.013***	-0.011***	-0.002***	-0.007***	-0.017***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.002)	(0.002)
eduhh_somecollege	-0.004**	-0.004***	-0.021***	-0.006***	-0.009***	-0.000	-0.001	-0.009***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)	(0.001)	(0.002)
NewEngland	0.034***	-0.004	0.004	-0.004	0.026***	-0.002**	0.018***	-0.001
	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)	(0.001)	(0.003)	(0.004)
MiddleAtlantic	0.010***	-0.027***	-0.006**	-0.004	0.019***	0.001*	0.011***	-0.011***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
ENCentral	0.049***	0.016***	0.005*	0.023***	0.028***	-0.005***	0.012***	0.006**
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
WNCentral	0.043***	0.018***	0.007**	0.014***	-0.003	-0.005***	0.000	0.024***
	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)	(0.003)
SouthAtlantic	0.041***	0.020***	0.009***	0.017***	0.010***	-0.003***	0.008***	0.028***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
ESCentral	0.049***	0.026***	-0.005	0.012***	0.000	-0.002**	-0.005**	0.036***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)
WSCentral	0.027***	0.037***	0.002	0.000	0.006**	-0.004***	-0.006***	0.021***
	(0.004)	(0.002)	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
Mountain	0.010***	0.024***	-0.010***	0.017***	-0.000	-0.001**	0.010***	0.015***
	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)	(0.003)
White	0.004	0.005*	0.011***	-0.010***	0.006*	-0.001**	0.000	-0.001
	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)
Black	-0.005	0.014***	-0.022***	0.001	-0.006	-0.006***	0.005*	-0.002
	(0.005)	(0.003)	(0.004)	(0.004)	(0.004)	(0.001)	(0.003)	(0.004)
Asian	0.001	-0.027***	0.001	-0.002	0.001	-0.000	0.013***	0.012**
	(0.006)	(0.004)	(0.005)	(0.005)	(0.004)	(0.001)	(0.004)	(0.005)

Table 3.5 Continued

	(1) Peanuts	(2) Pecans	(3) Almonds	(4) Cashews	(5) Walnuts	(6) Macadamia Nuts	(7) Pistachios	(8) Mixed Nuts
hispanic_reg	-0.010**	-0.000	0.002	-0.016***	0.001	-0.001	0.008***	-0.002
	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)	(0.003)
no_child	0.054***	0.017***	0.016***	0.029***	0.029***	0.002***	0.007***	0.035***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
hsize	0.033*	0.009	0.004	0.011	0.009	0.000	0.005	0.014
	(0.019)	(0.007)	(0.005)	(0.009)	(0.007)	(0.000)	(0.004)	(0.009)
income	0.000*	0.000**	0.000*	0.000	0.000*	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Bootstrap standard errors in parentheses. * p>0.1 **p>0.05 ***p>0.01
Note: up_nuts = unit value for respective nut category (also used as proxy for price)
Marginal effects; Standard errors in parentheses * p<0.1 ** p<0.05 *** p<0.01

61,380 households are located across United States and revealed different preferences for nuts in this study. Peanuts were preferred by households from central area with a higher propensity. Pecans were purchased with higher probability by households locating in west south-central area. Households from East North Central tend to purchase more almonds who also have higher propensity purchasing cashews. A higher propensity of macadamia nuts purchases is associated with households located in Middle Atlantic area.

As for the racial background of households and children presence, households without children had a larger propensity purchasing nuts as expected. As nuts being consumed substituting unhealthy snacks, households with children might purchases more of potato chips and unhealthy ones to satisfy their children's needs. Race was not a factor driving demand for overall tree nuts products but households with head being white/Caucasian were found to purchase pecans, almonds, and walnuts with a higher possibility. And black households prefer more of pecans and less of almonds and macadamia nuts. Hispanic households had a larger propensity purchasing pistachios and smaller for peanuts and cashews.

3.5.3. Evaluation of Model Prediction

We also provide the prediction evaluations for all probit models, that how accurate of this binary choice model to predict the true probability of households purchasing nuts is evaluated. We categorize the prediction into 4 scenarios in a 2×2 matrix. The four scenarios are, (a) true positive: the predicted probability for outcome $(\mathrm{Z} = 1)$ is larger than cut-off values[11]; (b) false negative: the predicted probability for outcome $(Z = 1)$ is less than than cut-off values; (c) false positive: the predicted probability for outcome $(Z = 0)$ is larger than cut-off values; (d) true negative: the predicted probability for outcome $(Z = 0)$ is smaller than cut-off values. We also calculated the sensitivity, defined as the proportion of observed positive outcomes that were predicted to be positive. Specificity are defined as the proportion of observed negative outcomes that were predicted to be negative. The correctly classified scenarios by our Probit model are (a) and (d) and we calculated the success rate by diving this number by the total number of observations. As shown in Table 6, the success rate (correctly classified) ranged from 48.56% (pistachios) to 94.52% (macadamia nuts); sensitivity varied from 39.94% (macadamia nuts) to 64.85% (pecans); specificity increased from 48.18% (pistachios) to 94.84% (macadamia nuts).

[11] We chose the cut-off values using the market penetration reported in Table 2b for respective nuts. The expectation-prediction for Probit mode deviates with cut-off levels.

Table 3.6 Prediction Evaluations for Probit Models

	Peanuts (cut-off = 0.23)		**Pecans** (cut-off = 0.07)	
	Prob. (Z=1)	Prob. (Z=0)	Prob. (Z=1)	Prob. (Z=0)
Predicted Prob. > cut-off	30,580 (true positive)	83,489 (false negative)	11,758	83,208
Predicted Prob. < cut-off	25,929 (false positive)	105,522 (true negative)	6,372	144,182
Sensitivity	54.12%		64.85%	
Specificity	55.83%		63.41%	
Correctly classified	55.43%		63.51%	
	Almonds (cut-off = 0.13)		**Cashews** (cut-off = 0.12)	
	Prob. (Z=1)	Prob. (Z=0)	Prob. (Z=1)	Prob. (Z=0)
Predicted Prob. > cut-off	17,073	96,862	14,652	95,903
Predicted Prob. < cut-off	14,554	117,031	14,203	120,762
Sensitivity	53.98%		50.78%	
Specificity	54.71%		55.74%	
Correctly classified	54.62%		55.15%	
	Walnuts (cut-off = 0.09)		**Macadamia Nuts** (cut-off = 0.01)	
	Prob. (Z=1)	Prob. (Z=0)	Prob. (Z=1)	Prob. (Z=0)
Predicted Prob. > cut-off	12,240	81,053	556	12,609
Predicted Prob. < cut-off	9,422	142,805	836	231,519
Sensitivity	56.50%		39.94%	

Table 3.6 Continued

Specificity	63.79%		94.84%	
Correctly classified	63.15%		94.52%	
	Pistachios (cut-off = 0.07)		**Mixed Nuts** (cut-off = 0.01)	
	Prob. (Z=1)	Prob. (Z=0)	Prob. (Z=1)	Prob. (Z=0)
Predicted Prob. > cut-off	9,594	117,913	16,454	101,563
Predicted Prob. < cut-off	8,373	109,640	11,245	116,258
Sensitivity	53.40%		59.40%	
Specificity	48.18%		53.37%	
Correctly classified	48.56%		54.05%	

Note: The number of obs is 245,520. The cut-off for respective nuts is based the market penetration data in Table 2. Numbers in each entry is the count of obs that fall into each scenario.

Sensitivity, defined as the proportion of observed positive outcomes that were predicted to be positive, equals to true positive/(true+false positive). Specificity, defined as the proportion of observed negative outcomes that were predicted to be negative, equals to true negative/(true+false negative).

Source: Nielsen Homescan Panel 2015.

3.6. Concluding Remarks

We took advantage of scanner data and binary choice approaches investigating the demand for peanuts and tree nuts products in the United States by identifying demand drivers and profiling households.[12] Based the estimation results and marginal effects of socio-demographic factors, households with older age, higher education level, higher income level, and without children were estimated to purchase peanuts and tree nuts with a larger propensity. The demand for different nuts products was distributed unevenly among different races of households and regions of country.

These findings not only contributed to the literature by adding an up-to-date research regarding the household-level demand of nuts with a more thorough analysis, larger data set, and detailed/diversified nut categories, but also contribute to the industries for nuts food processor and purveyors, including their price strategies, market segmentation, and targeting of households. Simply, these companies could increase their prices and target households with older head, higher educations, more members, larger income level, and without children. In detailed category, the producers of walnuts and cashews is better to spend efforts on targeting households locating in central area while macadamia nuts have a better market in Pacific area.

[12] The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

A set of limitations of our studies rest in the model technique we used and representativeness of our data. We used a binary choice method to capture consumers' choices at household-level. We were able to identify demand drivers and profile existing consumers. Alternative model techniques would be able to investigate the price elasticities and demand interrelationships. In addition, as noted in the data section, our sample in the Nielsen are concentrated on older households, Caucasians, and households without children. Our study also does not consider away-from-home purchases. Despite all these limitations, we believe our study add to the literature by providing a household-level analysis for consumer profiling and demand for nuts.

3.7. References

Bliss, C. I. 1934a. "The method of probits." *Science* 79: 38–39.

—. 1934b. "The method of probits—a correction." *Science* 79: 409–410

Dhaliwal, H. S., 1972. "An Econometric Investigation of Demand Interrelationships Among Tree Nuts and Peanuts." Ph.D. thesis, Oregon State University.

Fraser, G.E.; Sabaté, J.; Beeson, W.L. Strahan, T.M. 1992. "A Possible Protective Effect of Nut Consumption on Risk of Coronary Heart Disease. The Adventist Health Study." *Archives of Internal Medicine* 152, 1416-1424.

Green, R.D. 1999. "Demand for California Agricultural Commodities." *Outlook* 99:11.

He, S., Fletcher, S.M. and Rimal, A., 2005. "Snack peanut consumption: type preference and consumption manners." *Journal of food distribution research*, *36*(856-2016-57439).

Kim, Y. and Dharmasena, S., 2018. “Price Discovery and Integration in US Pecan Markets.” *Journal of Food Distribution Research*, 49(856-2018-3108).

King, J. C., J. Blumberg, L. Ingwersen, M. Jenab, and K. L. Tucker. 2008. “Tree Nuts and Peanuts as Components of a Healthy Diet.” *The Journal of Nutrition* 138(9), 1736S–1740S.

Kris-Etherton, P. M., Hu, F. B., Ros, E., and J. Sabaté. 2008. “The Role of Tree Nuts and Peanuts in the Prevention of Coronary Heart Disease: Multiple Potential Mechanisms.” *The Journal of Nutrition* 138(9), 1746S–1751S.

Lee, I.M., 1950. “Statistical Analysis of Annual Average Returns to Growers of Almonds, 1924/25 to 1948/49. No. 103.” University of California, College of Agriculture.

Lerner, E. B., 1959. “An Econometric Analysis of the Demand for Pecans with Special Reference to the Demand Interrelationships Among Domestic Tree Nuts.” Ph.D. thesis, Oklahoma State University.

Long, J. S. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage.

Long, J. S., and J. Freese. 2014. *Regression Models for Categorical Dependent Variables Using Stata*. 3rd ed. College Station, TX: Stata Press

Mattes, R.D., P.M. Kris-Etherton, and G.D. Foster. 2008. “Impacts of Peanuts and Tree Nuts on Body Weight and Healthy Weight Loss in Adults.” *Journal of Nutrition* 138(9), 1741S-1745S.

[dataset] Nielsen at the Kilts Center for Marketing, Consumer Panel Data 2015, https://www.chicagobooth.edu/research/kilts/datasets/nielsen

Rimal, A.P. and Fletcher, S.M., 2002. "Snack peanuts purchase pattern: effects of nutritional considerations and household characteristics." *Journal of Agricultural and Applied Economics*, *34*(1), pp.51-65.

Ros, E. 2010. "Health Benefits of Nut Consumption." *Nutrients*, 2(7), 652-682.

Russo, C., Green, R., Howitt, R. E., 2008. "Estimation of Supply and Demand Elasticities of California Commodities." Department of Agricultural and Resource Economics, University of California, Davis, Working Paper No. 08-001.

Sabaté, J., 2003. "Nut consumption and body weight." *The American journal of clinical nutrition*, *78*(3), pp.647S-650S.

[software] StataCorp. 2015. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.

Statista 2019. *Food Report 2019 - Snack Food*. Statista Consumer Market Outlook.

U.S. Department of Health and Human Services, 2015. *2015–2020 Dietary Guidelines for Americans.* Washington DC.

Wells, G.J., Miller, S.E. and Thompson, C.S., 1986. "Farm Level Demand for Pecans Reconsidered." *Journal of Agricultural and Applied Economics*, 18(1), pp.157-160.

4. STUDY III: HOUSEHOLD DEMAND FOR PEANUTS AND TREE NUTS IN THE UNITED STATES: A DYNAMIC PANEL TOBIT ANALYSIS

4.1. Introduction

Almonds provide a good source of calcium to fight against osteoporosis; Peanuts are reported to boost memory and reduce blood pressure; Pecans strength bone and keep nerves and muscles healthy; Macadamia nuts are recommended to help nervous system and skin development[13]. With many benefits of nuts being confirmed by nutritionists, consumptions of such products are on the rise due to the fact that nuts have many healthy contents, such as fiber, minerals, and unsaturated fat. With increasing awareness of health concerns, and given many healthy contents that nuts have, they are also recommended in the consumer's daily life. However, little is known about the market and consumers' behaviors.

We aim to examine the demand for nuts products and identify the drivers of such demand. Naturally, only households who made actual purchases can be observed and their information could be documented. However, examining only these individuals will yield bias of results that characteristics of household whose purchases are not observed are left out in any available data set. Alternatively, econometric model, such as Tobit, enables us to address and control for such bias and investigate the demand for peanuts and tree nuts which is a critical component of U.S. agricultural economy. Meanwhile,

[13] Market Research.com: https://www.marketresearch.com/Food-Beverage-c84/Food-c167/Nuts-c1644

most of such products were purchased at the household level, then a household level scanner data suit the needs of this study.

There is a dominant market for nut and seed products in North American while the global market in terms of market value and volume is growing tremendously. According to the United States Department of Agricultural (USDA) Economic Research Service, the annual per capita consumptions of peanuts rose from 5.7 pounds to 7.2 pounds, and tree nuts increased from 1.8 pounds to 4.7 pounds through the period from 1970 to 2016 in the United States. Nuts products are sold as many single varieties and as mixes and mainly being consumed in the form of snacks as alternatives for other products, such as potato chips and beef jerkies.

Nuts market accounts for a significant aspect of agricultural economy not only in the United States but also in the world. It has not been studied very well and has not received enough attention from scholars in the field of agricultural economics. We attempt to fill in this gap by mainly focusing on the demand of peanuts and tree nuts in the United States using household level scanner data and dynamic Tobit model addressing sample selection and unobserved heterogeneity.

In order to achieve the over-arching objective of this study, examining the demand for peanuts and tree nuts in the United States, we used a dynamic panel Tobit model analysis. Such technique enables the ability, (1) to investigate the effects of demographic factors; (2) to obtain own-price elasticities, cross-price elasticities, and income elasticities; (3) to address habitual patterns of nuts purchases with sample selection and unobserved heterogeneity.

The nut categories in this analysis correspond to almonds, cashews, macadamia nuts, mixed nuts, peanuts, pecans, pistachios, walnuts. We also aggregated all of the tree nuts as one category and explored its relationship with peanuts. To address these objectives, a panel data of household purchase and expenditure for the aforementioned nut categories are derived from the Nielsen Homescan Panel for the calendar year from 2009 to 2015.

This study is organized as following. First, we introduced the literature related to the demand aspect and price of nuts products. Second, dynamic unobserved effect Tobit model was introduced and developed. In the following sections, we described the data used and provided the empirical results. In the last section, we discussed our findings with limitations and provided directions for future studies.

4.2. Related Literature

Nuts market is one of the most important segments of the agricultural economy in the United States. However, it has not been studied very well. We summarized the literature related to the demand aspect of nuts products.

Lee (1950), Lerner (1959), Dhaliwal (1972), Well, Miller, and Thompson (1986), and Russo, Green, and Howitt (2008) addressed and investigated the demand interrelationship of various nuts products using annual data through 1924 to 2001 and different techniques, including single equation and seemingly unrelated regression.

Green (1999), Wood (2001), Crespi and Chacon-Cascante (2004), Florkowski and Sarmiento (2005), Kim and Dharmasena (2018) explored specific nuts products in certain areas, such as almonds in California and pecans in Oklahoma, Georgia, and,

Table 4.1 Summary of Literatures Related to Price and Demand for Peanuts and Tree Nuts Products

Topics	Studies	Data	Model	Main Findings
Demand/Demand Interrelationships	Lee (1950)	Annual data 1924-1949	Single equation	Estimated demand elasticities -0.4602 to -5.0251
	Lerner (1959)	Annual data 1970-2001	Seemingly unrelated regression	Own-price elasticities -1.188 to -23.042; Income elasticities 2.3228 to 20.1245; Walnuts were found to be substitutes for pecans and filberts. Pecans were estimated to be compliments for filberts.
	Dhaliwal (1972)	Annual data 1970-2001	Single equation	Own-price ranged from -0.286 to -1.923; Numerous substitute and complementary relationships were revealed.
	Wells, Miller, and Thompson (1986)	Annual data 1970-1982	Price-dependent demand function	The own-price flexibility of pecans at the farm level was estimated to be -0.97.
	Russo, Green, and Howitt (2008)	Annual data 1970-2001	Seemingly unrelated regression and Single equation	The estimated own-price elasticities for almonds ranging from -0.480 to -0.350 and for walnuts varying from -0.266 to -0.284 using single equation.
Regional Demand	Green (1999)	N/A	N/A	The own-price elasticity of demand for U.S. almonds was found to be -0.83.
	Crespi and Chacon-Cascante (2004)	Annual data 1962-1997	New empirical industrial organization	The market power of the Almond Board of California was significantly less than expected of profit-maximizing cartels.
	Florkowski and Sarmiento (2005)	Survey data of 177 observations in 1998	Spatial analysis	Identified the linkages between the price of in-shell pecans and the characteristics of the orchard, production costs, and location
	Kima and Dharmasena (2017)	Seasonal data 2005-2016	Machine-learning and graphical methods	Current pecan prices received by growers in Texas were directly caused by pecan prices received in Oklahoma, Georgia and Louisiana. Past period pecan prices in Georgia were found to influence current prices in other states.

Note: The list of references is provided at the end of the body text.

Louisiana by utilizing new empirical industrial organization, spatial analysis, and machine learning and graphical methods.

To the best of our knowledge, there are no other studies documenting and studying the demand for nuts products. Our study adds to the literature by investigating the effects of demographic factors, capturing habitual patterns of purchases, addressing sample selection, and controlling for unobserved heterogeneity. We summarized the findings and listed them in the Table 4.1.

4.3. Data

The panel data used for this study is derived based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business[14]. We used the data for the calendar years from 2009 to 2015. Serving the purpose of investigating the demand of peanuts and tree nuts products, we divided the nuts products in our data set into nine categories using product module code and UPC codes provided by Nielsen. The nine categories corresponding to this study are (1) peanuts, (2) pecans, (3) almonds, (4) cashews, (5) macadamia nuts, (6) walnuts, (7) pistachios, (8) mixed nuts, (9) tree nuts[15].

In the Nielsen data, for each household, information recoded includes quantities they buy, how much paid, and coupon value. First, we selected the common households that stayed in the

---

[14] Disclaimer: Researcher(s) own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.
[15] We aggregated all the tree nuts products into one category.

panel of Nielsen data from 2009 to 2015[16]. Second, we aggregated the data for each existing household using household code given by Nielsen annually. Then we summed up the total quantity and total expenditure that each household spent on each nut category in these seven calendar years. Then we divided the total expenditure by the total quantity and obtained the unit value, also taken as a proxy for price, for each category. Then we arrived at a panel of 23,811 households for the period from 2009 to 2015, a total of 166,677 observations.

We provide the descriptive statistics for quantity and unit value variable of respective category in Table 4.2. First, we replaced with missing quantities with zeroes. On average, households purchased 59.10 pounds of peanuts and 102.15 pounds of tree nuts products. Among tree nuts, mixed nuts were consumed the most, followed by almonds, pistachios, cashews, walnuts, pecans, and macadamia nuts. Then, we imputed the missing unit values for each category (see Appendix B). Among all the nuts products, peanuts had the lowest unit value, on the average, with eighteen cents per ounce. Among all the tree nuts products, macadamia nuts exhibited the highest unit price, on the average, with about 83 cents per ounce, followed by pecans, pistachios, walnuts, almonds, cashews, and mixed nuts. On average, the unit price of all tree nuts products was around forty cents per ounce.

---

[16] Nielsen updates their sample each year that part of households stays in the sample and other is replaced with new ones. In order to have a balanced panel data, only those households that stay in the sample from year 2009 to 2015 are included in this study.

Table 4.2 Descriptive Statistics of Unconditional Quantities (oz) and Unit Values ($/oz)

Variables	Mean	Std. Dev.	% of outliers[1]
Quantity (ounce sold)			
Peanuts	59.1	192.41	0.38%
Pecans	7.18	27.79	0.51%
Almonds	20.81	80.57	0.37%
Cashews	16.96	52.35	0.75%
Walnuts	12.36	45.44	0.60%
Macadamia Nuts	0.52	7.06	0.31%
Pistachios	11.56	45.29	0.75%
Mixed Nuts	22.37	68.7	0.74%
Tree Nuts	102.15	181.67	0.38%
Unit Value ($/oz)[1]			
Peanuts	0.16	0.08	0.18%
Pecans	0.57	0.14	0.44%
Almonds	0.43	0.15	0.22%
Cashews	0.39	0.11	0.16%
Walnuts	0.45	0.12	0.42%
Macadamia Nuts	0.83	0.15	0.04%
Pistachios	0.45	0.13	0.27%
Mixed Nuts	0.35	0.1	0.32%
Tree Nuts	0.41	0.19	0.20%

Note: We replaced the unobserved purchases with zeroes. We imputed the missing prices variables and please see the Appendix B for details. We do have zero price paid since coupon values were taken into consideration.

[1] We count the number of observations that is larger than mean+5*standard deviation as outliers. (one tail variable screening). Then divide this number by the number of total obs, 166,677 to arrive the %.

Note: The number of obs is 23,811 households × 7 periods (year 2009 to 2015). Total is 166,677.

Source: Nielsen Homescan Panel 2009-2015, calculations by the author.

There are 23,811 unique households in the Nielsen data for the years from 2009 to 2015. Not every household purchased nuts products at each time point, meaning our data is censored, hence might give rise to sample selection bias. We also reported the descriptive statistics for unconditional quantities and unit values that missing purchases are replaced with zeroes in Table 4.2.

Serving the purpose of investigating the effects of household's characteristics on demand, varieties of demographic information were provided in Table 4.3, including household size, household annual income, age of household head, education level of household head, regions of country, ethnicity, and presence of children under the age of 18. The average household size is 2.16. The majority of households are associated with older households head where 79% of them is older than 50 years old. Geographically, the panel is concentrated in East North Central and South Atlantic regions[17].

In Table 4.3, we also made suggestions about the hypotheses of these demographic factors. Income and household size are expected to be positively correlated with nuts purchases that with increasing income level and more members, households would increase their purchases. Education is also assumed to have a positive relationship with nuts purchases. Households with more education are more likely to take in healthy information and change their lifestyle by having more purchases. Households with different ages of head, locating in different regions, and different racial backgrounds might yield inconsistent patterns of purchases, that signs could be negative or positive.

---

[17] Nielsen categorizes the regions of country based on sources provided by U.S. Census Bureau. See Table 4.4.

Table 4.3 Descriptive Statistics of Household Size, Household Income, and Socio-Demographic Characteristics

Variable	Label	Mean	Std. Dev.
(+) Household Composition	*hsize*	2.16	1.15
(+) Annual Income	*lincome*	58,081	28,724
(−) Under 35	*agehh_under35*	0.02	0.14
(−) 35 to 49	*agehh_35to49*	0.2	0.4
(+) 50 to 64	*agehh_50to64*	0.5	0.5
(+) 65 and above*	*agehh_65andabove*	0.29	0.45
High school and less	*eduhh_hsandless*	0.2	0.4
At least some college*	*eduhh_atleastcollege*	0.8	0.4
(±) New England	*NewEngland*	0.05	0.21
(±) Middle Atlantic	*MiddleAtlantic*	0.13	0.34
(±) East North Central	*ENCentral*	0.19	0.39
(±) West North Central	*WNCentral*	0.09	0.29
(±) South Atlantic	*SouthAtlantic*	0.19	0.4
(±) East South Central	*ESCentral*	0.06	0.23
(±) West South Central	*WSCentral*	0.1	0.3
(±) Mountain	*Mountain*	0.07	0.26
(±) Pacific area*	*Pacific*	0.12	0.33
(±) White	*White*	0.84	0.36
(±) Black	*Black*	0.09	0.29
(±) Asian	*Asian*	0.03	0.16
(±) Other races*	*Other*	0.04	0.19
(±) Hispanic origin	*hispanic_reg*	0.04	0.2
(+) No children	*no_child*	0.86	0.35

Note: Min and max are not reported for dummy variables. *The base group that were excluded from analysis. The positive/negative signs indicate the expectation of these variables. For instance, households are expected to have a higher consumption of peanuts and tree nuts as their ages go up and income level increases. Source: Nielsen Homescan Panel 2009-2015, calculations by the author.

As shown in Table 4.3, on the average, annual income of households in the Nielsen data is around fifty-eight thousand dollars. Market penetration is defined as the percentage of households who had actual purchases to the total number of households, exhibited in Table 4.4. We calculated the income level for each household using the income categories provided by Nielsen, see Table 4.5.

Besides annual income ($) and household size, all of other socio-demographics were transformed into dummy variables used in the estimation process. Consequently, the mean is interpreted as the percentage of the panel in each group. For example, more than 80% households have education level at least some college. More than four fifth of households are white/Caucasian. Based on the information of presence of children, 86% households do not have any children under the age of 18.

In the following section, we explained the dynamic Tobit model we implemented to examine the demand of nuts products as well as to address unobserved heterogeneity and data-censoring problem.

Table 4.4 Annual Market Penetration by Nuts Category

	Peanuts	**Pecans**	**Almonds**	**Cashews**	**Macadamia Nuts**	**Pistachios**	**Walnuts**	**Mixed Nuts**	**Tree Nuts**
2009	52%	27%	39%	36%	4%	16%	33%	35%	82%
2010	52%	25%	39%	35%	3%	18%	31%	34%	81%
2011	52%	24%	41%	32%	2%	25%	30%	33%	81%
2012	50%	23%	41%	29%	2%	28%	29%	33%	81%
2013	50%	26%	38%	32%	2%	25%	29%	33%	81%
2014	52%	24%	36%	32%	2%	22%	27%	32%	80%
2015	51%	24%	32%	31%	2%	20%	27%	31%	80%

Note: The number of households across six year is 23,811.

Source: Nielsen Homescan Panel 2009-2015, calculations by the author.

Table 4.5 U.S. Census Bureau Regions and Divisions

Regions	States
New England	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont
Middle Atlantic	New Jersey, New York, and Pennsylvania
East North Central	Indiana, Illinois, Michigan, Ohio, and Wisconsin
West North Central	Iowa, Nebraska, Kansas, North Dakota, Minnesota, South Dakota, and Missouri
South Atlantic	Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, and West Virginia
East South Central	Alabama, Kentucky, Mississippi, and Tennessee
West South Central	Arkansas, Louisiana, Oklahoma, and Texas
Mountain	Arizona, Montana, Colorado, Utah, Idaho, Nevada, New Mexico, and Wyoming
Pacific area	Alaska, California, Hawaii, Oregon, and Washington

Source: U.S. Census Bureau.

Table 4.6 Income Levels in Nielsen Data

	Household Annual Income reported in Nielsen
Levels	$5,000
	$5,000-7,999
	$8,000-8,999
	$10,000-11,999
	$12,000-14,999
	$15,000-19,999
	$20,000-24,999
	$25,000-29,999
	$30,000-34,999
	$35,000-39,999
	$40,000-44,999
	$45,000-49,999
	$50,000-59,999
	$60,000-69,999
	$70,000-99,999
	above $100,000

Note: Sixteen levels of income choices were given in the survey and household chose the category fits them. We used the median as the income level for each household.

Source: Nielsen Homescan Panel 2009-2015; calculations by the author.

4.4. Methodology and Estimation

4.4.1. Dynamic Unobserved Effects Tobit Model

We utilize a panel version of Tobit model to investigate the demand for peanuts and tree nuts using Nielsen Homescan Panel data. The unobserved effects Tobit model introduced in Wooldridge (2010) based on Amemiya (1984) is used, and the model is expressed as follows,

$$(1)\quad q_{iht} = \max(0, \boldsymbol{X}'_{ht}\boldsymbol{\beta_i} + \delta_i q_{ih,t-1} + c_{ih} + \varepsilon_{iht}),$$

$$i = 1,2,\dots,9 \quad h = 1,2,\dots,23811 \quad \mathrm{t} = 1,2,\dots,7$$

where q_{iht} is the quantity made by household h for nut category i in year t, $q_{ih,t-1}$ is the habitual purchase at the lag one period, c_{ih} is the time-invariant component in the random-effect model, known as the unobserved effect. ε_{iht} assumed to be normally distributed with zero mean and variance (σ_ε^2) conditional on $\boldsymbol{X}_{ht}$ and c_{ih}.

In the vector of independent variables $\boldsymbol{X}_{ht}$ consists of unit price of respective nuts products, household size, annual household income, age of household head, education level of household head, region, ethnicity, presence of children under the age of 18 for household h in the panel year of t of Nielsen Homescan Panel 2009-2015. The aggregated tree nuts category was estimated differently by only including the unit price of tree nuts and peanuts and demographics factors.

According to Wooldridge (2010), the model in equation (1) is suitable only for corner solution applications, and owing to the censored data in our sample, a dynamic linear model is as follows,

$$(2)\quad q^*_{iht} = \boldsymbol{X}'_{ht}\boldsymbol{\beta_i} + \delta_i q_{ih,t-1} + c_{ih} + \varepsilon_{iht}$$

where q^*_{iht} is the non-zero purchases of household observed. In order to control for heterogeneity and tackle data-censoring problem, we followed the methods by Wooldridge (2010). First, we allow the possible correlation between $\boldsymbol{X}_{ht}$ and c_{ih}. And the latter component has a distribution as

(3) $c_{ih} \mid \boldsymbol{X}_h \sim \text{N}\,(\gamma + q_{ih1}\gamma_0 + \overline{\boldsymbol{X}}_h\gamma_i, \sigma_a^2)$

which also takes a linear form, $c_{ih} = \gamma + q_{ih1}\gamma_0 + \overline{\boldsymbol{X}}_h\gamma_i + a_h$, and σ_a^2 is the variance of a_h following a normal distribution as N $(0, \sigma_a^2)$. This is the method explained in Wooldridge (2010) to control for heterogeneity and tackle initial value problem that the unobserved effect, c_{ih} is given the initial value, q_{io}, and the exogenous variables in all time periods, $\boldsymbol{X}_h$.

Subsequently, we have our dynamic unobserved effects Tobit model, that takes a form as

(4) $q_{iht} = \max(0, \gamma_i + \boldsymbol{X}'_{ht}\boldsymbol{\beta_i} + \delta_i q_{ih,t-1} + q_{ih1}\gamma_0 + \overline{\boldsymbol{X}}_h\gamma_i)$

and the full model takes the form as follows,

$$(5)\quad \ln q_{iht} = \max\Bigg(0, \gamma_i + \sum_{j=1}^{8} \beta_j \ln P_{hjt} + \beta_{1i} \ln HZ_{ht} + \beta_{2i}\, ln\, INC_{ht} + \beta_{3i} NE_{ht}$$

$$+ \beta_{4i} MA_{ht} + \beta_{5i} ENC_{ht} + \beta_{6i} WNC_{ht} + \beta_{7i} SA_{ht} + \beta_{8i} ESC_{ht}$$

$$+ \beta_{9i} WSC_{ht} + \beta_{10i} Mou_{ht} + \beta_{11i} WH_h + \beta_{12i} BL_h + \beta_{13i} AS_h + \beta_{14i} HR_h$$

$$+ \beta_{15i} NC_{ht} + \beta_{16i} Age_{1ht} + \beta_{17i} Age_{2ht} + \beta_{18i} Age_{3ht} + \beta_{19i} Edu_{ht}$$

$$+ \delta_i \ln q_{ih,t-1} + \gamma_{0i} \ln q_{ih1} + \sum_{j=1}^{8} \omega_j \overline{P_{hj}} + \beta_{20i} \ln \overline{HZ_h} + \beta_{21i}\, ln\, \overline{INC_h}$$

$$+ \beta_{22i} \overline{NE_h} + \beta_{23i} \overline{MA_h} + \beta_{24i} \overline{ENC_h} + \beta_{25i} \overline{WNC_h} + \beta_{26i} \overline{SA_h}$$

$$+ \beta_{27i} \overline{ESC_h} + \beta_{28i} \overline{WSC_h} + \beta_{29i} \overline{Mou_h} + \beta_{30i} \overline{NC_h} + \beta_{31i} \overline{Age_{1h}}$$

$$+ \beta_{32i} \overline{Age_{2h}} + \beta_{33i} \overline{Age_{3h}} + \beta_{34i} \overline{Edu_h} + \varepsilon_{iht}\Bigg),$$

where

i – Peanuts, pecans, almonds, cashews, walnuts, macadamia nuts, pistachios, and mixed nuts;

h – Total 23,811 households;

t – Calendar year of 2009 to 2015;

γ_i – Constant terms for product i

P_{hjt} – Unit price paid by household h for product j at time period t;

HZ_{ht} – Household size for household h at time period t;

INC_{ht} – Household h annual income at time period t;

NE_{ht} – Household h locating in New England at time period t;

MA_{ht} – Household h locating in Middle Atlantic at time period t;

ENC_{ht} – Household h locating in East North Central at time period t;

WNC_{ht}	–	Household h locating in West North Central at time period t;
SA_{ht}	–	Household h locating in South Atlantic at time period t;
ESC_{ht}	–	Household h locating in East South Central at time period t;
WSC_{ht}	–	Household h locating in West South Central at time period t;
Mou_{ht}	–	Household h locating in Mountain at time period t;
WH_{h}	–	Household h being white/Caucasian;
BL_{h}	–	Household h being Black;
AS_{h}	–	Household h being Asian;
HR_{h}	–	Household h being Hispanic origins;
NC_{ht}	–	Household h without children under the age of 18 at time period t;
Age_{1ht}	–	Age of household h under the age of 35 at time period t;
Age_{2ht}	–	Age of household h between the age of 35 and 49 at time period t;
Age_{3ht}	–	Age of household h between the age of 50 and 64 at time period t;
Edu_{ht}	–	Education of household h high school at time period t;
$\overline{*_h}$	–	Time-average for respective variable by group of household h;
$\ln *$	–	Logarithmic function
ε_{hit}	–	Disturbance terms.

As the coefficients of Tobit model cannot be explained directly, we compute the marginal effects following McDonald and Moffitt (1980). The unconditional marginal effects using all observations take form as

$$(6) \quad \text{Unconditional Marginal Effect } of\ x_j = \beta_j \Phi(\frac{\widehat{q_{iht}}}{\sigma})$$

where $\widehat{q_{iht}}$ is the linear prediction from Tobit model (equation (5)), σ corresponds to the variance of error terms, $\Phi(\cdot)$ is the cumulative density function of standard normal distribution, and β_j is the estimated coefficient for respective explanatory variable. The conditional marginal effects using only non-zero observations take form as

(7) Conditional Marginal Effect $of\ x_j$

$$= \beta_j \left(1 - \frac{\widehat{q_{iht}}}{\sigma} \frac{\phi\left(\frac{\widehat{q_{iht}}}{\sigma}\right)}{\Phi\left(\frac{\widehat{q_{iht}}}{\sigma}\right)} - \left(\frac{\phi\left(\frac{\widehat{q_{iht}}}{\sigma}\right)}{\Phi\left(\frac{\widehat{q_{iht}}}{\sigma}\right)} \right) \left(\frac{\phi\left(\frac{\widehat{q_{iht}}}{\sigma}\right)}{\Phi\left(\frac{\widehat{q_{iht}}}{\sigma}\right)} \right) \right)$$

where $\phi(\cdot)$ indicates the standard normal probability density function.

The dynamic unobserved Tobit model is used to examine the household demand for peanuts and tree nuts products in this study as well as to control for heterogeneity and sample selection bias. In the following sections, we provide the estimation results and discussion as to how we interpret these empirical findings.

4.5. Empirical Results

The dynamic unobserved effects Tobit model was estimated in StataMP Version 14 using the command XTTOBIT. A log-log functional form was exploited in this study. First, we take the logarithm form on dependent variables[18] (q_{iht}), all price variables[19] for respective nuts category, household size, and household annual income. Second, the household size and household annual income were transformed using logarithm form as well. Owing to the fact this model is basically a random-effects, it is calculated using

[18] Due to the fact that we replaced unobserved purchases with zeroes, taking logarithm form will automatically drop these observations, please see Table 3 as refer to market penetration. In order to preserve these observations, we add the value of one on each observation, then take log of all of them, in which missing quantities now is transformed to be zero in the logarithm form.
[19] Owing to the method we derive quantity and unit value of nuts products from Nielsen dataset, we need to impute the missing unit values, that used as explanatory variables. Please refer to Appendix B as to how we impute and fill in missing unit prices.

Gauss-Hermite quadrature in Stata. Since nine detailed nuts products are included in this study, we conducted the Tobit model on each individual category separately. In order to tackle data-censoring problem, we specify the lower limit of quantity being zero in XTTOBIT command.

4.5.1. Parameter Estimates, Model Fitness, and Joint Tests

We provided the estimation results in Table 4.7, including parameter estimates, model fitness, and associated p-values. Owing to the non-linear estimation procedure, we calculated a proxy for model fitness, R^2. Post estimation, we obtained predicted values of dependent variable and calculated the correlation between it and its actual value. Then the square of this correlation was used as the R^2 for each equation, varying from 0.090 to 0.424. We also reported the variance components in the random-effects model as Sigma_u and Sigma_e for cross-sectional variation and overall variation components.

Exhibited in Table 4.8, joint tests were conducted for various socio-demographic factors that have more than two groups, including age of household head, ethnicity, and regions, indicative of significant difference between these factors and their base group jointly. Age of household head is a driving force of demand for almonds, macadamia nuts, and overall tree nuts. Ethnicity was jointly significant for pecans, almonds, walnuts, macadamia nuts, and pistachios. There are significant differences among regions of the demand for walnuts.

Table 4.7 Parameter Estimates and Standard Errors of Dynamic Unobserved Effects Tobit Model for Respective Nuts Category

	(1) Peanuts	(2) Pecans	(3) Almonds	(4) Cashews	(5) Walnuts	(6) Macadamia Nuts	(7) Pistachios	(8) Mixed Nuts	(9) Tree Nuts
lp_treenuts									-0.657***
									(0.015)
lp_peanuts	-0.492***	0.037	0.076**	0.106**	0.023	-0.065	0.055	0.066	0.090***
	(0.028)	(0.042)	(0.035)	(0.042)	(0.040)	(0.139)	(0.052)	(0.042)	(0.017)
lp_pecans	-0.109**	-0.885***	0.078	-0.295***	-0.232***	-0.204	0.501***	-0.056	
	(0.047)	(0.049)	(0.051)	(0.062)	(0.057)	(0.197)	(0.078)	(0.062)	
lp_almonds	0.024	0.021	-0.969***	0.060	-0.093**	-0.011	-0.586***	0.055	
	(0.032)	(0.043)	(0.032)	(0.043)	(0.041)	(0.145)	(0.053)	(0.044)	
lp_cashews	0.048	0.127**	0.032	-0.432***	-0.029	-0.082	0.322***	-0.143**	
	(0.045)	(0.061)	(0.051)	(0.051)	(0.058)	(0.192)	(0.077)	(0.061)	
lp_walnuts	0.016	0.057	-0.181***	0.022	-0.569***	0.245	-0.351***	0.055	
	(0.047)	(0.060)	(0.052)	(0.063)	(0.048)	(0.197)	(0.078)	(0.063)	
lp_macadamia	-0.123	0.054	0.180*	-0.700***	-0.207**	-1.472***	1.657***	-0.352***	
	(0.085)	(0.109)	(0.094)	(0.110)	(0.103)	(0.232)	(0.138)	(0.113)	
lp_pistachios	0.149***	0.069	-0.302***	0.308***	-0.044	0.280	-0.660***	0.024	
	(0.044)	(0.057)	(0.048)	(0.059)	(0.054)	(0.184)	(0.058)	(0.058)	
lp_mixed	0.001	0.071	0.062	-0.018	0.074	0.568***	-0.026	-0.422***	
	(0.040)	(0.054)	(0.045)	(0.053)	(0.052)	(0.184)	(0.067)	(0.047)	
lhsize	0.136***	0.184***	0.164***	0.066	0.006	0.307	0.166*	0.136*	0.060**
	(0.051)	(0.070)	(0.059)	(0.069)	(0.066)	(0.236)	(0.087)	(0.071)	(0.028)
lhincome	0.086***	0.289***	0.323***	0.123***	0.192***	0.054	0.248***	0.240***	0.184***
	(0.020)	(0.028)	(0.022)	(0.025)	(0.026)	(0.077)	(0.033)	(0.026)	(0.011)
NewEngland	0.393	-0.256	-0.389	0.561	-1.182**	2.204	0.626	0.728	0.016
	(0.433)	(0.588)	(0.482)	(0.568)	(0.554)	(2.056)	(0.690)	(0.595)	(0.241)

Table 4.7 Continued

	(1) Peanuts	(2) Pecans	(3) Almonds	(4) Cashews	(5) Walnuts	(6) Macadamia Nuts	(7) Pistachios	(8) Mixed Nuts	(9) Tree Nuts
MiddleAtlantic	0.123	-0.558	-0.601	0.221	-1.404***	-0.073	-0.472	0.478	-0.438**
	(0.350)	(0.469)	(0.379)	(0.455)	(0.438)	(1.594)	(0.592)	(0.478)	(0.191)
ENCentral	0.242	0.166	-0.112	0.515	-0.477	0.072	0.239	0.594	-0.103
	(0.306)	(0.406)	(0.336)	(0.407)	(0.392)	(1.315)	(0.521)	(0.421)	(0.168)
WNCentral	0.425	0.332	-0.694*	0.214	-1.200***	-0.046	-1.002*	0.564	-0.283
	(0.357)	(0.467)	(0.391)	(0.458)	(0.455)	(1.543)	(0.595)	(0.476)	(0.192)
SouthAtlantic	0.057	0.374	0.068	0.599	-0.551	0.610	-0.375	0.827**	-0.034
	(0.281)	(0.371)	(0.304)	(0.369)	(0.362)	(1.107)	(0.467)	(0.382)	(0.153)
ESCentral	-0.235	-0.245	0.158	1.082**	-0.371	3.835**	-0.078	0.081	-0.074
	(0.380)	(0.501)	(0.417)	(0.511)	(0.499)	(1.637)	(0.650)	(0.524)	(0.211)
WSCentral	0.494	0.216	-0.136	0.621	-0.536	0.866	-0.059	0.466	-0.015
	(0.328)	(0.413)	(0.349)	(0.421)	(0.404)	(1.356)	(0.554)	(0.436)	(0.174)
Mountain	0.424	0.219	-0.164	0.811**	-0.510	-1.245	-0.352	0.632	0.040
	(0.282)	(0.375)	(0.305)	(0.375)	(0.359)	(1.069)	(0.470)	(0.385)	(0.153)
L.lnoz_peanuts	0.333***								
	(0.006)								
L.lnoz_pecans		0.244***							
		(0.010)							
L.lnoz_almonds			0.418***						
			(0.008)						
L.lnoz_cashews				0.433***					
				(0.009)					
L.lnoz_walnuts					0.276***				
					(0.009)				
L.lnoz_macadamia						0.752***			
						(0.051)			
L.lnoz_pistachios							0.511***		
							(0.011)		

Table 4.7 Continued

	(1) Peanuts	(2) Pecans	(3) Almonds	(4) Cashews	(5) Walnuts	(6) Macadamia Nuts	(7) Pistachios	(8) Mixed Nuts	(9) Tree Nuts
L.lnoz_mixed								0.429***	
								(0.009)	
L.lnoz_treenuts									0.300***
									(0.004)
White	-0.037	0.084	-0.004	-0.050	0.021	-0.215	-0.109	0.024	0.013
	(0.066)	(0.089)	(0.071)	(0.084)	(0.086)	(0.239)	(0.105)	(0.087)	(0.036)
Black	-0.023	0.132	-0.410***	0.028	-0.165	-0.980***	-0.145	0.041	-0.029
	(0.078)	(0.104)	(0.084)	(0.098)	(0.102)	(0.293)	(0.124)	(0.102)	(0.043)
Asian	-0.102	-0.669***	-0.109	-0.071	-0.267**	0.514	0.412***	0.106	0.046
	(0.100)	(0.142)	(0.105)	(0.126)	(0.133)	(0.327)	(0.154)	(0.130)	(0.055)
hispanic_reg	-0.056	-0.044	0.048	-0.299***	0.022	-0.143	0.045	0.071	-0.017
	(0.066)	(0.089)	(0.069)	(0.084)	(0.086)	(0.234)	(0.103)	(0.085)	(0.036)
no_child	-0.064	0.096	-0.032	-0.005	-0.030	0.127	-0.054	0.062	0.009
	(0.056)	(0.077)	(0.064)	(0.076)	(0.074)	(0.259)	(0.093)	(0.079)	(0.031)
agehh_under35	-0.203	0.132	0.298**	-0.064	0.293	0.912	0.273	-0.152	-0.228***
	(0.133)	(0.187)	(0.149)	(0.187)	(0.182)	(0.587)	(0.226)	(0.194)	(0.072)
agehh_35to49	-0.107	0.044	0.140*	-0.150*	-0.044	0.568*	0.271**	-0.055	-0.147***
	(0.065)	(0.089)	(0.074)	(0.088)	(0.085)	(0.301)	(0.110)	(0.090)	(0.035)
agehh_50to64	-0.050	0.040	0.160***	-0.105*	0.028	0.562***	0.127*	-0.056	-0.059**
	(0.042)	(0.057)	(0.049)	(0.057)	(0.054)	(0.199)	(0.073)	(0.058)	(0.023)
eduhh_hsandless	-0.014	0.151**	0.004	0.042	0.062	-0.105	-0.095	0.070	0.051*
	(0.055)	(0.075)	(0.065)	(0.074)	(0.070)	(0.266)	(0.095)	(0.076)	(0.030)
Constant	-4.672***	-5.580***	-5.480***	-4.420***	-4.824***	-4.728***	-6.826***	-6.151***	-1.668***
	(0.341)	(0.460)	(0.368)	(0.422)	(0.444)	(1.157)	(0.540)	(0.436)	(0.148)
sigma_u	1.660***	2.077***	1.661***	1.889***	2.098***	3.080***	2.386***	1.958***	0.933***
	(0.016)	(0.023)	(0.018)	(0.022)	(0.022)	(0.088)	(0.030)	(0.023)	(0.009)

Table 4.7 Continued

	(1) Peanuts	(2) Pecans	(3) Almonds	(4) Cashews	(5) Walnuts	(6) Macadamia Nuts	(7) Pistachios	(8) Mixed Nuts	(9) Tree Nuts
sigma_e	2.512***	2.700***	2.622***	2.977***	2.708***	3.931***	3.348***	3.062***	1.557***
	(0.008)	(0.013)	(0.010)	(0.013)	(0.012)	(0.065)	(0.017)	(0.013)	(0.004)
R^2	0.352	0.228	0.311	0.256	0.257	0.090	0.214	0.284	0.424

Standard errors in parentheses
* p<0.1 ** p<0.5 *** p<0.01

Note: lp_nuts is the log of the unit price of each category, lnoz_nuts is the log of the quantity, L.is the lag one operator in Stata.

Source: Nielsen Homescan Panel 2009-2015, estimated in Stata using XTTOBIT command; calculations by the author.

Table 4.8 Joint Tests Associated with Socio-Demographic Factors

	Age of Household Head		Ethnicity		Regions of Country	
Joint Tests	Chi-sq stats	p-value	Chi-sq stats	p-value	Chi-sq stats	p-value
Peanuts	3.38	0.34	1.14	0.77	7.41	0.49
Pecans	0.77	0.86	**43.41**	0.00	9.98	0.27
Almonds	**12.53**	0.01	**64.04**	0.00	10.62	0.22
Cashews	4.08	0.25	2.18	0.54	8.81	0.36
Walnuts	5.74	0.13	**15.09**	0.00	**16.96**	0.03
Macadamia Nuts	**8.42**	0.04	**28.53**	0.00	11.22	0.19
Pistachios	6.02	0.11	**19.65**	0.00	9.43	0.31
Mixed Nuts	1.26	0.74	0.82	0.85	7.35	0.50
Tree Nuts	**20.04**	0.00	3.37	0.34	12.01	0.15

Note: The joint tests are used to test the significant difference between socio-demographic factors and their base.

The null hypothesis are as follows,

Coefficients associated with NewEngland = MiddleAtlantic = ENCentral = WNCentral = SouthAtlantic = ESCentral = WSCentral = Mountain = 0;

Coefficients associated with White = Black = Asian = 0;

Coefficients associated with agehh_under35 = agehh_35to49 = agehh_50to64 = 0.

Source: Nielsen Homescan Panel 2009-2015, evidences from Table 3, joint tested conducted in Stata 14 by the author.

4.5.2. Marginal Effects

The first eight nut categories were estimated similarly in that all eight price variables are included. The last category, tree nuts was estimated separately by only including the price of its own and peanuts, owing to the fact we aggregated all tree nuts products together as one category. We calculated the unconditional marginal effects using all observations (see equation (6)) and conditional marginal effects using non-zero purchases (see equation (7)), as shown in Table 4.9 and 4.10. We limit our explanation based solely on conditional marginal effects using the mean.

Since a log-log functional form was used, the marginal effects of price can be interpreted as elasticities. The own-price elasticities were estimated ranging from -0.126 (cashews and mixed nuts) to -0.526 (overall tree nuts). There are many substitute and complementary interrelationships revealed in the results. Significantly, peanuts were found to be substitutes for pecans, almonds, cashews, walnuts, pistachios, mixed nuts and overall tree nuts. Pecans were found to be substitutes for almonds and pistachios and complements for peanuts, cashews, and walnuts, macadamia nuts, and mixed nuts. Almonds were found to be compliments for walnuts, macadamia nuts, and pistachios. Cashews were found to be substitutes for peanuts, pecans, almonds, and pistachios and compliments for walnuts, macadamia nuts, and mixed nuts. Macadamia nuts were found to be substitutes for pecans, almonds, and pistachios and compliments for peanuts, cashews, walnuts, and mixed nuts. Walnuts were found to be compliments for almonds and pistachios. Pistachios were found to be substitutes for peanuts, pecans, macadamia nuts, mixed nuts, and cashews and compliments for almonds and walnuts. Mixed nuts were found to be compliments for cashews and pistachios. Household annual income was also transformed using logarithm forms, correspondingly, the income elasticities varied from 0.036 to 0.147, indicative of necessities. The estimate for macadamia nuts was not significant (see Table 4.7).

For all nuts products, households increase their consumptions as household size gets bigger. Most of coefficients associated with regions of country are not statistically

Table 4.9 Unconditional (Mean) Marginal Effects of Price, Income, and Selected Socio-Demographics

	Peanuts	Pecans	Almonds	Cashews	Walnuts	Macadamia Nuts	Pistachios	Mixed Nuts	Tree Nuts
lp_treenuts									-0.600
lp_peanuts	-0.285	0.010	0.031	0.036	0.007	-0.002	0.013	0.023	0.082
lp_pecans	-0.063	-0.227	0.032	-0.100	-0.072	-0.005	0.119	-0.020	
lp_almonds	0.014	0.005	-0.397	0.021	-0.029	0.000	-0.139	0.019	
lp_cashews	0.028	0.033	0.013	-0.147	-0.009	-0.002	0.076	-0.050	
lp_walnuts	0.009	0.015	-0.074	0.008	-0.176	0.006	-0.083	0.019	
lp_macadamia	-0.071	0.014	0.074	-0.239	-0.064	-0.034	0.393	-0.124	
lp_pistachios	0.086	0.018	-0.124	0.105	-0.014	0.006	-0.157	0.008	
lp_mixed	0.001	0.018	0.025	-0.006	0.023	0.013	-0.006	-0.148	
lhsize	0.079	0.047	0.067	0.023	0.002	0.007	0.039	0.048	0.054
lhincome	0.050	0.074	0.132	0.042	0.059	0.001	0.059	0.084	0.168
NewEngland	0.228	-0.065	-0.159	0.191	-0.365	0.051	0.149	0.256	0.014
MiddleAtlantic	0.071	-0.143	-0.246	0.075	-0.433	-0.002	-0.112	0.168	-0.400
ENCentral	0.140	0.042	-0.046	0.175	-0.147	0.002	0.057	0.209	-0.094
WNCentral	0.246	0.085	-0.284	0.073	-0.370	-0.001	-0.238	0.199	-0.259
SouthAtlantic	0.033	0.096	0.028	0.204	-0.170	0.014	-0.089	0.291	-0.031
ESCentral	-0.136	-0.063	0.065	0.369	-0.115	0.089	-0.018	0.029	-0.068
WSCentral	0.286	0.055	-0.056	0.212	-0.165	0.020	-0.014	0.164	-0.013
Mountain	0.246	0.056	-0.067	0.277	-0.157	-0.029	-0.083	0.222	0.037
White	-0.021	0.022	-0.002	-0.017	0.006	-0.005	-0.026	0.008	0.012
Black	-0.013	0.034	-0.168	0.009	-0.051	-0.023	-0.034	0.014	-0.026
Asian	-0.059	-0.171	-0.045	-0.024	-0.082	0.012	0.098	0.037	0.042
hispanic_reg	-0.032	-0.011	0.020	-0.102	0.007	-0.003	0.011	0.025	-0.016
no_child	-0.037	0.025	-0.013	-0.002	-0.009	0.003	-0.013	0.022	0.008
agehh_under35	-0.117	0.034	0.122	-0.022	0.090	0.021	0.065	-0.054	-0.208
agehh_35to49	-0.062	0.011	0.057	-0.051	-0.014	0.013	0.064	-0.019	-0.134
agehh_50to64	-0.029	0.010	0.066	-0.036	0.009	0.013	0.030	-0.020	-0.054
eduhh_hsandless	-0.008	0.039	0.002	0.014	0.019	-0.002	-0.023	0.025	0.047

Note: We calculated marginal effects following McDonald and Moffit (1980) and reported the mean.

Table 4.10 Conditional (Mean) Marginal Effects of Price, Income, and Selected Socio-Demographics

	Peanuts	Pecans	Almonds	Cashews	Walnuts	Macadamia Nuts	Pistachios	Mixed Nuts	Tree Nuts
lp_treenuts									-0.526
lp_peanuts	-0.217	0.009	0.025	0.031	0.006	-0.007	0.013	0.020	0.072
lp_pecans	-0.048	-0.219	0.026	-0.086	-0.064	-0.023	0.119	-0.017	
lp_almonds	0.010	0.005	-0.321	0.018	-0.026	-0.001	-0.139	0.016	
lp_cashews	0.021	0.031	0.010	-0.126	-0.008	-0.009	0.076	-0.043	
lp_walnuts	0.007	0.014	-0.060	0.007	-0.157	0.027	-0.083	0.016	
lp_macadamia	-0.054	0.013	0.060	-0.204	-0.057	-0.162	0.394	-0.105	
lp_pistachios	0.066	0.017	-0.100	0.090	-0.012	0.031	-0.157	0.007	
lp_mixed	0.000	0.018	0.020	-0.005	0.020	0.063	-0.006	-0.126	
lhsize	0.060	0.046	0.054	0.019	0.002	0.034	0.039	0.041	0.048
lhincome	0.038	0.072	0.107	0.036	0.053	0.006	0.059	0.072	0.147
NewEngland	0.173	-0.063	-0.129	0.163	-0.325	0.243	0.149	0.217	0.012
MiddleAtlantic	0.054	-0.138	-0.199	0.064	-0.386	-0.008	-0.112	0.143	-0.350
ENCentral	0.107	0.041	-0.037	0.150	-0.131	0.008	0.057	0.177	-0.083
WNCentral	0.187	0.082	-0.230	0.062	-0.330	-0.005	-0.238	0.168	-0.227
SouthAtlantic	0.025	0.093	0.023	0.174	-0.152	0.067	-0.089	0.247	-0.027
ESCentral	-0.104	-0.061	0.052	0.315	-0.102	0.423	-0.018	0.024	-0.059
WSCentral	0.218	0.054	-0.045	0.181	-0.148	0.096	-0.014	0.139	-0.012
Mountain	0.187	0.054	-0.054	0.236	-0.140	-0.137	-0.084	0.188	0.032
White	-0.016	0.021	-0.001	-0.015	0.006	-0.024	-0.026	0.007	0.011
Black	-0.010	0.033	-0.136	0.008	-0.045	-0.108	-0.035	0.012	-0.023
Asian	-0.045	-0.166	-0.036	-0.021	-0.073	0.057	0.098	0.032	0.036
hispanic_reg	-0.025	-0.011	0.016	-0.087	0.006	-0.016	0.011	0.021	-0.014
no_child	-0.028	0.024	-0.011	-0.001	-0.008	0.014	-0.013	0.019	0.007
agehh_under35	-0.089	0.033	0.099	-0.019	0.081	0.101	0.065	-0.045	-0.182
agehh_35to49	-0.047	0.011	0.046	-0.044	-0.012	0.063	0.064	-0.016	-0.118
agehh_50to64	-0.022	0.010	0.053	-0.030	0.008	0.062	0.030	-0.017	-0.047
eduhh_hsandless	-0.006	0.037	0.001	0.012	0.017	-0.012	-0.023	0.021	0.04

significant (see Table 4), meaning there is no significant difference detected between these regional groups and the base group, the Pacific area. In addition, household locating in New England, Middle Atlantic, and West North Central had less consumptions of walnuts compared to the one locating in Pacific area. Less consumption of almonds were revealed in the area of West North Central. Cashews were purchased more in the area of East South Central and Mountain compared to the Pacific area. Households from East South Central area had larger consumption of macadamia nuts. Walnuts were not well preferred by the households living in the New England, Middle Atlantic, and West North Central area.

Households with different ethnicity yield different preference of nuts products in our study. African Americans had less consumption of almonds and macadamia nuts compared to other race[20]. Asians have large consumption of pistachios and less of pecans and walnuts. Households who are Hispanic origins prefer less cashews.

Over 86% of households in our sample does not have children under the age of 18 and the children dummy variable is not significant across all nuts category, indicative of insignificant differences between households with and without children. About 2 percent of households' head are younger than the age of 35, they prefer more of almonds and less of tree nuts products compared to the households of over 65 years old. Accounting for 20% of our sample, households' head age between 35 and 49 years who had larger consumption of almonds, macadamia nuts, and pistachios compared to

[20] The base group in race is other, see Table 2.

households with head being over 65 years old. Overall, older households tend to increase their consumption regarding the overall tree nuts products.

Meanwhile, we also found that if a household's head has at least some college education, these household consumed more overall tree nuts. Additionally, we also investigated the habitual consumption of households by including their consumption in previous period. As shown in Table 4, all dependent variables are positively correlated with their lag one observations, meaning households have a habitual purchase by continuously buying what they have bought before.

Based on these empirical findings, we provide an overall discussion in the following section as well as limitation and future studies.

4.6. Concluding Remarks

We examined the demand of various nuts products in the United States using household level scanner data and dynamic unobserved effects Tobit model[21]. The panel data used in this study is provided and derived from the Nielsen Homescan Panel 2009-2015. In this study, we estimated the demand of nine nuts products, investigated the effects of socio-demographic factors, studied the habitual behavior of nuts purchase, addressed sample selection bias, and controlled for unobserved heterogeneity. According to the best of our knowledge, none of studies in the literature covers any aforementioned contributions of this research.

---

[21] The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

Results of Tobit model revealed inelastic demand of peanuts and all of tree nuts products. Meanwhile, estimated income elasticities showed that peanuts and tree nuts are necessity. In addition, based on socio-demographic effects, households with older age, higher education level, and higher income level consumed more nuts. Presence of children was not a driver of nuts purchases. Overall, household with different socio-demographic background yield inconsistent preference pattern in this study.

Although substitution among nut products was far more common, some complementarity among nut products was evident. Information concerning the substitution and complementary patterns among peanuts and tree nuts is important strategically for stakeholders in the nuts market.

These findings contributed to the nuts industries for stakeholders and processors in terms of pricing strategy, consumer targeting, and market segmentations. One thing to be noted and market practitioners must be careful about is that the demand of nuts products estimates were sensitive to the model choice that different models might yield inconsistent results.

Our study contributed to the literature by demostrating a more up-to-date analysis with recent data and filling up the gap existing in the literature for the demand of nuts products in the Unites States. Several limitations need attention in future studies. Alternative model is worth of exploring, such as mapping and semi-nonparametric methods to deal with data-censoring issue other than the technique used in this study. It will be meaningful comparing estimates across different models. This study is also limited in the data use and the limitation rests in the representativeness of household samples and the categorization of demographics. The household in Nielsen data is associated with older age level, higher income level, and no children under the age of 18 years old.

4.7. References

Amemiya, T., 1984. "Tobit models: A survey." *Journal of econometrics*, *24*(1-2), pp.3-61.

Crespi, J.M., and A. Chacon-Cascante. 2004. "Do US marketing orders have much market power? An examination of the Almond Board of California." *Agribusiness* 20:1.

Dhaliwal, H. S., 1972. "An econometric investigation of demand interrelationships among tree nuts and peanuts." Ph.D. thesis, Oregon State University.

Florkowski, W.J., and C. Sarmiento. 2005. "The examination of pecan price differences using spatial correlation estimation." *Applied Economics* 37:271–278.

Green, R.D. 1999. "Demand for California Agricultural Commodities." *Outlook* 99:11.

Kim, Y. and Dharmasena, S., 2018. "Price Discovery and Integration in US Pecan Markets." *Journal of Food Distribution Research*, *49*(856-2018-3108).

Lee, I. M., 1950. "Statistical Analysis of Annual Average Returns to Growers of Almonds, 1924/25 to 1948/49." 103, University of California, College of Agriculture, Agriculture Experiment Station.

Lerner, E. B., 1959. "An econometric analysis of the demand for pecans with special reference to the demand interrelationships among domestic tree nuts." Ph.D. thesis, Oklahoma State University.

McDonald, J.F. and Moffitt, R.A., 1980. "The uses of Tobit analysis." *The review of economics and statistics*, pp.318-321.

[dataset] Nielsen at the Kilts Center for Marketing, 2009-2015, Consumer Panel Data, https://www.chicagobooth.edu/research/kilts/datasets/nielsen

Russo, C., Green, R., Howitt, R. E., 2008. "Estimation of supply and demand elasticities of california commodities."

SAS 9.4 (software). 2013. SAS 9.4 User's Guide. Cary, NV: SAS Institute Inc.

Stata (software). StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.

Wells, G. J., Miller, S. E., Thompson, C. S., 1986. "Farm level demand for pecans reconsidered." *Journal of Agricultural and Applied Economics*, 18(1), 157–160.

Wood, B.W., 2001. "Production unit trends and price characteristics within the United States pecan industry." *HortTechnology*, *11*(1), pp.110-118.

Wooldridge, J.M., 2010. Econometric analysis of cross section and panel data. MIT press.

5. CONCLUSIONS

Our study contributed to the literature by putting together a more up-to-date and thorough economic analysis of peanuts and tree nuts markets in the United States. The household-level scanner data were provided by Nielsen.[22]

We first developed the LA/GAIDS model to address the demand for various nut products in the United States incorporating variety-seeking behavior and pre-commitment levels focusing on the year from 2004 to 2015. Second, we used the same data focusing on a panel year and binary choice approaches to investigate the demand for peanuts and tree nuts products by identifying demand driving forces and profiling households. Last, we examined the demand of various nuts products again using household level scanner data and dynamic unobserved effects Tobit model focusing on the year from 2009 to 2015. The main findings are as follows.

Variety-seeking behavior was only revealedfor peanuts. Pre-commitment levels were evident for pecans, almonds, walnuts, macadamia nuts, pistachios, and other nuts. Although most of pre-committed quantities were negative, almost all of them generally exhibited an upward trend over the period from 2004 through 2015. As such, households were relatively more committed to purchasing peanuts and tree nuts over this period. Moreover, the demand for peanuts, almonds, cashews, walnuts, pistachios, other nuts

---

[22] The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

and mixed nuts was found to be inelastic. In general, the various nut products were net substitutes. Although substitution among nut products was far more common, some complementarity was also evident.

Regarding the effects of socio-demographic factors, households with older head, higher education level, and higher income level consumed more nut products. Presence of children was not found to be a driver of nuts purchases. Overall, household with different socio-demographic background yielded inconsistent preference pattern at different time point of a year and in different locations.

These findings contributed to the nuts industries for stakeholders and processors in terms of pricing strategy, consumer targeting, and market segmentations. One thing to be noted and market practitioners must be careful about is that the demand of nuts products estimates were sensitive to the model choice that different models might yield somewhat different results. Our studies contributed to the literature by demonstrating a more up-to-date analysis with recent data and filling up the gap existing in the literature for the demand of nuts products in the Unites States.

Several limitations might be concerned and addressed in the future studies. One limitation of our analysis concerns the implicit assumption of separability of peanuts and tree nuts from other snack products. Additionally, although the LA/GAIDS model accounts for price, total expenditure, seasonality, variety-seeking behavior and pre-commitment, other explanatory factors were excluded from the analysis. Branded and generic advertising expenditures were not included in this analysis due to unavailability of monthly data over the period 2004 to 2015. Perhaps other macroeconomic factors could have been included in the LA/GAIDS model. Another set of limitations of our studies rest in the binary choice model and Tobit model we used and representativeness of our data. Our sample in the Nielsen are concentrated on older households, and households without children. Our study also does not consider away-from-home purchases. On top of that, alternative model is worth of exploring, such as mapping and semi-nonparametric methods to deal with data-censoring issue other than the technique used in this study. It will be meaningful comparing estimates across different models. Despite all these limitations, we believe our study add to the literature by providing a household-level analysis for nuts markets by examining demand and profiling consumers.

APPENDIX A

PRICE IMPUTATION OF CROSS-SECTIONAL DATA ASSOCIATED WITH

STUDY II

Table 5.1 Price Imputation Peanuts Study II

The SAS System							
The REG Procedure							
Dependent Variable: lp_Peanuts							
Number of Observations Read	245,520						
Number of Observations Used	56,314						
Number of Observations with Missing Values	189,206						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	13.00	33.14	2.55	14.62	0.00		
Error	56300.00	9816.28	0.17				
Corrected Total	56313.00	9849.42					
Root MSE	0.42	R-Square	0.00				
Dependent Mean	-1.79	Adj R-Sq	0.00				
Coeff Var	-23.27						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-1.79	0.03	-58.92	0.00	.	0.00
q1	1	-0.04	0.00	-8.06	0.00	0.67	1.49
q2	1	-0.03	0.00	-6.17	0.00	0.68	1.48
q3	1	-0.03	0.00	-5.16	0.00	0.68	1.48
lhsize	1	-0.01	0.00	-2.47	0.01	0.92	1.09
lincome	1	0.00	0.00	-0.13	0.90	0.91	1.09
MiddleAtlantic	1	0.02	0.01	2.19	0.03	0.32	3.09
ENCentral	1	0.02	0.01	1.68	0.09	0.25	4.05
WNCentral	1	0.04	0.01	4.00	0.00	0.38	2.62
SouthAtlantic	1	0.03	0.01	3.52	0.00	0.23	4.36
ESCentral	1	0.04	0.01	3.89	0.00	0.44	2.26
WSCentral	1	0.06	0.01	5.77	0.00	0.35	2.89
Mountain	1	0.03	0.01	2.42	0.02	0.44	2.25
Pacific	1	0.07	0.01	7.09	0.00	0.33	2.99

Table 5.2 Price Imputation Pecans Study II

The SAS System							
The REG Procedure							
Dependent Variable: lp_Pecans							
Number of Observations Read	245,520						
Number of Observations Used	18,093						
Number of Observations with Missing Values	227,427						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	13.00	31.12	2.39	15.81	0.00		
Error	18079.00	2737.04	0.15				
Corrected Total	18092.00	2768.16					
Root MSE	0.39	R-Square	0.01				
Dependent Mean	-0.48	Adj R-Sq	0.01				
Coeff Var	-80.38783						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.29	0.05	-5.28	0.00	.	0.00
q1	1	0.02	0.01	2.26	0.02	0.87	1.14
q2	1	0.05	0.01	6.69	0.00	0.88	1.14
q3	1	0.07	0.01	8.72	0.00	0.88	1.14
lhsize	1	0.02	0.01	2.39	0.02	0.92	1.09
lincome	1	-0.02	0.00	-4.07	0.00	0.91	1.10
MiddleAtlantic	1	0.01	0.02	0.65	0.52	0.36	2.78
ENCentral	1	-0.01	0.02	-0.54	0.59	0.21	4.71
WNCentral	1	-0.02	0.02	-1.20	0.23	0.34	2.98
SouthAtlantic	1	-0.01	0.02	-0.43	0.67	0.18	5.41
ESCentral	1	0.02	0.02	0.86	0.39	0.36	2.77
WSCentral	1	-0.03	0.02	-1.53	0.13	0.25	4.01
Mountain	1	-0.08	0.02	-4.45	0.00	0.36	2.77
Pacific	1	-0.07	0.02	-3.89	0.00	0.33	3.07

Table 5.3 Price Imputation Almonds Study II

The SAS System							
The REG Procedure							
Dependent Variable: lp_Almonds							
Number of Observations Read	245,520						
Number of Observations Used	31,466						
Number of Observations with Missing Values	214,054						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	13.00	108.39	8.34	49.80	<.0001		
Error	31452.00	5265.43	0.17				
Corrected Total	31465.00	5373.81					
Root MSE	0.41	R-Square	0.02				
Dependent Mean	-0.65	Adj R-Sq	0.02				
Coeff Var	-63.19						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.65	0.04	-14.79	0.00	.	0.00
q1	1	-0.15	0.01	-23.34	0.00	0.66	1.52
q2	1	-0.08	0.01	-11.78	0.00	0.66	1.51
q3	1	-0.04	0.01	-6.09	0.00	0.67	1.49
lhsize	1	0.01	0.00	1.19	0.23	0.92	1.09
lincome	1	0.01	0.00	1.48	0.14	0.92	1.09
MiddleAtlantic	1	-0.01	0.01	-0.67	0.50	0.33	3.02
ENCentral	1	-0.02	0.01	-1.42	0.16	0.27	3.71
WNCentral	1	0.02	0.01	1.37	0.17	0.39	2.54
SouthAtlantic	1	-0.01	0.01	-0.56	0.58	0.24	4.13
ESCentral	1	0.02	0.01	1.57	0.12	0.49	2.03
WSCentral	1	0.01	0.01	1.10	0.27	0.36	2.80
Mountain	1	0.03	0.01	2.45	0.01	0.44	2.26
Pacific	1	-0.02	0.01	-1.77	0.08	0.33	3.01

Table 5.4 Price Imputation Cashews Study II

The SAS System							
The REG Procedure							
Dependent Variable: lp_Cashews							
Number of Observations Read	245,520						
Number of Observations Used	28,715						
Number of Observations with Missing Values	216,805						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	13.00	45.31	3.49	29.70	0.00		
Error	28701.00	3368.17	0.12				
Corrected Total	28714.00	3413.48					
Root MSE	0.34	R-Square	0.01				
Dependent Mean	-0.81	Adj R-Sq	0.01				
Coeff Var	-42.13554						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.78	0.04	-22.25	0.00	.	0.00
q1	1	-0.07	0.01	-12.01	0.00	0.70	1.44
q2	1	-0.04	0.01	-6.52	0.00	0.70	1.44
q3	1	0.01	0.01	1.65	0.10	0.70	1.43
lhsize	1	-0.01	0.00	-2.15	0.03	0.91	1.10
lincome	1	0.01	0.00	1.64	0.10	0.91	1.10
MiddleAtlantic	1	-0.02	0.01	-1.73	0.08	0.31	3.20
ENCentral	1	-0.07	0.01	-6.98	0.00	0.24	4.10
WNCentral	1	-0.06	0.01	-5.47	0.00	0.39	2.56
SouthAtlantic	1	-0.04	0.01	-4.02	0.00	0.22	4.52
ESCentral	1	-0.06	0.01	-4.67	0.00	0.46	2.17
WSCentral	1	-0.06	0.01	-5.05	0.00	0.36	2.74
Mountain	1	-0.08	0.01	-7.04	0.00	0.42	2.39
Pacific	1	-0.09	0.01	-8.13	0.00	0.34	2.95

Table 5.5 Price Imputation Walnuts Study II

The SAS System

The REG Procedure

Dependent Variable: lp_Walnuts

Number of Observations Read	245,520
Number of Observations Used	21,591
Number of Observations with Missing Values	223,929

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	13.00	139.99	10.77	80.15	0.00
Error	21577.00	2898.86	0.13		
Corrected Total	21590.00	3038.85			
Root MSE	0.37	R-Square	0.05		
Dependent Mean	-0.63	Adj R-Sq	0.05		
Coeff Var	-58.36				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.43	0.04	-9.79	0.00	.	0.00
q1	1	0.12	0.01	18.47	0.00	0.82	1.21
q2	1	0.16	0.01	22.62	0.00	0.83	1.20
q3	1	0.13	0.01	19.56	0.00	0.83	1.21
lhsize	1	0.03	0.01	4.97	0.00	0.91	1.10
lincome	1	-0.02	0.00	-6.12	0.00	0.91	1.10
MiddleAtlantic	1	-0.05	0.01	-3.90	0.00	0.35	2.87
ENCentral	1	-0.09	0.01	-7.61	0.00	0.29	3.43
WNCentral	1	0.01	0.01	0.75	0.45	0.48	2.10
SouthAtlantic	1	0.00	0.01	0.33	0.74	0.28	3.61
ESCentral	1	0.03	0.01	1.98	0.05	0.54	1.85
WSCentral	1	-0.03	0.01	-2.15	0.03	0.42	2.37
Mountain	1	0.00	0.01	0.12	0.90	0.50	2.01
Pacific	1	-0.05	0.01	-3.94	0.00	0.40	2.49

Table 5.6 Price Imputation Macadamia Nuts Study II

The SAS System							
The REG Procedure							
Dependent Variable: lp_Maca							
Number of Observations Read	245,520						
Number of Observations Used	1,388						
Number of Observations with Missing Values	244,132						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	13.00	5.01	0.39	3.63	0.00		
Error	1374.00	145.90	0.11				
Corrected Total	1387.00	150.91					
Root MSE	0.33	R-Square	0.03				
Dependent Mean	-0.05	Adj R-Sq	0.02				
Coeff Var	-706.70						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.29	0.17	-1.70	0.09	.	0.00
q1	1	0.03	0.02	1.36	0.17	0.76	1.32
q2	1	0.03	0.02	1.39	0.16	0.75	1.34
q3	1	0.01	0.02	0.27	0.79	0.75	1.33
lhsize	1	-0.01	0.02	-0.77	0.44	0.88	1.14
lincome	1	0.02	0.02	1.62	0.11	0.88	1.14
MiddleAtlantic	1	-0.15	0.04	-3.46	0.00	0.36	2.80
ENCentral	1	0.01	0.04	0.34	0.73	0.36	2.79
WNCentral	1	0.04	0.05	0.79	0.43	0.50	1.98
SouthAtlantic	1	-0.03	0.04	-0.73	0.47	0.33	3.07
ESCentral	1	-0.09	0.05	-1.75	0.08	0.59	1.70
WSCentral	1	0.01	0.05	0.32	0.75	0.44	2.25
Mountain	1	-0.01	0.05	-0.29	0.77	0.44	2.29
Pacific	1	-0.06	0.04	-1.46	0.14	0.32	3.10

Table 5.7 Price Imputation Pistachios Study II

The SAS System							
The REG Procedure							
Dependent Variable: lp_Pist							
Number of Observations Read	245,520						
Number of Observations Used	17,919						
Number of Observations with Missing Values	227,601						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	13.00	22.52	1.73	12.87	0.00		
Error	17905.00	2410.27	0.13				
Corrected Total	17918.00	2432.79					
Root MSE	0.37	R-Square	0.01				
Dependent Mean	-0.53	Adj R-Sq	0.01				
Coeff Var	-69.42						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.29	0.05	-5.94	0.00	.	0.00
q1	1	-0.01	0.01	-2.01	0.04	0.68	1.47
q2	1	0.04	0.01	4.99	0.00	0.70	1.43
q3	1	0.03	0.01	4.40	0.00	0.71	1.41
lhsize	1	0.00	0.01	0.78	0.43	0.92	1.08
lincome	1	-0.03	0.00	-6.25	0.00	0.92	1.08
MiddleAtlantic	1	0.06	0.01	3.94	0.00	0.31	3.25
ENCentral	1	0.01	0.01	0.98	0.33	0.27	3.66
WNCentral	1	0.04	0.02	2.62	0.01	0.44	2.26
SouthAtlantic	1	0.06	0.01	4.54	0.00	0.24	4.09
ESCentral	1	0.06	0.02	3.35	0.00	0.52	1.91
WSCentral	1	0.06	0.02	3.82	0.00	0.40	2.51
Mountain	1	0.02	0.02	0.96	0.34	0.44	2.25
Pacific	1	0.04	0.01	2.94	0.00	0.35	2.88

Table 5.8 Price Imputation Mixed Nuts Study II

The SAS System							
The REG Procedure							
Dependent Variable: lp_Mixed							
Number of Observations Read	245,520						
Number of Observations Used	27,583						
Number of Observations with Missing Values	217,937						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	13.00	67.81	5.22	30.11	0.00		
Error	27569.00	4776.54	0.17				
Corrected Total	27582.00	4844.35					
Root MSE	0.42	R-Square	0.01				
Dependent Mean	-0.88	Adj R-Sq	0.01				
Coeff Var	-47.03						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-1.55	0.05	-33.94	0.00	.	0.00
q1	1	-0.03	0.01	-4.68	0.00	0.71	1.41
q2	1	0.02	0.01	3.53	0.00	0.72	1.40
q3	1	0.02	0.01	2.63	0.01	0.72	1.39
lhsize	1	-0.02	0.01	-2.76	0.01	0.92	1.09
lincome	1	0.06	0.00	15.09	0.00	0.92	1.09
MiddleAtlantic	1	0.02	0.01	1.65	0.10	0.31	3.18
ENCentral	1	-0.03	0.01	-1.84	0.07	0.25	3.97
WNCentral	1	-0.01	0.01	-0.91	0.36	0.35	2.83
SouthAtlantic	1	0.00	0.01	-0.37	0.71	0.20	5.00
ESCentral	1	-0.02	0.02	-1.59	0.11	0.40	2.47
WSCentral	1	0.02	0.01	1.18	0.24	0.30	3.29
Mountain	1	0.01	0.02	0.89	0.38	0.39	2.59
Pacific	1	0.04	0.01	2.55	0.01	0.30	3.37

APPENDIX B

PRICE IMPUTATION OF PANEL DATA ASSOCIATED WITH STUDY III

Table 5.9 Price Imputation Peanuts Study III

The SAS System							
The REG Procedure							
Dependent Variable: lp_peanutsnew							
Number of Observations Read	166,677						
Number of Observations Used	85,264						
Number of Observations with Missing Values	81,413						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	16.00	753.39	47.09	274.39	0.00		
Error	85247.00	14629.00	0.17				
Corrected Total	85263.00	15382.00					
Root MSE	0.41	R-Square	0.05				
Dependent Mean	-1.86	Adj R-Sq	0.05				
Coeff Var	-22.22						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-1.82	0.03	-71.02	0.00	.	0.00
D2009	1	-0.17	0.01	-32.42	0.00	0.58	1.73
D2010	1	-0.19	0.01	-36.27	0.00	0.58	1.73
D2011	1	-0.14	0.01	-25.42	0.00	0.58	1.73
D2012	1	0.05	0.01	8.65	0.00	0.59	1.71
D2013	1	0.01	0.01	2.14	0.03	0.58	1.71
D2014	1	0.00	0.01	0.08	0.94	0.58	1.73
lhsize	1	-0.03	0.00	-8.02	0.00	0.90	1.11
lincome	1	0.00	0.00	0.90	0.37	0.90	1.11
MiddleAtlantic	1	0.01	0.01	1.38	0.17	0.31	3.26
ENCentral	1	0.01	0.01	1.44	0.15	0.23	4.39
WNCentral	1	0.03	0.01	3.80	0.00	0.33	2.99
SouthAtlantic	1	0.00	0.01	0.60	0.55	0.22	4.47
ESCentral	1	0.02	0.01	1.98	0.05	0.44	2.29
WSCentral	1	0.03	0.01	3.64	0.00	0.33	3.00
Mountain	1	-0.01	0.01	-1.05	0.29	0.42	2.40
Pacific	1	0.03	0.01	3.77	0.00	0.32	3.17

Table 5.10 Price Imputation Pecans Study III

The SAS System							
The REG Procedure							
Dependent Variable: lp_pecans							
Number of Observations Read	166,677						
Number of Observations Used	41,224						
Number of Observations with Missing Values	125,453						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	16.00	732.34	45.77	321.44	0.00		
Error	41207.00	5867.70	0.14				
Corrected Total	41223.00	6600.04					
Root MSE	0.38	R-Square	0.11				
Dependent Mean	-0.59	Adj R-Sq	0.11				
Coeff Var	-63.51						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.53	0.04	-14.67	0.00	.	0.00
D2009	1	-0.32	0.01	-46.30	0.00	0.56	1.80
D2010	1	-0.20	0.01	-29.00	0.00	0.57	1.75
D2011	1	0.02	0.01	2.35	0.02	0.58	1.72
D2012	1	0.08	0.01	10.63	0.00	0.59	1.70
D2013	1	-0.06	0.01	-7.94	0.00	0.57	1.76
D2014	1	-0.02	0.01	-3.18	0.00	0.58	1.73
lhsize	1	0.00	0.00	-0.58	0.56	0.90	1.11
lincome	1	0.00	0.00	1.52	0.13	0.90	1.12
MiddleAtlantic	1	0.01	0.01	0.42	0.67	0.32	3.11
ENCentral	1	-0.03	0.01	-2.51	0.01	0.17	5.75
WNCentral	1	-0.01	0.01	-1.01	0.31	0.27	3.68
SouthAtlantic	1	-0.04	0.01	-3.53	0.00	0.17	5.77
ESCentral	1	-0.04	0.01	-2.88	0.00	0.34	2.93
WSCentral	1	-0.08	0.01	-6.90	0.00	0.24	4.25
Mountain	1	-0.10	0.01	-8.05	0.00	0.33	3.07
Pacific	1	-0.05	0.01	-4.00	0.00	0.30	3.39

Table 5.11 Price Imputation Almonds Study III

The SAS System							
The REG Procedure							
Dependent Variable: lp_almonds							
Number of Observations Read	166,677						
Number of Observations Used	62,939						
Number of Observations with Missing Values	103,738						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	16.00	984.45	61.53	324.49	0.00		
Error	62922.00	11931.00	0.19				
Corrected Total	62938.00	12915.00					
Root MSE	0.44	R-Square	0.08				
Dependent Mean	-0.91	Adj R-Sq	0.08				
Coeff Var	-48.06						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.84	0.03	-24.96	0.00	.	0.00
D2009	1	-0.33	0.01	-48.74	0.00	0.53	1.90
D2010	1	-0.33	0.01	-49.23	0.00	0.53	1.90
D2011	1	-0.35	0.01	-51.84	0.00	0.52	1.94
D2012	1	-0.33	0.01	-49.03	0.00	0.52	1.93
D2013	1	-0.22	0.01	-32.27	0.00	0.53	1.89
D2014	1	-0.10	0.01	-14.93	0.00	0.54	1.84
lhsize	1	-0.01	0.00	-3.48	0.00	0.89	1.12
lincome	1	0.02	0.00	7.09	0.00	0.89	1.12
MiddleAtlantic	1	-0.02	0.01	-1.71	0.09	0.31	3.18
ENCentral	1	-0.04	0.01	-4.33	0.00	0.24	4.24
WNCentral	1	-0.01	0.01	-0.64	0.52	0.34	2.92
SouthAtlantic	1	-0.05	0.01	-5.29	0.00	0.23	4.44
ESCentral	1	-0.02	0.01	-2.24	0.03	0.47	2.13
WSCentral	1	-0.04	0.01	-4.01	0.00	0.33	3.05
Mountain	1	-0.08	0.01	-7.75	0.00	0.39	2.56
Pacific	1	-0.08	0.01	-7.94	0.00	0.29	3.42

Table 5.12 Price Imputation Cashews Study III

The SAS System							
The REG Procedure							
Dependent Variable: lp_cashews							
Number of Observations Read	166,677						
Number of Observations Used	53,742						
Number of Observations with Missing Values	112,935						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	16.00	958.49	59.91	524.92	0.00		
Error	53725.00	6131.29	0.11				
Corrected Total	53741.00	7089.78					
Root MSE	0.34	R-Square	0.14				
Dependent Mean	-0.97	Adj R-Sq	0.13				
Coeff Var	-34.74						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-1.05	0.03	-39.77	0.00	.	0.00
D2009	1	-0.33	0.01	-61.29	0.00	0.55	1.81
D2010	1	-0.32	0.01	-59.13	0.00	0.55	1.80
D2011	1	-0.21	0.01	-37.15	0.00	0.58	1.74
D2012	1	-0.02	0.01	-4.34	0.00	0.59	1.69
D2013	1	-0.07	0.01	-12.00	0.00	0.58	1.73
D2014	1	-0.06	0.01	-11.26	0.00	0.57	1.74
lhsize	1	-0.03	0.00	-9.32	0.00	0.89	1.12
lincome	1	0.02	0.00	10.24	0.00	0.89	1.12
MiddleAtlantic	1	0.02	0.01	2.00	0.05	0.31	3.24
ENCentral	1	-0.02	0.01	-3.01	0.00	0.24	4.22
WNCentral	1	-0.02	0.01	-2.60	0.01	0.36	2.77
SouthAtlantic	1	-0.02	0.01	-2.07	0.04	0.23	4.35
ESCentral	1	-0.02	0.01	-2.13	0.03	0.47	2.14
WSCentral	1	-0.01	0.01	-1.22	0.22	0.37	2.70
Mountain	1	-0.01	0.01	-1.72	0.09	0.42	2.41
Pacific	1	-0.01	0.01	-1.49	0.14	0.34	2.97

Table 5.13 Price Imputation Walnuts Study III

The SAS System							
The REG Procedure							
Dependent Variable: lp_walnutsnew							
Number of Observations Read	166,677						
Number of Observations Used	49,214						
Number of Observations with Missing Values	117,463						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	16.00	1514.68	94.67	700.06	0.00		
Error	49197.00	6652.84	0.14				
Corrected Total	49213.00	8167.52					
Root MSE	0.37	R-Square	0.19				
Dependent Mean	-0.85	Adj R-Sq	0.19				
Coeff Var	-43.23						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.71	0.03	-22.95	0.00	.	0.00
D2009	1	-0.44	0.01	-71.27	0.00	0.54	1.87
D2010	1	-0.38	0.01	-60.64	0.00	0.55	1.82
D2011	1	-0.23	0.01	-36.57	0.00	0.55	1.80
D2012	1	-0.09	0.01	-14.61	0.00	0.56	1.78
D2013	1	-0.06	0.01	-9.33	0.00	0.56	1.77
D2014	1	0.02	0.01	2.94	0.00	0.58	1.73
lhsize	1	-0.01	0.00	-2.39	0.02	0.90	1.12
lincome	1	0.01	0.00	2.66	0.01	0.89	1.12
MiddleAtlantic	1	-0.04	0.01	-4.60	0.00	0.32	3.12
ENCentral	1	-0.08	0.01	-10.03	0.00	0.25	3.95
WNCentral	1	0.01	0.01	1.26	0.21	0.40	2.47
SouthAtlantic	1	-0.01	0.01	-0.77	0.44	0.26	3.86
ESCentral	1	-0.02	0.01	-2.34	0.02	0.51	1.96
WSCentral	1	-0.05	0.01	-5.96	0.00	0.40	2.51
Mountain	1	-0.06	0.01	-6.21	0.00	0.44	2.25
Pacific	1	-0.05	0.01	-5.54	0.00	0.37	2.72

Table 5.14 Price Imputation Macadamia Nuts Study III

The SAS System
The REG Procedure
Dependent Variable: lp_macanuts

Number of Observations Read	166,677
Number of Observations Used	4,302
Number of Observations with Missing Values	162,375

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	16.00	168.28	10.52	92.26	0.00
Error	4285.00	488.49	0.11		
Corrected Total	4301.00	656.77			
Root MSE	0.34	R-Square	0.26		
Dependent Mean	-0.24	Adj R-Sq	0.25		
Coeff Var	-140.44				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.20	0.10	-2.04	0.04	.	0.00
D2009	1	-0.48	0.02	-24.98	0.00	0.45	2.22
D2010	1	-0.39	0.02	-19.51	0.00	0.49	2.05
D2011	1	-0.18	0.02	-8.58	0.00	0.53	1.90
D2012	1	-0.03	0.02	-1.22	0.22	0.55	1.83
D2013	1	-0.05	0.02	-2.19	0.03	0.52	1.92
D2014	1	0.00	0.02	0.06	0.95	0.52	1.92
lhsize	1	-0.01	0.01	-1.13	0.26	0.90	1.12
lincome	1	0.01	0.01	1.24	0.21	0.89	1.12
MiddleAtlantic	1	0.03	0.03	1.29	0.20	0.36	2.82
ENCentral	1	0.06	0.03	2.16	0.03	0.29	3.39
WNCentral	1	0.09	0.03	3.17	0.00	0.47	2.12
SouthAtlantic	1	0.02	0.03	0.70	0.49	0.28	3.56
ESCentral	1	0.07	0.03	2.24	0.03	0.55	1.81
WSCentral	1	0.05	0.03	1.94	0.05	0.40	2.47
Mountain	1	0.03	0.03	1.06	0.29	0.40	2.53
Pacific	1	0.03	0.03	1.29	0.20	0.28	3.59

Table 5.15 Price Imputation Pistachios Study III

The SAS System
The REG Procedure
Dependent Variable: lp_pistachios

Number of Observations Read	166,677
Number of Observations Used	36,679
Number of Observations with Missing Values	129,998

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	16.00	1404.10	87.76	619.91	0.00
Error	36662.00	5189.98	0.14		
Corrected Total	36678.00	6594.08			
Root MSE	0.38	R-Square	0.21		
Dependent Mean	-0.83	Adj R-Sq	0.21		
Coeff Var	-45.35				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-0.46	0.04	-12.71	0.00	.	0.00
D2009	1	-0.62	0.01	-75.44	0.00	0.61	1.63
D2010	1	-0.52	0.01	-65.32	0.00	0.59	1.69
D2011	1	-0.38	0.01	-52.25	0.00	0.53	1.88
D2012	1	-0.31	0.01	-43.50	0.00	0.51	1.97
D2013	1	-0.22	0.01	-30.03	0.00	0.53	1.89
D2014	1	-0.09	0.01	-12.11	0.00	0.55	1.81
lhsize	1	-0.01	0.00	-2.68	0.01	0.90	1.11
lincome	1	0.00	0.00	-1.42	0.16	0.90	1.11
MiddleAtlantic	1	0.01	0.01	1.20	0.23	0.32	3.17
ENCentral	1	-0.03	0.01	-3.20	0.00	0.27	3.77
WNCentral	1	0.03	0.01	2.36	0.02	0.41	2.44
SouthAtlantic	1	0.01	0.01	1.29	0.20	0.26	3.85
ESCentral	1	0.00	0.01	0.10	0.92	0.54	1.86
WSCentral	1	0.02	0.01	1.63	0.10	0.40	2.50
Mountain	1	-0.05	0.01	-4.64	0.00	0.42	2.38
Pacific	1	-0.09	0.01	-8.72	0.00	0.32	3.08

Table 5.16 Price Imputation Mixed Nuts Study III

The SAS System							
The REG Procedure							
Dependent Variable: lp_mix							
Number of Observations Read	166,677						
Number of Observations Used	54,714						
Number of Observations with Missing Values	111,963						
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	16.00	1006.09	62.88	435.63	0.00		
Error	54697.00	7895.21	0.14				
Corrected Total	54713.00	8901.30					
Root MSE	0.38	R-Square	0.11				
Dependent Mean	-1.07	Adj R-Sq	0.11				
Coeff Var	-35.38						
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-1.45	0.03	-48.36	0.00	.	0.00
D2009	1	-0.34	0.01	-55.73	0.00	0.55	1.81
D2010	1	-0.33	0.01	-53.36	0.00	0.55	1.81
D2011	1	-0.25	0.01	-40.04	0.00	0.56	1.78
D2012	1	-0.10	0.01	-16.44	0.00	0.57	1.77
D2013	1	-0.09	0.01	-14.07	0.00	0.56	1.77
D2014	1	-0.05	0.01	-7.61	0.00	0.57	1.76
lhsize	1	-0.05	0.00	-12.65	0.00	0.90	1.11
lincome	1	0.06	0.00	20.71	0.00	0.90	1.11
MiddleAtlantic	1	-0.01	0.01	-0.98	0.33	0.32	3.13
ENCentral	1	-0.03	0.01	-4.00	0.00	0.24	4.16
WNCentral	1	-0.03	0.01	-2.69	0.01	0.34	2.98
SouthAtlantic	1	-0.03	0.01	-4.08	0.00	0.21	4.66
ESCentral	1	-0.05	0.01	-5.24	0.00	0.43	2.32
WSCentral	1	-0.02	0.01	-2.41	0.02	0.33	3.07
Mountain	1	-0.02	0.01	-2.43	0.02	0.40	2.49
Pacific	1	0.01	0.01	1.30	0.19	0.31	3.24

Table 5.17 Price Imputation Tree Nuts Study III

The SAS System
The REG Procedure
Dependent Variable: lp_treenuts

Number of Observations Read	166,677
Number of Observations Used	134,477
Number of Observations with Missing Values	32,200

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	16.00	2239.78	139.99	1050.46	0.00
Error	134460.00	17918.00	0.13		
Corrected Total	134476.00	20158.00			
Root MSE	0.37	R-Square	0.11		
Dependent Mean	-0.96	Adj R-Sq	0.11		
Coeff Var	-38.07				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-1.04	0.02	-57.72	0.00	.	0.00
D2009	1	-0.35	0.00	-92.92	0.00	0.58	1.74
D2010	1	-0.31	0.00	-83.44	0.00	0.58	1.73
D2011	1	-0.21	0.00	-56.83	0.00	0.58	1.73
D2012	1	-0.12	0.00	-31.22	0.00	0.58	1.73
D2013	1	-0.09	0.00	-23.17	0.00	0.58	1.73
D2014	1	-0.04	0.00	-11.32	0.00	0.58	1.72
lhsize	1	-0.04	0.00	-20.46	0.00	0.90	1.12
lincome	1	0.03	0.00	17.98	0.00	0.89	1.12
MiddleAtlantic	1	-0.01	0.01	-2.20	0.03	0.31	3.26
ENCentral	1	-0.05	0.01	-9.05	0.00	0.24	4.20
WNCentral	1	-0.07	0.01	-11.88	0.00	0.35	2.83
SouthAtlantic	1	-0.03	0.01	-5.47	0.00	0.23	4.34
ESCentral	1	-0.04	0.01	-6.40	0.00	0.46	2.16
WSCentral	1	-0.04	0.01	-7.58	0.00	0.34	2.95
Mountain	1	-0.08	0.01	-14.24	0.00	0.40	2.47
Pacific	1	-0.07	0.01	-12.81	0.00	0.31	3.23