ABSTRACT

CLAY, DAWN MARIE. U.S. Consumption Habits of Whole Grains and Fruit Juices and the Impact of USDA Dietary Guidelines (Under the direction of Dr. Michael Wohlgenant.)

Given the rising national and personal healthcare costs, the health status of individuals has become a more salient issue. For decades, the USDA has attempted to influence consumers to make healthy food choices. Habits could either strengthen or counteract many of the dietary health objectives of the Dietary Guidelines for Americans (DGA). I measured habit formation and the potential impact of present bias on specific consumer food and beverage choices. The actions of the USDA in the DGA were analyzed to determine their effect on improving consumer health outcomes across many socio-economic groups. Long term, time-consistent but diminishing habits were found for whole grains, other grains and sugar-sweetened beverage commodities and myopic preferences were shown for fruit juices. Spatial dependence was found for other grains and sugar-sweetened beverages with lower neighborhood effects after the 2010 DGA were published.

The Marshallian own-price elasticities were only statistically significantly different between the linear approximation to the Almost Ideal Demand System (LA/AIDS) and the linear approximation to the Exact Affine Stone Index (LA/EASI) model for tea/coffee, fruit juices and bottled water commodities. Engel curves generated from LA/EASI were more sophisticated than what could be modeled using a linear or quadratic relationship. Therefore, adding demographic variables into the LA/EASI model allowed the Engel curves to be more responsive to different income elasticities as income changed.

Three Essays on American Consumption Habits of Whole Grains and Fruit Juices and the Impact of USDA Dietary Guidelines for Americans

by
Dawn Marie Clay

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Economics

Raleigh, North Carolina

2015

APPROVED BY:

Dr. Michael Wohlgenant Committee Chair	Dr. Walter Thurman
Dr. Nora Traum	Dr. Ivan Kandilov

DEDICATION

I dedicate this dissertation to my grandparents: Grandma and Grandpa, Mama and Dakky (family term). Without you, I would not have spent my life pursuing higher education.

BIOGRAPHY

Marie Clay was educated at the University of Michigan-Dearborn and Ann Arbor where she earned bachelors and master's degrees. She was able to attend North Carolina State University and work for the USDA Economic Research Service for internships during the summers.

ACKNOWLEDGMENTS

I would like to thank my dissertation committee members for helpful insights, advice and many comments on approaches for this endeavor. Dr. Wohlgenant on several occasions helped me with learning to avoid rabbit trails and decide when to pursue new paths of research inquiry when finding statistical or methodological difficulties. Other committee members helped to refocus me from topics that would not have been feasible and would have detracted from cohesiveness of the overall dissertation. Dr. Thurman was so helpful in leading me to realize my logical errors so that I would be able to follow troubleshoot those problems myself in future work. Dr. Traum spent several sessions guiding me to discover my limits as I tried to determine areas of research where I could become an expert and where I could not achieve such levels of understanding. She saved me many months of potential headaches and failure. Dr. Kandilov contributed to the cohesiveness of the committee that spanned the two main branches of the economics department.

I would also like to acknowledge the USDA Economic Research Service ERS for financial and technical support, NCSU for funding and training during the course of the last 5 years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Dissertation Introduction

The USDA desires to assist Americans in improving health outcomes by eating healthier foods and beverages. Pursuant to that goal, the USDA published the Dietary Guidelines for Americans (DGA) every 5 years since 1980 along with the Department of Health and Human Services. The USDA Economic Research Service also commissioned a survey to study food acquisitions and purchases of Americans. The 2012 USDA National Household Food Acquisition and Purchase Survey (FoodAPS) dataset was collected by Mathematica Policy Research to observe purchase behavior of households for seven days from food-at-home and food-away-from-home venues for each member of the household. Chapters 1 and 2 study habit formation for healthy foods and beverages before and after the 2010 DGA and chapter 3 uses the USDA FoodAPS dataset in two consumer demand models to discover price and income sensitivities of consumers for healthy and unhealthy nonalcoholic beverages. Summaries of each chapter are written in the following paragraphs.

American diets have higher amounts of cholesterol, fat, sugar and sodium that have led to heart attacks, strokes and diabetes. The USDA Dietary Guidelines for Americans attempts to inform individuals of low-calorie nutrient-dense foods and beverages that will help improve health outcomes. If households develop habits for healthier foods and beverages, the effects of the USDA DGA would have far reaching benefits for consumers' health. Chapter 1 explores rational habit formation for grains, fruit juices and sugar-sweetened beverages before and after the 2010 DGA were published. It also tests time preference that might mitigate the long term habit formation for whole grains and fruit juices.

Consumers' preferences are affected by the choices of others. Thus consumers can be drawn to purchase goods based not solely on their past purchases but based on purchase decisions by other consumers. Chapter 2 addresses the question which asks if households make purchase decisions based on purchase decisions of others in their neighborhoods. A given distance band from a block group centroid was chosen where there would be at least one nearest neighbor for each household in the dataset for a given grain or beverage commodity. An inverse distance weighting matrix was created exogenously from this household distance band from

respective block group centroids to use in estimating several spatial econometric models of external habit formation.

Chapter 3 addresses concerns of price sensitivity of consumers when purchasing six nonalcoholic beverage commodities. A new dataset, USDA's FoodAPS was used to estimate demand models in the linear approximation to the Almost Ideal Demand System (LA/AIDS) (Deaton and Muellbauer, 1980) and the linear approximation to the EASI (LA/EASI) (Lewbel and Pendakur, 2009). Given that the LA/AIDS is nested in the LA/EASI, tests of LA/AIDS as a restricted test of LA/EASI were conducted to determine if using a model with higher order polynomial estimation of the Engel curves and interaction terms with the price and income variables would statistically significantly modify those Engel curves and Marshallian price and income elasticity estimates as compared to the LA/AIDS model. The AIDS and LA/AIDS are very frequently used demand systems for food and beverages and total household consumption. Comparing a more flexible functional form in the LA/EASI that is more responsive to consumer demographic information provides a meaningful possibility to measure budget share distributions as a reaction to policy changes affecting income and prices.

Chapter 2 Habit Formation with Time-Preference Modeling

2.1 Introduction

Are the Dietary Guidelines for Americans (DGA) affecting dietary choices of Americans and if so, are these long-term changes? The USDA has published the DGA for several decades with the intent of informing consumers of the dietary choices they should make for healthier living outcomes. I seek to determine if the DGA has had any impact on US dietary choices and if the DGA has effectively induced either short or long term habit formation.

One concern affecting food and beverage choices is time preference. In addition, individuals may have an aversion to healthy foods because of their tastes and preferences. Under time pressures, the desire for large amounts of caffeine and quick energy from soft drinks and energy drinks can increase. Thus, planning to consume healthier items such as whole grains and fruit juices could be too difficult to follow during stressful times. Consumers might want to have the healthier outcomes that whole grains and fruit juices provide, but might enjoy the taste of refined grains and sugar-sweetened beverages. If under less stressful conditions, these individuals could focus more heavily on the current period with a lower regard for future utility maximization. The shift could move temporarily to today and away from the future because the benefits of today's consumption is higher than the cost of poorer health outcomes in the near future.

Another concern affecting healthy food and beverage choices is discerning which items are healthier than their counterparts. Whole grains are healthier than refined grains, but it might be difficult to decide which foods have whole grains and which are made from refined grains. Determining the difference between these two types of grains can require reading the list of ingredients to see if they contain whole wheat or enriched flour. Not all Americans are aware of the subtle differences between these goods.

Choosing between fruit juice and sugar-sweetened beverages could be equally challenging when trying to maintain a healthy diet. 100 percent fruit juices and sugar-sweetened beverages have natural sugars and added sugars respectively, but these sugars may be metabolized similarly. 100% fruit juices have been linked to weight gain in children who are overweight or

obese. The 2010 DGA warns of the potential dangers to those children when consuming 100 percent fruit juices. Some nutritionists recommend avoiding fruit juice if possible and instead consuming fruit. This way the natural sugars can be used to chew and further digest the fruit so that those sugars do not present as negative of health outcomes.

However, even with these negative results from fruit juice consumption, there are also benefits. Fruit juice contains vitamins A and C, folate and dietary fiber (if the juice has pulp). These benefits are not in sports drinks or soft drinks at the same levels if at all. However, some consumers might think that fruit juice is just as harmful as other drinks and decide to continue consuming sugar-sweetened beverages such as juice drinks and fruit smoothies.

Given perceived time pressures, individuals could make decisions based on two types of strategies: maximizing utility or responding to environmental cues. There is evidence that when deciding what and how much to eat/drink, consumers choose foods and beverages based on primal instincts and also based on environmental signals. Primal urges such as hunger and thirst play a definite role in timing of food choices, but the amount of food consumed is influenced by the amount of food served and the amount of food eaten by others in close proximity. The choice to consume food and drinks is not simply to assuage hunger and thirst, but to receive pleasure from the consumption and also to create or change moods or relive past positive experiences. Therefore, impulses from multiple consumptions motivations could have mitigated effects on future consumption goals as the person gets close to that time period of consumption.

Individuals may lose sight of healthy eating goals during times of hunger or when environmental pressures are strong enough to dissuade healthy eating. Societal indicators for when to eat can be cultural such as not letting a friend eat a meal alone, or in the workplace such as visiting a restaurant to conduct an informal business meeting. Therefore, even when individuals are not hungry, they might be triggered to consume food to be perceived as friendly or to close a business deal. Stores will have samples of food at their deli counters and other areas of the stores to entice shoppers to purchase foods that they might not have purchased otherwise. Environmental cues can encourage individuals to eat independent of hunger or thirst. In order to keep on a diet or maintain other food consumption goals during the current

consumption period, consumers might require outside agents to act on their behalf to represent their past consumption decisions regarding the current consumption choices.

Students at schools where nutritious lunches are packed by parents could trade these foods for tastier, but less healthy food items from friends during their school lunch periods. A better result is to have students believe that they have options so that they make decisions regarding food choices with guidance from those who can make healthy food items more appealing even when consumers are subject to environmental signals toward unhealthy foods (Guthrie and Newman, 2013). Thus, the parents could pack foods that are healthier, or give the students the choices between the healthy foods that they choose. This could create a perception for the students that they want to eat the foods that were packed instead of eating those lunches because they have no other options. If the students believe that they have options in the foods that they have for lunch, they might be more inclined to stay with those food choices instead of trading those items for less healthy options.

The theoretical framework has been well established regarding various utility-maximization modeling techniques for changing tastes across time. The empirical applications of such theoretical underpinnings as simple trend terms and myopic and rational addictions models have not developed as rapidly, but have been growing in recent years. This paper intends to extend such research as it relates to habit formation and the adherence of US consumers to the USDA Dietary Guidelines for Americans (DGA) where the source of data is the IRI data from National Consumer Panel.

In this study, the nudges on consumers will be in the form of DGA that are placed on food products and in pamphlets. The remainder of the essay is organized as follows: section 2 contains the literature review; the theoretical model derivations are explained in section 3; sections 4 and 5 contain the data source details and methods; results are in section 6 and section 7 concludes.

2.2. Literature Review

USDA DGA in 2005 (DGA, 2005) discussed increasing whole grains consumption and decreasing consumption of enriched grains. These guidelines shown in Table 6 exhibit the side-by-side comparison of 100 percent whole grain wheat flour to enriched, bleached, all-purpose

white flour by calories, vitamins and mineral amounts. The whole grain wheat flour has fewer calories, and more dietary fiber, magnesium, potassium and niacin. It has less iron, thiamine, riboflavin and folate than the enriched white flour with folate being the major nutrient lacking in the whole grain flour. A chart in Appendix A-2 shows the total amount of ounce-equivalents of grains by calorie level and then the ounce-equivalents (oz-eq) by whole grains and other grains. In each calorie level, whole grains are recommended as half of the grains to be consumed.

The 2010 USDA DGA (DGA, 2011) recommended the same total oz-eq of grains and the same proportions of whole grains and other grains as the 2005 DGA. However, more emphasis was placed on reducing the enriched grains as shown in text boxes and figures. The different types of grains – whole, refined and enriched – were defined in detail in a textbox. Figure 4.1 illustrates three ways to meet the grains oz-eq recommendation using shading of bread slices. These were new ways to communicate the grains recommendation.

Whole grains were recommended, but the enriched grains were also shown to be beneficial and to have some nutrients in higher levels than the whole grains. One item that both documents have is the use of oz-eq. This might not be easily understood by the public.

Watts et al. (2011) studied the impact of DGA on American eating habits. They report that 48 percent of Americans had heard of the 2005 DGA and 71 percent had heard of the 2010 DGA. However, 48 percent know little about them, 20 percent knew a fair amount and just 3 percent know a lot about them. Thus just being aware of these guidelines did not mean that Americans knew them well.

The DGA for 2005 and 2010 differed mainly in presentation style for grains. If individuals were aware but unfamiliar with the DGA, they would not benefit from the encouragements to improve their diets. Some websites have parts of the DGA specified, but not all Americans view these sites. Even if individuals are knowledgeable of the DGA content, studies referenced in the DGA for 2010 using National Health and Nutrition Examination Survey (NHANES) 2001-2004 and 2005-2006 data showed that whole grains were under-consumed by 85 percentage points and refined grains were over-consumed by 100 percentage points compared

to the guideline amounts. Thus, the goal has not been reached by the USDA for recommended levels of whole grain consumption by Americans.

2.2.1. Myopic Habit Formation Models

In myopic habit models, the consumer has a previous experience with the food in question. The current period consumption level is based on the consumption of the individual in the past who does not anticipate consumption levels in the future, but has allowed past consumption to guide the current amount of consumption (Becker, Grossman and Murphy, 1994). This consumer is possibly naïve (Strotz, 1955) because it is very likely that future consumption of the product will occur. This was not a one-time purchase, so there could be a psychological stock of habits built for the product that leads the individual to desire the product more in the next period than in the current one. Myopic preferences are formed solely based on past consumption experiences that can increase the consumption in the current period, but with no anticipation for desire of the goods in the future period. This is a shortsighted preference formation compared to farsighted preferences that are needed for rational habit development.

2.2.2. Rational Habit Formation Models

Farsighted individuals who allow past and anticipated future consumption of a good to influence marginal rates of substitution for current consumption are said to have rational habits formed for the product (Becker, Grossman and Murphy, 1994). Rational habit formation has been studied and a firm theoretical foundation has been laid by researchers. Becker (1996) established addiction as the strongest degree of habit formation. In an earlier paper, Becker and Murphy (1988) developed a rational addiction model. Becker, Grossman and Murphy (1994) derived and estimated a rational addiction model of current cigarette consumption on a one period lag and lead consumption values assuming a quadratic utility function in addition to a myopic model that only considers lagged cigarette consumption. These models were estimated to determine if rational addiction is a societal phenomenon for cigarettes or simply myopic addiction.

Zhen and Wohlgenant (2005) developed a theoretical model for consumer demand for meat using an underlying quadratic utility function. The model was used to study how food safety has affected consumer demand for meat. The empirical analysis of this theoretical work is

found in the Zhen and Wohlgenant (2006) paper. Each of these models assumes that rational addictions have constant discounting. Therefore, time preference is always the same and is not determined endogenously.

A different assumption is that individuals have present bias for certain goods, so time preference is inconsistent and this time inconsistency can be modeled in various ways as shown by Frederick, Lowenstein and O'Donoghue (2002). Present bias is where consumers prefer consumption in the current period more heavily than in future periods. With present bias, the preferences are time-inconsistent and timing in decision-making matters for optimizing lifetime utility. The quasi-hyperbolic discounting model (Laibson, 1997) will be used to estimate a time-inconsistent model.

2.2.3. Nudges

Outside agents such as government agencies or healthcare professionals can urge decisions for foods that benefit positive health outcomes via economic and social incentives and information dissemination where the consumers can still make choices. These are referred to as nudges and are less intrusive and heavy handed than mandates to achieve certain health outcomes or punishing negative dietary choices as with taxing sugar-sweetened beverages. When individuals believe that they do not have choices across a set of food options, they can determine to eat less healthy foods in the future when these choices are possible. An example is a project at Cornell University's Center for Behavioral Economics on smarter strategies for lunch choices that demonstrated ways for students to improve their decisions when they are given choices instead of punishments for choosing low nutrient dense foods (Guthrie and Newman, 2013).

Mancino and Kuchler (2012) measured the effects of the 2005 DGA on whole-grain bread consumption using 2003-2007 Nielsen Homescan panel data. In this study, they assumed that information nudges from the 2005 DGA were common knowledge and found that the DGA increased whole grain bread consumption was independent from the decreased product price.

Mancino and Kuchler (2012) divided their dataset into pre-2005 DGA and post-2005 DGA categories and estimated two Almost Ideal Demand System (AIDS) models where each had three goods: refined grain, multigrain and whole-grain bread products. Cross-price elasticities

for whole grain and refined grain bread became statistically significant in the AIDS estimation using data after 2005 DGA. This was evidence that consumers were more inclined to treat whole grains as substitutes for refined grains. The change in consumption patterns might not have been a direct result of the 2005 DGA, but based on market changes or news coverage regarding the benefits of whole grains. Assuming the awareness of the DGA was common knowledge, the nudges were essential in driving the conclusion in their study.

Experimental design is often expensive to implement and yields results that can often be applied only to very specific subsections of the US population. By using national household panel data and the USDA DGA, Mancino and Kuchler (2012) generated conclusions that were applicable nationwide. The current chapter will incorporate the behavioral economics technique used by Mancino and Kuchler (2012) where the 2010 USDA DGA will serves as a nudge and the dividing line for the pre- and post- datasets from Information Resources, Inc. (IRI) to form conclusions about the effects of the 2010 DGA on consumer habit formation for whole grains and fruit juices.

2.3 Model

2.3.1 Discount Models

In the quasi-hyperbolic discount model, all future periods have a high discount rate (lower discount factor, β) from the perspective of the current period. Adjacent periods revert to a lower discount rate (higher discount factor, δ) that would be the same as the condition in the exponential discount model that assumes time-consistent preferences. As time passes and the near future becomes the current period, the consumer places more weight on the current consumption level and less on the future consumption levels.

The quasi-hyperbolic discount function reduces to the exponential discount function when $\beta = 1$. The reduced form parameters of the two models can be estimated to test for present bias versus time consistent preferences that require estimation of some of the structural parameters (namely β and δ). The β estimate has been empirically lower than the δ estimate where δ has been estimated closer to 1 and β has been estimated to be closer to two-thirds (Laibson, 1997).

In order to determine if the estimated δ was sustained each month or if the time inconsistent model was a better fit with the data where the β discount factor was included to account for present bias, at least two lead terms were needed. Therefore, the utility maximization equation that was modeled after the BGM paper included two lags. As with the Becker, Grossman and Murphy (1994) model (hereafter referred to as the BGM model), the current study used a quadratic utility function that was time-inseparable. This allowed, but did not require, intertemporal impacts in the utility function such as with habit formation. The model derivation also includes quasi-hyperbolic discounting.

2.3.2 Model Derivations

(1) $$\max U = u_{it}(y_{it}, c_{it}, c_{i,t-1}, c_{i,t-2}, e_{it}) + E_t\left[\beta \sum_{\tau=1}^{\infty} \delta^{\tau} u_{it}(y_{i,t+\tau}, c_{i,t+\tau}, c_{i,t+\tau-1}, c_{i,t+\tau-2}, e_{i,t+\tau})\right]$$

$$\text{subject to } (1+r)A_0 = \sum_{\tau=1}^{\infty} \left(\frac{1}{1+r}\right)^{\tau} (y_{t+\tau} + p_{t+\tau} c_{t+\tau}), \text{ where } \left(\frac{1}{1+r}\right)^{\tau} = \gamma^{\tau}$$

(2) $$\frac{\partial \mathcal{L}}{\partial y_{it}}: \ \frac{\partial U}{\partial y_{it}} = \lambda$$

(3) $$\frac{\partial \mathcal{L}}{\partial C_{it}}: \ \frac{\partial U}{\partial C_{it}} + \beta\delta \left(\frac{\partial U}{\partial C_{i,t+1}}\right) + \beta\delta^2 \left(\frac{\partial U}{\partial C_{i,t+2}}\right) = \lambda P_{it}$$

Similar to the BGM model, solving the top FOC for y_{it} and then generating the equilibrium equations for $y_{i,t+1}$ and $y_{i,t+2}$ allowed substitution into the c_{it} FOC. Solving that equation for c_{it} generated the linear difference equation to be estimated that had possible habit formation in addition to time-inconsistent preferences.

(4) $$c_{it} = \theta_1 c_{i,t-1} + \theta_2 c_{i,t-2} + \beta\delta\theta_3 c_{i,t+1} + \beta\delta^2\theta_2 c_{i,t+2} + \alpha p_{it} - \theta_0 + \underbrace{\xi_1 e_{it} + \beta\delta\xi_2 e_{i,t+1} + \beta\delta^2\xi_3 e_{i,t+2} + \varepsilon_{it}}_{error\ term}$$

where c_{it} is the current period of consumption of a given commodity, $c_{i,t-1}$ is the 1-month lag of aggregated household commodity consumption, $c_{i,t-2}$ is the 2-month lag of aggregated household commodity consumption, $c_{i,t+1}$ is the 1-month lead aggregated household commodity consumption, $c_{i,t+1}$ is the 2-month lead aggregated household commodity consumption, p_t is the commodity price and the e_{it}, $e_{i,t+1}$ and $e_{i,t+2}$ terms are the unobservable current and future life cycle variables that are correlates of consumption.

Only $\beta\delta^2$ was identified. When trying to determine a value for β, this study used the reduced form model where $\beta\delta^2 = \frac{\Upsilon_4}{\Upsilon_2}$. Given that δ^2 was not identified, the value from Laibson (1997) was used. Thus δ = 0.99 was used whenever an estimate of β was calculated. When δ was estimated,

(5) $$c_{it} = \Upsilon_1 c_{i,t-1} + \Upsilon_2 c_{i,t-2} + \Upsilon_3 c_{i,t+1} + \Upsilon_4 c_{i,t+2} + \alpha p_{it} - \theta_0 + u_{it}$$

$$\text{where } u_{it} = \xi_1 e_{it} + \beta\delta\xi_2 e_{i,t+1} + \beta\delta^2\xi_3 e_{i,t+2} + \varepsilon_{it}$$

Assuming time consistent preferences under rational habits where β = 1, we have:

(6) $$\max U = u_{it}(y_{it}, c_{it}, c_{i,t-1}, c_{i,t-2}, e_{it}) + E_t\left[\sum_{\tau=1}^{\infty} \delta^\tau u_{it}(y_{i,t+\tau}, c_{i,t+\tau}, c_{i,t+\tau-1}, c_{i,t+\tau-2}, e_{i,t+\tau})\right]$$

$$\text{subject to } (1+r)A_0 = \sum_{\tau=1}^{\infty} \left(\frac{1}{1+r}\right)^\tau \left(y_{i,t+\tau} + p_{i,t+\tau} c_{i,t+\tau}\right), \text{ where } \left(\frac{1}{1+r}\right)^\tau = \gamma^\tau$$

(7) $$\frac{\partial \mathcal{L}}{\partial y_{it}}: \ \frac{\partial U}{\partial y_{it}} = \lambda$$

(8) $$\frac{\partial \mathcal{L}}{\partial c_{it}}: \ \frac{\partial U}{\partial c_{it}} + \delta\left(\frac{\partial U}{\partial c_{i,t+1}}\right) + \delta^2\left(\frac{\partial U}{\partial c_{i,t+2}}\right) = \lambda p_{it}$$

(9) $$c_{it} = \theta_1 c_{i,t-1} + \theta_2 c_{i,t-2} + \delta\theta_1 c_{i,t+1} + \delta^2\theta_2 c_{i,t+2} + \alpha p_{it} - \theta_0 + \underbrace{\xi_1 e_{it} + \beta\delta\xi_2 e_{i,t+1} + \beta\delta^2\xi_3 e_{i,t+2} + \varepsilon_{it}}_{error\ term}$$

where θ_1 , θ_2 , δ and δ^2 were identified.

This is the discrete form of the exponential discount model. The reduced form of the exponential discount model is:

$$(10)\mathrm{c_{it}} = \Pi_1 \mathrm{c_{i,t-1}} + \Pi_2 \mathrm{c_{i,t-2}} + \Pi_3 \mathrm{c_{i,t+1}} + \Pi_4 \mathrm{c_{i,t+2}} + \alpha \mathrm{p_{it}} - \theta_0 + u_{it}$$

where $u_{it} = \xi_1 \mathrm{e_{it}} + \delta\xi_2 \mathrm{e_{i,t+1}} + \delta^2\xi_3 \mathrm{e_{i,t+2}} + \varepsilon_{it}$

$$(11)\mathrm{c_{it}} = \theta_1 \mathrm{c_{i,t-1}} + \theta_2 \mathrm{c_{i,t-2}} + \alpha \mathrm{p_{it}} - \theta_0 + \underbrace{\xi_1 \mathrm{e_{it}} + \varepsilon_{it}}_{error\ term}$$

where θ_1 and θ_2 are identified, but not directly comparable to their counterparts in either of the rational habits models due to different structural parameters in their denominators.

$$(12)\ \mathrm{c_{it}} = \alpha_i + \theta_1 \mathrm{c_{i,t-1}} + \theta_2 \mathrm{c_{i,t-2}} + \beta\delta\theta_3 \mathrm{c_{i,t+1}} + \beta\delta^2\theta_2 \mathrm{c_{i,t+2}} + \alpha \mathrm{p_{it}} - \theta_0 + \underbrace{\xi_1 \mathrm{e_{it}} + \beta\delta\xi_2 \mathrm{e_{i,t+1}} + \beta\delta^2\xi_3 \mathrm{e_{i,t+2}} + \varepsilon_{it}}_{error\ term}$$

Taking the mean of each variable within each household demeans the data. If the structural model is demeaned, it removes the fixed effect so that it is no longer correlated with the structural model's composite error term.

$$(13)(\mathrm{c_{it}} - \overline{\mathrm{c_{it}}}) = (\alpha_i - \overline{\alpha_i}) + \theta_1(\mathrm{c_{i,t-1}} - \overline{\mathrm{c_{i,t-1}}}) + \theta_2(\mathrm{c_{i,t-2}} - \overline{\mathrm{c_{i,t-2}}}) + \beta\delta\theta_3(\mathrm{c_{i,t+1}} - \overline{\mathrm{c_{i,t+1}}}) + \beta\delta^2\theta_2(\mathrm{c_{i,t+2}} - \overline{\mathrm{c_{i,t+2}}}) + \alpha(\mathrm{p_{it}} - \overline{\mathrm{p_{it}}}) - (\theta_0 - \overline{\theta_0}) + \underbrace{\xi_1(\mathrm{e_{it}} - \overline{\mathrm{e_{it}}}) + \beta\delta\xi_2(\mathrm{e_{i,t+1}} - \overline{\mathrm{e_{i,t+1}}}) + \beta\delta^2\xi_3(\mathrm{e_{i,t+2}} - \overline{\mathrm{e_{i,t+2}}}) + (\varepsilon_{it} - \overline{\varepsilon_{it}})}_{error\ term}$$

Rewriting the model with a tilde above each demeaned variable generates the structural equation with the fixed effects removed.

$$(14)\ \tilde{\mathrm{c}}_{\mathrm{it}} = \theta_1 \tilde{\mathrm{c}}_{\mathrm{i,t-1}} + \theta_2 \tilde{\mathrm{c}}_{\mathrm{i,t-2}} + \beta\delta\theta_3 \tilde{\mathrm{c}}_{\mathrm{i,t+1}} + \beta\delta^2\theta_2 \tilde{\mathrm{c}}_{\mathrm{i,t+2}} + \alpha \tilde{\mathrm{p}}_{\mathrm{it}} + \underbrace{\xi_1 \tilde{\mathrm{e}}_{\mathrm{it}} + \beta\delta\xi_2 \tilde{\mathrm{e}}_{\mathrm{i,t+1}} + \beta\delta^2\xi_3 \tilde{\mathrm{e}}_{\mathrm{i,t+2}} + \tilde{\varepsilon}_{it}}_{error\ term}$$

The reduced form quasi-hyperbolic discount model with removed fixed effects is:

$$(15)\mathrm{c_{it}} = \Upsilon_1 \tilde{\mathrm{c}}_{\mathrm{i,t-1}} + \Upsilon_2 \tilde{\mathrm{c}}_{\mathrm{i,t-2}} + \Upsilon_3 \tilde{\mathrm{c}}_{\mathrm{i,t+1}} + \Upsilon_4 \tilde{\mathrm{c}}_{\mathrm{i,t+2}} + \varpi \tilde{\mathrm{p}}_{\mathrm{it}} + \tilde{u}_{it}$$

$$\text{where } \tilde{u}_{it} = \tilde{\mathrm{e}}_{\mathrm{it}} + \beta\delta\xi_2 \tilde{\mathrm{e}}_{\mathrm{i,t+1}} + \beta\delta^2\xi_3 \tilde{\mathrm{e}}_{\mathrm{i,t+2}} + \tilde{\varepsilon}_{it}$$

Given that the exponential model is nested in the quasi-hyperbolic model and results when β = 1, that substitution into equation (15) yields equation (16):

$$(16)\, c_{it} = \Upsilon_1 \tilde{c}_{i,t-1} + \Upsilon_2 \tilde{c}_{i,t-2} + \Upsilon_3 \tilde{c}_{i,t+1} + \Upsilon_4 \tilde{c}_{i,t+2} + \varpi \tilde{p}_{it} + \tilde{u}_{it}$$

where $\tilde{u}_{it} = \tilde{e}_{it} + \delta\xi_2 \tilde{e}_{i,t+1} + \delta^2 \xi_3 \tilde{e}_{i,t+2} + \tilde{\varepsilon}_{it}$

This is the exponential discount model with fixed effects removed. The only difference between equations (15) and (16) occur in the error terms and are not observable econometrically. Therefore, testing the structural parameters was essential to forming conclusions about the results.

2.3.3 Model Derivations for Sensitivity Analysis

Using a sensitivity test on the number of lead and lag terms, two additional leads and two additional lag terms were added to the difference equations. This was the result of adding two additional lag terms to the life cycle model that was maximized.

$$(17)\ \max U = u_{it}(Y_{it}, c_{it}, c_{i,t-1}, c_{i,t-2}, c_{i,t-3}, c_{i,t-4}, e_{it}) + E_t[\beta \textstyle\sum_{\tau=1}^{\infty} \delta^{\tau} u_{it}(Y_{i,t+\tau}, c_{i,t+\tau}, c_{i,t+\tau-1}, c_{i,t+\tau-2}, c_{i,t+\tau-3}, c_{i,t+\tau-4}, e_{i,t+\tau})]$$

$$(18)\, c_{it} = \Upsilon_1 \tilde{c}_{i,t-1} + \Upsilon_2 \tilde{c}_{i,t-2} + \Upsilon_3 \tilde{c}_{i,t+1} + \Upsilon_4 \tilde{c}_{i,t+2} + \varpi \tilde{p}_{it} - \theta_0 + \tilde{u}_{it}$$

where $\tilde{u}_{it} = \tilde{e}_{it} + \beta\delta\xi_2 \tilde{e}_{i,t+1} + \beta\delta^2 \xi_3 \tilde{e}_{i,t+2} + \tilde{\varepsilon}_{it}$

subject to $(1+r)A_0 = \sum_{\tau=1}^{\infty} \left(\frac{1}{1+r}\right)^{\tau} \left(Y_{i,t+\tau} + P_{i,t+\tau} c_{i,t+\tau}\right)$, where $\left(\frac{1}{1+r}\right)^{\tau} = \gamma^{\tau}$

$$(19)\ \frac{\partial \mathcal{L}}{\partial y_{it}}:\ \frac{\partial U}{\partial y_{it}} = \lambda$$

$$(20)\ \frac{\partial \mathcal{L}}{\partial c_{it}}:\ \frac{\partial U}{\partial c_{it}} + \beta\delta \left(\frac{\partial U}{\partial c_{i,t+1}}\right) + \beta\delta^2 \left(\frac{\partial U}{\partial c_{i,t+2}}\right) + \beta\delta^3 \left(\frac{\partial U}{\partial c_{i,t+3}}\right) + \beta\delta^4 \left(\frac{\partial U}{\partial c_{i,t+4}}\right) = \lambda P_{it}$$

$$(21)\, c_{it} = \theta_1 c_{i,t-1} + \theta_2 c_{i,t-2} + \theta_3 c_{i,t-3} + \theta_4 c_{i,t-4} + \beta\delta\theta_5 c_{i,t+1} + \beta\delta^2\theta_6 c_{i,t+2} + \beta\delta^3\theta_7 c_{i,t+3} + \beta\delta^4\theta_4 c_{i,t+4} + \alpha p_{it} - \theta_0 + \underbrace{\xi_1 e_{it} + \beta\delta\xi_2 e_{i,t+1} + \beta\delta^2\xi_3 e_{i,t+2} + \beta\delta^3\xi_4 e_{i,t+3} + \beta\delta^4\xi_5 e_{i,t+4} + \varepsilon_{it}}_{error\ term}$$

where only $\beta\delta^4$ was identified. When trying to determine a value for β, this study used the reduced form model where $\beta\delta^4 = \frac{\Upsilon_8}{\Upsilon_4}$. Given that δ^4 was not identified, the value from Laibson (1997) was used. Thus δ = 0.99 was used whenever an estimate of β was calculated.

The reduced form for this quasi-hyperbolic discount model with four consumption leads and four consumption lags is:

$$(22)\, c_{it} = \Upsilon_1 c_{i,t-1} + \Upsilon_2 c_{i,t-2} + \Upsilon_3 c_{i,t-3} + \Upsilon_4 c_{i,t-4} + \Upsilon_5 c_{i,t+1} + \Upsilon_6 c_{i,t+2} + \Upsilon_7 c_{i,t+3} + \Upsilon_8 c_{i,t+4} + \alpha p_{it} - \theta_0 + u_{it}$$

where $u_{it} = \xi_1 e_{it} + \beta\delta\xi_2 e_{i,t+1} + \beta\delta^2\xi_3 e_{i,t+2} + \beta\delta^3\xi_4 e_{i,t+3} + \beta\delta^4\xi_5 e_{i,t+4} + \varepsilon_{it}$

Assuming time consistent preferences under rational habits where β = 1, the resulting model is:

$$(23)\ \max U = u_{it}(Y_{it}, c_{it}, c_{i,t-1}, c_{i,t-2}, c_{i,t-3}, c_{i,t-4}, e_{it}) + E_t[\textstyle\sum_{\tau=1}^{\infty} \delta^\tau u_{it}(Y_{i,t+\tau}, c_{i,t+\tau}, c_{i,t+\tau-1}, c_{i,t+\tau-2}, c_{i,t+\tau-3}, c_{i,t+\tau-4}, e_{i,t+\tau})]$$

subject to $(1+r)A_0 = \sum_{\tau=1}^{\infty} \left(\frac{1}{1+r}\right)^\tau (Y_{t+\tau} + P_{t+\tau} c_{t+\tau})$, where $\left(\frac{1}{1+r}\right)^\tau = \gamma^\tau$

$$(24)\ \frac{\partial \mathcal{L}}{\partial y_{it}}:\ \frac{\partial U}{\partial y_{it}} = \lambda$$

$$(25)\ \frac{\partial \mathcal{L}}{\partial c_{it}}:\ \frac{\partial U}{\partial c_{it}} + \delta\left(\frac{\partial U}{\partial c_{i,t+1}}\right) + \delta^2\left(\frac{\partial U}{\partial c_{i,t+2}}\right) + \delta^3\left(\frac{\partial U}{\partial c_{i,t+1}}\right) + \delta^4\left(\frac{\partial U}{\partial c_{i,t+2}}\right) = \lambda P_{it}$$

$$(26)\, c_{it} = \theta_1 c_{i,t-1} + \theta_2 c_{i,t-2} + \theta_3 c_{i,t-3} + \theta_4 c_{i,t-4} + \delta\theta_1 c_{i,t+1} + \delta^2\theta_2 c_{i,t+2} + \delta^3\theta_1 c_{i,t+3} + \delta^4\theta_2 c_{i,t+4} + \alpha p_{it} - \theta_0 + \underbrace{\xi_1 e_{it} + \delta\xi_2 e_{i,t+1} + \delta^2\xi_3 e_{i,t+2} + \delta^3\xi_4 e_{i,t+3} + \delta^4\xi_5 e_{i,t+4} + \varepsilon_{it}}_{error\ term}$$

where θ_1, θ_2 , δ, δ^2, δ^3 and δ^4 are identified.

This is the discrete form of the exponential discount model. The reduced form of the exponential discount model is:

$$(27)\, c_{it} = \Pi_1 c_{i,t-1} + \Pi_2 c_{i,t-2} + \Pi_3 c_{i,t+1} + \Pi_4 c_{i,t+2} + \alpha p_{it} - \theta_0 + u_{it}$$

where $u_{it} = \xi_1 e_{it} + \delta\xi_2 e_{i,t+1} + \delta^2\xi_3 e_{i,t+2} + \delta^3\xi_4 e_{i,t+3} + \delta^4\xi_5 e_{i,t+4} + \varepsilon_{it}$

Assuming myopic habits, $\beta = 0$, or equivalently $\delta = \delta^2 = \delta^3 = \delta^4 = 0$ jointly, which by substitution into either the quasi-hyperbolic discount model or the exponential discount model respectively gives:

$$(28) c_{it} = \theta_1 c_{i,t-1} + \theta_2 c_{i,t-2} + \theta_1 c_{i,t-3} + \theta_2 c_{i,t-4} + \alpha p_{it} - \theta_0 + \underbrace{\xi_1 e_{it} + \varepsilon_{it}}_{\substack{error \\ term}}$$

where θ_1 and θ_2 are identified, but not directly comparable to their counterparts in either of the rational habits models due to different structural parameters in their denominators.

$$(29)\ c_{it} = \alpha_i + \theta_1 c_{i,t-1} + \theta_2 c_{i,t-2} + \beta\delta\theta_3 c_{i,t+1} + \beta\delta^2\theta_2 c_{i,t+2} + \alpha p_{it} - \theta_0 + \underbrace{\xi_1 e_{it} + \beta\delta\xi_2 e_{i,t+1} + \beta\delta^2\xi_3 e_{i,t+2} + \varepsilon_{it}}_{in\ the\ error\ term}$$

Taking the mean of each variable within each household demeaned the data. If the structural model is demeaned, it removes the fixed effect so that it is no longer correlated with the structural model's composite error term.

$$(30)\ (c_{it} - \overline{c_{it}}) = (\alpha_i - \overline{\alpha_i}) + \theta_1\left(c_{i,t-1} - \overline{c_{i,t-1}}\right) + \theta_2\left(c_{i,t-2} - \overline{c_{i,t-2}}\right) + \beta\delta\theta_3\left(c_{i,t+1} - \overline{c_{i,t+1}}\right) + \beta\delta^2\theta_2\left(c_{i,t+2} - \overline{c_{i,t+2}}\right) + \alpha(p_{it} - \overline{p_{it}}) - (\theta_0 - \overline{\theta_0}) + \underbrace{\xi_1(e_{it} - \overline{e_{it}}) + \beta\delta\xi_2\left(e_{i,t+1} - \overline{e_{i,t+1}}\right) + \beta\delta^2\xi_3\left(e_{i,t+2} - \overline{e_{i,t+2}}\right) + (\varepsilon_{it} - \overline{\varepsilon_{it}})}_{error\ term}$$

Rewriting the model with a tilde above each demeaned variable generates the structural equation with the fixed effects removed.

$$(31)\ \tilde{c}_{it} = \theta_1\tilde{c}_{i,t-1} + \theta_2\tilde{c}_{i,t-2} + \theta_3\tilde{c}_{i,t-3} + \theta_4\tilde{c}_{i,t-4} + \beta\delta\theta_5\tilde{c}_{i,t+1} + \beta\delta^2\theta_6\tilde{c}_{i,t+2} + \beta\delta^3\theta_7\tilde{c}_{i,t+1} + \beta\delta^4\theta_4\tilde{c}_{i,t+2} + \alpha\tilde{p}_{it} + \underbrace{\xi_1\tilde{e}_{it} + \beta\delta\xi_2\tilde{e}_{i,t+1} + \beta\delta^2\xi_3\tilde{e}_{i,t+2} + \beta\delta^3\xi_4\tilde{e}_{i,t+3} + \beta\delta^4\xi_5\tilde{e}_{i,t+4} + \tilde{\varepsilon}_{it}}_{in\ the\ error\ term}$$

The reduced form quasi-hyperbolic discount model with removed fixed effects is:

$$(32) c_{it} = \Upsilon_1\tilde{c}_{i,t-1} + \Upsilon_2\tilde{c}_{i,t-2} + \Upsilon_1\tilde{c}_{i,t-3} + \Upsilon_2\tilde{c}_{i,t-4} + \Upsilon_3\tilde{c}_{i,t+1} + \Upsilon_4\tilde{c}_{i,t+2} + \Upsilon_3\tilde{c}_{i,t+3} + \Upsilon_4\tilde{c}_{i,t+4} + \varpi\tilde{p}_{it} + \tilde{u}_{it}$$

where $\tilde{u}_{it} = \xi_1\tilde{e}_{it} + \beta\delta\xi_2\tilde{e}_{i,t+1} + \beta\delta^2\xi_3\tilde{e}_{i,t+2} + \beta\delta^3\xi_4\tilde{e}_{i,t+3} + \beta\delta^4\xi_5\tilde{e}_{i,t+4} + \tilde{\varepsilon}_{it}$

When β = 1, the quasi-hyperbolic discount model reduced to the exponential discount model. With this substitution, equation (32) became:

$$(33)\, c_{it} = \Upsilon_1\tilde{c}_{i,t-1} + \Upsilon_2\tilde{c}_{i,t-2} + \Upsilon_1\tilde{c}_{i,t-3} + \Upsilon_2\tilde{c}_{i,t-4} + \Upsilon_3\tilde{c}_{i,t+1} + \Upsilon_4\tilde{c}_{i,t+2} + \Upsilon_3\tilde{c}_{i,t+3} + \Upsilon_4\tilde{c}_{i,t+4} + \varpi\tilde{p}_{it} + \tilde{u}_{it}$$

where $\tilde{u}_{it} = \xi_1\tilde{e}_{it} + \delta\xi_2\tilde{e}_{i,t+1} + \delta^2\xi_3\tilde{e}_{i,t+2} + \delta^3\xi_4\tilde{e}_{i,t+3} + \delta^4\xi_5\tilde{e}_{i,t+4} + \tilde{\varepsilon}_{it}$

Equation (33) had the fixed effect removed from the exponential discount model. If testing indicated the presence of fixed effects for the models with 4 lead and 4 lag terms, equations (32) and (33) would be used to estimate the results and draw conclusions.

2.4 Data

This study used IRI Consumer Panel Network household data aggregated monthly from years 2008 and 2012, which are before and after the publication of the 2010 USDA Dietary Guidelines for Americans. Each respondent is the household head who gives answers that incorporate the purchases and other relevant consumer demographics of all individuals in their respective home. In this study, the focus is on long-term habits for increasing consumption of whole grains (as consumed in cereals, bread, rolls, pastas, rice, frozen and ready-to-cook products, flour and mixes) and decreasing or avoiding sugar-sweetened beverages (such as non-diet sodas, sports drinks and fruit drinks). Thus, the variables of interest are fruit juices and whole grains.

The instruments used for the endogenous quantities purchased were two commodity price lead and two lag terms, race, Hispanic status, education, marital status, household size, household income and month of purchase. Quarterly food-at-home prices database (QFAHPD) categories from the Nielsen Homescan data were used to select UPC's from the IRI food-at-home beverages and bakery items files for purchases at food retail outlets to determine sugar-sweetened beverages, fruit juices, whole grains and other grains products. Two 1-month lead and two 1-month lag terms were created for households by month for six months of 2008 and also for 2012. These years were chosen because they were before and after the publication release data of the 2010 DGA on January 31, 2011. Using 2012 data

allowed time for the information from the 2010 DGA to disseminate to the public. The year 2008 was chosen to be far from the discussion in the news about the upcoming 2010 DGA. When conducting sensitivity analysis, four 1-month lead terms and four 1-month lag terms were created to be used in the rational habit formation models.

These years of data were used to test any impact of the DGA on public habits and time preference for grains, fruit juices and sugar-sweetened beverages. The unit of observation was household purchases on a daily basis where shopping trips occur and are reported via scanner devices that were given to the households. The unit of analysis is the household purchases of the commodity goods in ounces aggregated per month. Prices per ounce of those commodities were averages by household per month. Households were on one dataset and the purchases of each commodity were separately merged into household dataset containing demographic data. Panelists were chosen if they were members for the entire calendar year. The years 2008 and 2012 were treated separately.

The quantities purchased were aggregated by household per month without regard for presence of zero purchases. However, due to the nature of the dataset, only nonzero amounts were recorded. By selecting the commodity, only positive values were available. Missing data was not imputed with a zero value because those households could have had purchases that were not recorded. The result was selection bias where households would only be chosen if they purchased the commodity being modeled.

Once the quantities were aggregated and the prices averaged, SAS proc panel was used to create eleven lags of those variables. Two lead and lags would be needed, so time periods of data would be lost. The fourth lag was used as the current period with the second and third lags used as the 1- and 2-month lead variables and the fifth and sixth lags were used as the 1- and 2-month lag variables. Six observations per household remained after lagging so the estimation with panel data each had T=6.

In order to measure several periods of habit strength and many periods of discounting of future periods of the selected commodities many lead and lag terms would be needed. If a household had zero purchases all the nonzero observations for that household would be

eliminated from the regression model. Thus, to regress a model with many lead and lags to determine long-term habit formation and time preference requires restricting the data to those who have nonzero purchases of a given commodity. These deleted households comprised 80, 62, 66, and 58 percent of the data respectively for whole grains, other grains, fruit juices and sugar-sweetened beverages using 2008 IRI data. Using 2012 IRI data, those percentages were 87, 68, 81 and 73 respectively. Therefore, households from the IRI panel were selected if they had at least one daily purchase per month for twelve months. The model results would then be specific to those households with long-term purchases even if those purchases were only a very few ounces per month.

Table 2-1 shows the total monthly commodity quantities purchased per household in ounces averaged across households. Also included in the table are commodity average prices paid ($/oz) averaged across households by month. Average prices and quantities did not change much within a given commodity from 2008 to 2012. The average price per ounce for fruit juices was almost twice the average price per ounce paid for sugar-sweetened beverages in both years. The average aggregated commodity quantity purchased of fruit juices was approximately one-third of that amount for sugar-sweetened beverages. Therefore, households on average paid more per ounce and received less per ounce of fruit juices compared to sugar-sweetened beverages. The difference between the monthly aggregated quantities purchased of whole and other grains was much closer with whole grains at about 80 percent of the other grains monthly average aggregated purchases. Average prices per ounce for whole and other grains were almost the same at 12 cents per ounce for whole grains and 11 cents per ounce for other grains in 2008. In 2012, the average price/oz paid was 12 cents for both whole and other grains.

Table 2-1. Commodity Averages, 2008 and 2012 IRI data

Commodities	Avg aggr quant/oz 2008	Avg prices per oz 2008	Avg aggr quant/oz 2012	Avg prices per oz 2012
Whole Grains	73.74	0.12	72.38	0.12
Other Grains	103.54	0.11	94.16	0.12
Fruit Juices	286.66	0.06	250.35	0.07
SS Beverages	912.30	0.04	812.85	0.04

Notes: IRI data 2008 and 2012, aggregated monthly averages

2.5 Methods

Testing the effect of Dietary Guidelines for Americans as a nudging treatment to improve eating habits was conducted using a pre-and post-2010 split of the IRI data following Mancino and Kuchler (MK) (2012). In their paper, MK use the 2005 DGA as a nudge assumed to be known by the public and a treatment used to divide the data as before and after the 2005 publication of the USDA DGA. MK show that consumption changes can be affected by both health information from the DGA and price changes emanating from increased production of whole grain bread in anticipation of consumers' increased desire for healthier bread loaves from the influence of the 2005 DGA publication.

Consumption in the current time period depends on unobservables that are correlated with future consumption such as the expectation of future prices. These unobservables are captured in the error term of the current period. Since habits in consumption of particular goods indicates that future consumption levels are correlated with current consumption of those goods, the rational habit formation model has future consumption correlated with expectation of future prices that was mentioned to be captured in the current period's error term. The same relationship can be argued to exist in past time periods for goods consumed habitually. Thus the habit model has current consumption correlated with past, present and future unobservable factors in the respective error terms among other factors that cause consumption to be endogenous. The error terms are serially correlated due to price expectation of previous periods. Serial correlation and endogeneity of the consumption variables with habits leads to biased and inconsistent OLS coefficient estimates. The estimation of a rational habits model requires at least consistent estimation of coefficients of lagged and lead consumption variables.

Many estimation procedures can be used to achieve this goal. Two-stage and three-stage least squares and generalized method of moments estimation are some of the possible procedures yielding consistent estimation in the presence of endogeneity. Becker, Grossman and Murphy (1994) used a two stage least squares (2SLS) procedure in their rational cigarette addiction model. This study is largely modeled after the BGM paper. However, Generalized Method of Moments (GMM) and Arellano and Bond (1991) will be used to estimate the current rational habit model due to the panel structure of the data. BGM estimated the myopic model using OLS, thus, the current paper will follow their estimation method.

The BGM rational habit model was adjusted in this work to include quasi-hyperbolic discounting as introduced in Laibson (1997) where the resulting endogeneity was addressed using GMM with lead and lagged prices used as instrumental variables. Exponential discounting was mentioned earlier as representing time consistent (time-stable) preferences. As the future time period approached the current period discounting between two adjacent periods remained constant. For goods with preferences that could be altered under the influence of temporary emotional states such as food and beverages, impatience could lead to increased discounting between the current period and adjacent future periods. These preferences are time-inconsistent and are often modeled by hyperbolic discounting utility functions. Laibson (1997) used a linearized form of the hyperbolic discounting function, which is loosely related to the nonlinear form because as per period discounting increases the closer one is to the current time period. In fact, only the adjacent period to the current one is discounted at a higher rate. All future periods thereafter are discounted at a constant rate as with exponential discounting models.

In order to test the models to determine the time preference that best fits the data, the δ value for each commodity was estimated using equation (10). Since β was not identified in the quasi-hyperbolic discount models, only a conditional value could be calculated from the $\beta\delta^2$ expression after substituting δ = 0.99 into the expression. It was tested to determine if it was statistically different from 1 as the nested case and null hypothesis for exponential discounting.

Those β values are reported in the Tables of the Results section. The reported δ and higher order δ values of those tables were not used to estimate the reported β values. No β values could be directly reported since they were not identified.

As the discount factors were identified in the exponential discounting model, it was used to test for the case of myopic habit formation. After estimation of the exponential model, the estimates for δ and δ^2 were tested using the testnl postestimation nonlinear hypothesis testing command in Stata to determine if they were jointly significantly different from 0. If they were not, then the myopic habit formation model was directly estimated using OLS as was performed by BGM.

The lead and lag variables in the optimization equation from the first order condition are each correlated with the unobservable demand shifters at time=t. It is also possible that these unobservable variables in the error term were serially correlated. These conditions lead to possible endogenous regressors. The price variables in the contemporaneous periods of the lead and lag quantities are correlated with those respective quantities purchased. They could be valid instruments for their contemporaneous quantities variables, but future prices could have posed an errors-in-variables problem that introduced measurement error into the rational habits models. This is because future prices are not known with certainty by consumers. Thus, the use of future prices as instruments for commodity consumption variables creates the classical errors-in-variables problem that BGM also incurred. BGM used future prices as instruments for their rational addictions model. They also did a robustness check by removing lead prices from the instruments of their rational addictions model. The same two sets of instruments were used in this study to check the sensitivity of the estimates to the use of lead prices as instruments.

Fixed effects estimators are unbiased when the regressors are exogenous and the errors are uncorrelated. Two of the commodity-years were not found to have endogenous regressors, but the models regress several leads and lags of adjacent periods of quantities purchased on the current period. The rational habits model is dynamic with lead and lag consumption terms. The

sample size approaches infinity, but T is fixed at 4 to 6 months due to lost time periods from lagging the commodity group dependent variables. Thus, application of the fixed effects estimator on these models could have resulted in Nickell bias (Nickell, 1981). This bias occurs when the individual effects are correlated with the regressors as when the individual effects stem from omitted variable bias. If the sample size N and time period T both approach infinity, Nickell (1981) demonstrated that the bias approaches zero. If T if fixed, the bias does not approach zero, but remains a problem in the regression estimates. Therefore, the results from the OLS model were used when the regressors were determined to be exogenous.

As an additional test of time preferences in the rational habit formation model, the quasi-hyperbolic discount model was estimated where $\beta\delta^2 = \frac{\Upsilon_4}{\Upsilon_2}$ was identified, thus $\beta = \frac{\Upsilon_4}{\Upsilon_2}\frac{1}{(0.99)^2}$ was estimated from equation (5) using δ = 0.99 as per the Laibson (1997) article for the models with two leads and two lag terms. In the models with four leads and four lag terms, $\beta\delta^4 = \frac{\Upsilon_8}{\Upsilon_4}$ was identified, thus $\beta\delta^4 = \frac{\Upsilon_8}{\Upsilon_4}$ was used to estimate $= \frac{\Upsilon_8}{\Upsilon_4}\frac{1}{(0.99)^4}$. $\beta = \frac{\Upsilon_4}{\Upsilon_2}\frac{1}{(0.99)^2}$ and $\beta = \frac{\Upsilon_8}{\Upsilon_4}\frac{1}{(0.99)^4}$ were each tested depending on the number of lead and lag terms in the rational habits model for grains commodities against 1 as a further test to determine if the null hypothesis of the nested exponential discount model for β = 1 was supported or rejected. Also, these respective β values were tested across years for the grains commodities as a way to determine if time preferences changed after the publishing of the 2010 DGA.

2.6 Results

The four commodities were tested for endogeneity of the lead and lag terms. For each commodity, the lead and lag terms were jointly tested for endogeneity using a robust regression-based version of the Wu-Hausman test of endogeneity. This test was robust to autocorrelated, heteroskedastic and clustered errors. The results of those tests are shown in

Table **2-2**. The whole grains and other grains commodities did not have evidence to reject the null hypothesis of efficient estimation using OLS. Fruit juices and sugar-sweetened beverages did have supporting evidence to reject the null hypothesis of efficient OLS for the

consistent estimation from GMM under the alternative hypothesis. Rational habits models were estimated for fruit juices and sugar-sweetened beverages using instruments with and without future prices. The estimates without future prices were statistically unstable and produced negative R^2 values. Therefore the results below only include estimates from instruments with past and future prices and with demographic variables.

Table 2-2 Endogeneity test results

Commodities	Year	F stat	P-value
Whole Grains	2008	0.92	0.45
	2012	1.18	0.32
Other Grains	2008	1.05	0.38
	2012	1.50	0.20
Fruit Juices	2008	11.03	0.00
	2012	8.25	0.00
SS Beverages	2008	6.65	0.00
	2012	15.56	0.00

Notes: F-statistics were generated from the robust regression-based test for endogeneity that was robust to heteroskedasiticity, autocorrelation and clustering of errors.

Each commodity rejected the null hypothesis of random effects for the alternative hypothesis of fixed effects correlation with the idiosyncratic errors. If there were only a few time periods and a lagged dependent variable with serial correlation and large N with a fixed effect correlation with the errors, the model might suffer from Nickell bias (Nickell, 1981). Thus, the model would have negatively biased estimates for the lagged dependent variables. This was the case with each of the models when estimating equations 15, 16, 32 and 33. The habit correlation coefficients of the lagged consumption variables were negative for each commodity and the R^2 values were negative for each commodity. Therefore, results from those models were excluded in this study.

Several habit formation models were estimated under different time preference and forward- and backward-looking assumptions for four commodity groups to determine the model most supported by the data per commodity group. The sugar-sweetened beverages β estimate using 2008 data was equal to 0.38 as shown in Table 2-3, which was not statistically

significantly different from 1. Also, the square of the δ estimate was not significantly different from the δ^2 estimate of the exponential discounting rational habits model. However, the test of $\delta = \delta^2 = 0$ for the nested case of myopic preferences failed to be rejected. Overall, these results supported a myopic habit formation for sugar-sweetened beverages.

Full model estimation results for Chapter 2 are located in the Appendix tables. In Table 2-4, the results of the myopic habits models for sugar-sweetened beverages showed a statistical difference across years for the 1-month lag habit coefficient. There was a statistical lowering of the habit coefficient for sugar-sweetened beverages after the 2010 DGA based on the myopic habit model. Price elasticity stayed the same from 2008 to 2012 for this commodity.

The 2012 data for the same commodity failed to reject the myopic preferences null hypothesis and also failed to reject the condition of no habit formation from the past to the current time period. Consuming sugar-sweetened beverages in the past increased consumption in the current period by 10 percent. Demographic variables contributing to consumption levels in 2008 were household size, region of the country and month of purchase. In 2012, county size and race were contributing factors in addition to those factors in 2008. These factors could be targeted when deciding where to mail brochures regarding DGA in the future.

Table 2-3 Sugar-sweetened Beverages, 2008 and 2012 IRI data (2 lead and lag terms)

Structural Parameters	2008 Estimates with IV	2012 Estimates with IV
δ	0.59	0.32
δ^2	0.37	-0.04
β	0.38[a]	-0.04[a]
θ_1	0.20	0.10
θ_2	0.20	0.11
E_d	-0.03	-0.03
R^2	0.75	0.28
N	45860	40032

Notes: *p<0.10 **p<0.05, ***p<0.01, [a]Not statistically different from 1, † for at least p<0.10 for testing structural parameters across years

Table 2-4. Sugar-sweetened Beverages, 2008 and 2012 IRI data, Myopic model (2 lag terms)

Structural Parameters	2008 Estimates with IV	2012 Estimates with IV
θ_1	0.35***	0.12**
θ_2	0.28***	0.12**
E_d	-0.03***	-0.03**
R^2	0.45	0.28
N	55032	40032

Notes: *p<0.10 **p<0.05, ***p<0.01, † for at least p<0.10 for testing structural parameters across years

Using 2008 data fruit juices showed evidence of endogeneity from the robust regression-based test of endogeneity; therefore, it was necessary to estimate the reduced form model from equation (10) with instruments using GMM. Table 2-4 shows the goodness-of-fit was 0.51 with $\delta = \frac{\Pi_3}{\Pi_1} = 3.14$ and $\delta^2 = \frac{\Pi_4}{\Pi_2} = 0.89$. This test had a p = 0.62 value which was evidence of a consistent δ value from month to month of future discounting in the exponential model. The test of myopic preferences failed to reject the null of joint δ and δ^2 equal to zero, which led to estimation of a myopic model for this commodity. The myopic model results shown in Table 2-5 had a slightly lower R^2 = 0.47 and higher habit correlation coefficients. The habit coefficient was shown to be stable from month to month for two periods, so there was a sustained preference for fruit juices with no consideration for the impact of future consumption on the current period.

In 2012, consumers were shown to still have myopic preferences for fruit juices with a slightly lower, but sustained habit formed (Tables 2-5 and 2-6). Switching from the exponential

habits model to the myopic model resulted in rejection of the no-habits condition that failed to be rejected in the previous model. The habits coefficient from 1-month quantity lag was not significantly different from 0 in the rational habits model, but was significant in the myopic model. The habits coefficient from 2-month lag was significant in both models. Household size, month of purchase and whether or not the head of the household was widowed were significant influences for quantities purchased of fruit juices in the rational habits model, but only household size and widowed remained significant in the myopic model. Price elasticity was statistically significant across years for fruit juices in both the rational habits model and the myopic habits model.

Table 2-5 Fruit Juices, 2008 and 2012 IRI data (2 lead and lag terms)

Structural Parameters	2008 Estimates with IV	2012 Estimates with IV
δ	3.14	1.94
δ_2	0.89	0.71
θ_1	0.07	0.07
θ_2	0.18	0.14
E_d	-0.16	-0.23†
R^2	0.51	0.45
N	31026	11496

Notes: *p<0.10 **p<0.05, ***p<0.01, [a]Not statistically different from 1, † for at least p<0.10 for testing structural parameters across years

Table 2-6. Fruit Juices, 2008 and 2012, Myopic model (2 lag terms)

Structural Parameters	2008 Estimates with IV	2012 Estimates with IV
θ_1	0.29***	0.20***
θ_2	0.29***	0.21***
E_d	-0.16***	-0.24***†
R^2	0.47	0.42
N	31026	11496

Notes: *p<0.10 **p<0.05, ***p<0.01, † for at least p<0.10 for testing structural parameters across years

All of the grain commodities showed support for rational habit formation with no present bias (time-inconsistent preferences). Habit strength was sustained over a two-month period in a range of 0.20-0.22. There was no endogeneity detected from the Wooldridge robust

regression-based test for the grains food groups, so estimation was performed using OLS with cluster-robust errors at the household level. Other grains with 2008 data in Table 2- showed strong evidence of time consistent habit formation with $\delta = \frac{\Pi_3}{\Pi_1} = 1.01$ and $\delta^2 = \frac{\Pi_4}{\Pi_2} = 1.04$ using the estimation of the reduced form model of equation (10), where $(1.01)^2$ was not significantly different form 1.04. This indicated that a consistent discounting of the next period for each time period as required in the exponential discounting model. The conditional estimated $\beta = 1.06$ was not significantly different from 1, which also supported the exponential discounting model. Myopic preferences as the nested case of the exponential model was rejected with $p = 0.00$ and the case for sustained habit formation was supported with $\theta_1 = 0.22$ and $\theta_2 = 0.20$. Household size, county size, month of purchase, Hispanic status, race and region of the country were significant factors in determining quantities purchased containing other grains.

Table 2-7. Other Grains, 2008 and 2012 IRI data (2 lead and lag terms)

Structural Parameters	2008 Estimates	2012 Estimates
δ	1.01***	0.95***†
δ^2	1.04***	0.95***†
β	1.06***[a]	0.97***[a]†
θ_1	0.22***	0.22***
θ_2	0.20***	0.20***
E_d	-0.06***	-0.07***†
R^2	0.52	0.47
N	81174	63462

Notes: *p<0.10 **p<0.05, ***p<0.01, [a]Not statistically different from 1, † for at least p<0.10 for testing structural parameters across years

Time consistent habit formation for other grains was also shown in 2012 with $\delta = \frac{\Pi_3}{\Pi_1} = 0.95$ and $\delta^2 = \frac{\Pi_4}{\Pi_2} = 0.95$ and no significant difference between $\left(\frac{\Pi_3}{\Pi_1}\right)^2$ and $\frac{\Pi_4}{\Pi_2}$. Differences in the β values across years were significant statistically, but did not affect the conclusions from the model given that both values were not statistically different from 1 as equal support for the exponential model. This was consistent with no present bias for other grain products. The habit coefficient values with 2008 data are the same as those from the 2012 data at 0.22 and 0.20

from the 1-month and 2-month lags respectively. There was no indication of lowered habits for other grains from 2008 to 2012 as a result of the published 2010 USDA DGA. It was interesting to note that in 2012, time of purchase and all of the same demographic variables except Hispanic status contributed in deciding the amount of other grains products purchased.

There was a similar case for whole grains in Table 2-8 as with other grains for sustained rational habits with time consistent preferences. Testing the exponential discounting model for 2008 whole grains data rejected myopic habits and the case for no habits. The exponential discount factor was consistent for the next two periods where $\delta^2 = \frac{\Pi_3}{\Pi_1}$ and the estimated $\delta^2 = \frac{\Pi_4}{\Pi_2}$ were not significantly different from each other. In a separate test conditional β = 1.09 was not different from 1 statistically, so the present-bias habit formation model was not supported. The null case for exponential discounting with β = 1 was supported by this test. The month-of-purchase time effect was significant in determining quantities purchased, but no demographic variables were significant except region of the country. Habit strength was consistent from the 1- and 2-month lags of quantities purchased.

Consumers showed sustained rational habits with consistent preferences over time for whole grain products in 2012. The exponential discount factor was consistent for the 1- and 2-month lead terms for quantities purchased containing whole grains. Myopic preferences were rejected with p = 0.00 and the conditional β = 1.04 was not significantly different from 1 with p = 0.34 as the null case for an extra test regarding exponential discounting habit formation modeling. The demographic variables were not significantly correlated with purchases of whole grains in 2012 as in 2008, except for region of the country and household size. Other grains purchases were affected by many demographic characteristics, but whole grain purchases were not influenced by demographics much at all.

Table 2-8 Whole Grains, 2008 and 2012 (2 lead and lag terms)

Structural Parameters	2008 Estimates	2012 Estimates
δ	0.93***	0.94***
δ^2	1.06***	1.02***
θ_1	0.22***	0.24***
θ_2	0.20***	0.20***
E_d	-0.07***	-0.05***
R^2	0.48	0.55
N	8124	5538

Notes: *p<0.10 **p<0.05, ***p<0.01, [a]Not statistically different from 1, † for at least p<0.10 for testing structural parameters across years

Sensitivity analysis results of Table 2-9 and Table 2- for whole grains and other grains respectively indicated a more stable time consistent habit formation model for these two commodities with statistically significant β values across years. The β values for other grains were 1.50 and 1.02 for before and after the 2010 DGA and for whole grains these values were 6.19 and 1.24 in the models where 4 lead and lag terms were included. These values were statistically significantly different from each other across years 2008 and 2012. The values were lowered considerably and significantly toward more plausible values even though both sets of β values for those commodities were not statistically different from the null case of 1 for the test of exponential discounting. Also, consumers' purchases of whole and other grains showed evidence of time-consistent rational habit formation for a 9-month span in both 2008 (but tapering) and 2012 (not tapering statistically). Whole grains and other grains had statistically significant differences across years at the farthest lead and lag months for discounting and habit coefficient parameters. Thus, the models were more stable with all exponential discount factors not statistically different from 0.90 and habit coefficients from the 2-month lead and lag to the 4-month lead and lag were not statistically different from 0.10 for whole grains. The habit coefficients from these same lead and lag months for other grains were not statistically significantly different from 0.105.

Table 2-9 Whole Grains, 2008 and 2012 IRI data (4 lead and lag terms)

Structural Parameters	2008 Estimates	2012 Estimates
δ	1.04***	0.90***
δ^2	1.35***	1.17***
δ^3	0.75***	0.50***
δ^4	5.95***	1.19***†
θ_1	0.15***	0.18***
θ_2	0.10***	0.11***
θ_3	0.15***	0.12***
θ_4	0.02***	0.08***†
E_d	-0.03***	-0.04***
R^2	0.84	0.85
N	5416	3692

Notes: *p<0.10 **p<0.05, ***p<0.01, [a]Not statistically different from 1, † for at least p<0.10 for testing structural parameters across years

Table 2-10 Other Grains, 2008 and 2012 IRI data (4 lead and lag terms)

Structural Parameters	2008 Estimates	2012 Estimates
δ	1.02***	0.99***
δ^2	1.05***	1.11***
δ^3	1.13***	0.99***
δ^4	1.44***	0.98***†
θ_1	0.14***	0.15***
θ_2	0.12***	0.12***
θ_3	0.12***	0.11***
θ_4	0.07***	0.09***†
E_d	-0.04***	-0.06***†
R^2	0.84	0.82
N	54116	42308

Notes: *p<0.10 **p<0.05, ***p<0.01, [a]Not statistically different from 1, † for at least p<0.10 for testing structural parameters across years

Sugar-sweetened beverages and fruit juices did not have significant results for the rational habits models with 2-month lead and lag consumption values, so these commodities were not tested using the expanded rational habits models with four lead and four lag terms.

2.7 Conclusions

Whole grains and enriched grains were important to improving the health status of Americans. The USDA Dietary Guidelines for Americans have attempted to inform and

encourage healthy eating habits. The 2010 DGA emphasized foods to consume in increased amounts such as whole grains and those to consume in lower amounts to have healthier diets. The awareness of and deeper knowledge from these guidelines are needed if they are to be useful as nudges urging consumers to have better nutrient consumption with fewer calories. The typical American diet for whole grains was only 15 percent of the recommended 3 oz-eq of the DGA and 200 percent of the 3 oz-eq of the refined grains based on 2001-2004 and 2005-2006 NHANES data. Therefore, some consumption of whole grains occurred, but not the healthy levels indicated in past DGA.

Offering samples of whole grain products might be a way to increase consumption of this grain over the consumption of refined grains. This current study has shown that habits are formed when consuming whole and refined grains. The average individuals who purchased whole grains items monthly for the calendar year had a steady likelihood of purchases of whole grains for at least nine months in 2008 and 2012. This habit duration was strengthened from 2008 to 2012 by one 1-month lead and one 1-month lag term with no statistically significant diminishing of habit strength. However, the same result was found for the average consumer purchasing other grains in 2008 and 2012. Thus, individuals who purchased whole grain items and other grains items regularly for the month deepened their habits after the publication of the 2010 DGA. The momentum to continue purchasing based on past consumption levels was supported by the results. If individuals could switch from refined grain products to whole grains, then there might be more Americans consuming whole grains in equal proportions to other grains in order to improve health status. Both whole and other grains are recommended by the 2010 DGA, but the daily proportions are not being realized in American diets. This could be because individuals were not exposed to whole grains products or did not find them to be satisfactory across several types of grain containing items such as cereals, pastas and flour, etc. The exponential discounting model results support that individuals on average did not have commitment concerns regarding consumption choices but do have lasting habits for both types of goods.

Decreased habit strength for sugar-sweetened beverages for the 1-month lag period across years 2008 and 2012 was supported. This statistically significant decrease in habit strength was

not met with an increase in habit strength for fruit juices. This decrease in habit strength was only for food at home purchases of sugar-sweetened beverages, so an increase in consumption of this commodity could have occurred at restaurants in 2012 and at vending machines. Even though the results are restricted to food at home purchases, they were encouraging and could be furthered with more dissemination of the 2010 DGA.

2.8 Policy Implications

The 2005 and 2010 DGA have detailed content on consumption levels of whole grains and other grains that would improve the health status of Americans. However, the usage of ounce-equivalents as the measure of consumption levels for these commodities is not a term that is used on all product labels. Individuals might find it had to compare across products to know if they are consuming the recommended amounts of these products. Grams of whole grains as recommended by the USDA are mentioned on some bread products for example. Other products show the amount of ounce-equivalents of whole grains that one would receive for two slices consumed. Thus, the consumption amounts are not easily compared to determine which products offer the highest levels of whole grains per slice of bread. This could discourage other-grains consumers from beginning to consume whole grains especially if they enjoy other grains products.

There is a further complication in trying to understand ounce-equivalents (oz-eq). One ounce of ready-to-eat foods with whole grains have 16 grams of whole grains and 12 grams of other products needed to form the final product. One ounce of whole grain flour contains 16 grams of whole grains without other fillers. Thus, to consume one ounce of whole grains, an individual could eat 16 grams of whole grains directly or 28 grams from ready-to-eat foods such as prepared bread loaves as discussed by the Whole Grains Council (ref).

The USDA could offer brochures that give short lists of grams needed to meet the 3 oz-eq of whole grains and the number of bread slices that equal 3 oz-eq. These could also be stated in television and radio public service announcements and on billboards so that individuals will have more understanding when making food choices at grocery stores. In addition, it might be helpful to include grains as an item on the Nutrition Facts Panel. Seeing the content of the

whole grains on the food packages in a form that can be directly compared across brands and types of grains might offer the consumers ways to increasing whole grains and decreasing consumption percentages of other grains in their diets for the long run.

Households had myopic preferences for both sugar-sweetened beverages and fruit juices, which might indicate that buying these items was an impulse buy based on taste preferences created in the past rather than choices based on consideration of future health conditions. It might be helpful to have the USDA offer information along with the MyPlate diagram that includes some of the consequences of long term consumption of sugar-sweetened beverages so that there is more than just a comparison of one can of soda or sports drink compared to one container of fruit juices. If consumers were made aware of the long-term hazards to drinking beverages with added sugars, they might find it easier to switch to more healthy drinks.

2.9 Future Work

Combining food commodities into one rational habit formation model and separating beverage commodities into another rational habits model could be helpful in capturing cross-price elasticities to determine if households switched from the less healthy to the healthier commodity within the food and beverage categories. There might have been a more intense switch from one type of commodity for foods and separately for beverages after the 2010 DGA were published that was not observed in the current study where each commodity was estimated individually from other commodities. Wohlgenant (2012) discussed the methods needed to estimate such a multivariate rational addictions model. This model could be applied as a rational habit formation model. Other models to be estimated in the future are beverages including milk, bottled water, tea/coffee and vegetable juices. Also, fruit and vegetables as food commodities can be estimated.

Chapter 3 External Habit Formation

3.1 Motivation

Food deserts have been posited, explored and mapped for a few years, thus the topic is not new. Lower income individuals have lowered access to fresh fruits and vegetables due to their housing locations and the poor quality stores nearby. Due to lack of private modes of transportation and good stores close by, these consumers must rely on the few stores within walking distance or a short bus ride of their dwellings. Many of these people will consume fresh fruit only when at fast-food restaurants because their local grocers do not carry fresh fruits and vegetables in their produce sections.

There are social welfare programs that are intrinsically motivated to study these groups of individuals such as those who receive food stamps. Middle class and high income persons are not generally the focus of food access programs because these consumers tend not to be the focus of large health studies in particular. If poor food choices have been made, then these individuals have access to healthcare and can discuss such matters with their physicians.

I propose that with growing healthcare costs in general and the increasing time pressures on higher income earners, fast-food restaurants and ready-to-eat meals are being consumed more and more by this demographic group. The opportunity cost of cooking a meal often rises for those whose incomes are growing more quickly than food prices.

Also, fast-food meals and discount grocery stores are generally considered inferior and are kept at the lower end of the price scale. So even if food prices in the immediate areas of higher income consumers rise at the rates of their incomes, it is not that expensive to join a membership club for discount prices in bulk or simply travel a few miles to a discount grocer or fast-food establishment. Thus, it is reasonable to conclude that these individuals are making poor food choices often due to lack of time and not because of income constraints.

Interdependent consumer preferences (ICP) could affect good and bad eating habits and should be considered when analyzing the spatial effects on habit formation. ICP's result when

there are social concerns, reductions in transaction costs, and when there is a signaling effect of brand ownership on inferred attribute levels (Yang and Allenby, 2002). The interdependent consumer preferences have been referenced by various names in the literature such as "peer influences (Duncan, Haller and Portes, 1968), bandwagon effect" (Leibenstein, 1950), "neighborhood effects" (Case, 1991), "conformity" (Bernheim, 1994), "friendship effects" (Nepal, Bista and Paris, 2012), and "contagion" (Argo, Dahl and Morales, 2008).

Preferences can be formed based on the individuals' tastes and also based on the views and information from others in the consumers' peer groups. There could be conformity if the consumers believe that the others in their network of influence could provide information to benefit the original consumer or if others could help lower the transaction costs of said consumers. Individuals want to have popularity, esteem and respect (Bernheim, 1994) and these could be enhanced if the consumers based their preferences on information from peers. The functional attributes are not the only salient factors of the consumers in preference formation. The intangible benefits received from the buyers is based on shared information of products and services that help consumers avoid consuming products with net negative effects and leads them to products with net positive effects that the consumer could not discover as quickly or at all without the interdependent connection to the network of peers.

Economists tend not to focus on such individual aspects of consumer theory, but let it suffice that the consumer is fully rational and maximizes utility given budget and time constraints. Psychologists and sociologists have pursued the study of preference formation and marketing has also tried to model preference formation in attempts to predict what goods have network effects and also how the network effects will alter individual consumption for certain goods. I will incorporate such research literature from these other disciplines to motivate the models used to describe behavior that is counter to what is expected under the assumption of rational utility theory.

Consumers who have economic activity of interest make decisions about utility maximization each period where they are located. These instant decisions are affected by prices and by their tastes and what they prefer over other goods available in their areas. There could

be ways that these individuals' current consumption influences future consumption. Thus, they have a temporal component to how their consumption data are mapped and then modeled.

These data are distributed with variance over time, thus serial data (called time series data) and other variables in the past and future can influence the predicted variable. The time dimension to the data separate these data from cross sectional data where all variables are determined in the same period. The direction of time series data is necessarily from past to future; never from future to past. Serial correlation of dependent variables are lagged variables that can be one to T time periods back where T is the maximum number of periods contained in the series. The lagged variables can be correlated with each other, which introduces multicollinearity into the econometric models.

Also, the error terms can be serially correlated where the past errors covary with current and/or future time periods. The variance-covariance matrix can possibly have non-zero off-diagonal terms. This violates the Gauss-Markov assumptions which leads to incorrect variance estimation and invalid inferential statistics as does multicollinearity. As a result, the time dimension cannot be ignored when estimating consumer demand models.

Anticipated rainfall can affect current trips to the store and how much is purchased in the current period and also in the next. This is a dynamic effect. Also, there might be an average effect of two or more phenomena over time. This is a time series of cross sectional data represented as panel data. These two types of time component data have similarities, but also have nuances that require different techniques to glean the optimal estimation values and statistical inferences from regression models.

Consumers and their economic activities and influences of those activities are not only distributed over time, but also across regions. These spatial data are specifically areal (involving two dimensional space) where centroids of areal units are located by latitudinal and longitudinal coordinates in fixed regions around the US. They can affect the consumption by other consumers. This is modeled by a spatial weights matrix premultiplied by a coefficient to create a spatial lag term of the dependent variable. The other factors in the environment are

distributed over regions and can affect the main economic outcome of interest which is the dependent variable being modeled. Spatial weights are assigned to these other variables that are captured in an X matrix and premultiplied by a coefficient.

Fixed effects, socio-demographic characteristics, and other factors can influence the dependent variable independent from the spatial dimension. These factors are added to the model without spatial weighting matrices.

Consumer units are also distributed across regions as are other factors which affect consumer decision making. Environments where consumers are located can influence their tastes and preferences and timing of their purchases. Consumers can also affect tastes and preferences of other consumers based on proximity to each other. So neighbors and coworkers can change the consumption choices of individuals. These spatial relationships can be constant or vary by region (heterogeneous). The changes to spatial relationships of economic factors and other influences can be structurally stable or unstable. The structured stable changes can be heteroskedastic, spatial clusters or spatial regime switching coefficients. The structurally unstable changes are not able to be used to gain insights about the outcome of interest.

The spatial dimension violates some of the Gauss Markov assumptions. The assumption of spherical errors is often violated for example. The spatial lag violates the assumption of independent observations. Spatial heterogeneity violated the assumption of constant relationships in repeated samples. For example, different geographic regions that are sampled could have spatial autocorrelation coefficients that change compared to the originally sampled region. The goal is to use models that can accommodate these violations of the Gauss-Markov assumptions by examining the spatial distribution of the areal data and then trying to determine if the spatial effects are with the lagged dependent variable or the main effects or remaining in the disturbance terms. It is also important to explore the constancy of the spatial effects parameters and if nonconstant, to decipher the pattern of changes to those spatial parameters via model fitting under regime switching, cluster analysis or spatial expansion of the data.

I intend to measure the impact of spatial effects (both instantaneous and delayed) on habits using the IRI Homescan (IRI) data for households, retailers and medical groups. Spatial effects will be estimated using a general serial and spatial econometric model that nests both spatial autoregressive (SAR) and spatial error (SEM) models using the general to specific process for determining what type of underlying spatial data generating process exists in the geocoded IRI spatio-temporal data.

3.2 Literature Review

3.2.1 Spatial Econometrics Defined

Spatial econometrics is different from spatial statistics and geographic analysis and is a subset of spatial economics which involves location theory, spatial competition and regional and urban economics (Duranton, 2008). It is therefore necessary to understand how spatial econometrics differs from these other spatial fields, before launching into specific concerns and challenges of model estimation in two dimensions.

Spatial statistics concerns clustering and spatial dependence, but is not limited to regression analysis. Regional and urban economics are concerned with the distribution of scarce resources to individuals in given locations where there is an underpinning of economic theory with agents maximizing utility across spaces. Thus, regional and urban economics fall under the umbrella of spatial economics as does spatial econometrics, where spatial statistics is not based in economic theory, but statistical theory. However, spatial econometrics differs from regional and urban economics in the area of tools used to measure spatial relationships. Spatial dependence in spatial econometrics occurs in the direct economic components of interest and in the error terms where economic relationships cannot be directly modeled.

Spatial relationships can be measured using established models and these relationships can then be tested using local and global inference tests. It is also possible to use the estimated relationships to predict the location of spatial dependence between known points for related variables, but this will not be discussed in this paper. Anselin (2001) has written extensively on spatial econometrics and his research will be the basis of the analytical framework for this chapter of the dissertation. Anselin (2001) outlines spatial effects, spatial models, specification

tests for these models and also techniques used for spatial prediction where the latter is not germane to this paper. Two relevant spatial models for this paper are spatial autocorrelation and spatial error models. The paths to these model choices will be discussed extensively.

3.2.2 Spatial Effects

Spatial effects are phenomena that covary over regions and are either stationary or structurally unstable over space. These phenomena are captured as data in vectors or matrices and mapped over geographic regions or simply on graphs. Their plots could be white noise, thus independent and stationary (having a set structure at any section of a given region) or having patterns of dependence. The dependence patterns can be constant (stable, so a set mean and variance for example) or nonconstant (structurally unstable across areas).

The existing spatial relationship is called spatial dependence and its stable or nonconstant structure comprise the two aspects of spatial heterogeneity. Spatial effects are spatial dependence and spatial heterogeneity (Anselin, 2001). Spatial heterogeneity is the difference in values of an economic factor across an area. Spatial dependence is the influence of value on an economic factor in one location from a value in a neighboring region.

Spatial interdependence is the relationship of specified or latent economic factors represented on a spatial graph including regions of the globe. The specified economic factors could be the lagged variable of interest or other environmental factors that could influence the main effect being estimated. Latent effects cannot be directly captured in models and are often captured by the error terms. If these effects exhibit autocorrelation, then the error terms are correlated with each other across regions. A possible result is that the spatial error autoregression model suffers from omitted variable bias.

Spatial factors that can be modeled explicitly are either spatial lagged terms or spatial main effects. Spatial autocorrelation of the dependent variable is one type of spatial dependence. Another type of spatial factor is main effects spatial dependence such as with socio-demographic variables and fixed effects that are distributed over regions and influence the dependent variable. These are typically modeled as spatial lag and spatial Durbin (SDM) models where spatial weights matrices are used for the dependent variable and the other spatial

factors. These two weights matrices can be identical, or can differ in functional form such as contiguity weighting for one and distance weighting for the other.

When the model parameters change over regions or the spatial units have different error variances, then there is heterogeneity of the spatial relationship. This could take the form of nonconstant error variances which is heteroskedasticity. Nonconstant model coefficients are shown in the form of spatial regimes, random coefficient models and spatial expansion methods. Thus, there is a functional form of the error variance that changes in accordance with the region's coordinates. As the region moves from east to west, for example, the error variance could increase or decrease linearly or by polynomial factors, by transcendental or logarithmic functional form.

Heteroskedasticity could be caused by model misspecification such as with missing variables (Anselin, 1988, pp. 119-120). These model misspecifications cause unexplained variation to remain in the error terms in ways that change across space resulting in heteroskedasticity of various forms that might not have known relationship changes. This could mean that the heteroskedasticity cannot be incorporated in the regression model leading to possible biased estimates which are overly precise.

The coefficients of the model could vary by region, which means different regions could be more responsive to changes in economic or environmental factors. It could also mean that neighborhood effects could be greater depending on the region. Cities in the Northeast of the US have less of an economic downturn on average for a given level of snowfall versus Southern cities. These differences are modeled as spatial regime switches. An indicator variable for spatial subregions would model where the coefficient change(s) would occur. Varying model coefficients have different intercepts and/or slope coefficients and can be estimated using spatial regimes models.

Individuals in an area can influence each other over time and space. Their error terms can also have temporal and spatial relationships. These are measured individually as separate dimensions of interaction that have no overlap. Thus, when the spatial influence occurs, the time component is held at zero. There is no time lapse, so the spatial impact of individuals' consumption on each other is instantaneous. Similarly, the impact of consumption over time

for an individual is measured where there is no neighborhood or other geographic effect. Thus, the spatial influence of consumption over time is held at zero.

However, a spatial effect might take time to diffuse through the population and a temporal effect might impact others in the immediate area. These are two examples of space-time interactions that would require space-time analysis techniques different from those used for pure space and time impacts.

3.2.4 Specification of spatial effects

Spatial weights encapsulate the interdependence of spatial units which makes them critical to understanding the spatial component of the model and if spatial relationships actually exist. Unfortunately, these weights are needed in order to run spatial regression models which assist in discovering the types of spatial and nonspatial interdependencies of economic and environmental factors. Also, the choice of weighting matrix avoids identification problems if it is made exogenously to the system of spatial equations of the model. Therefore, the model could be incorrectly specified and the inferences misleading if the weighting matrices are selected erroneously. It might be wise to run models with several weighting matrices to see how the model fit statistics vary and which model forms have the best fit results to determine what weighting matrices should be ultimately chosen.

Deriving the spatial weight matrix requires determining the functional form of spatial influence of the neighbors that will also reflect how quickly the relationship diminishes with distance. Cliff and Ord (1969) and Kelejian and Prucha (2010) describe several functional forms for weight matrices for spatial models. There are distance-based weights, weights based on boundaries and weights based on both boundaries and distance.

Approaches to determining the spatial function of the weight matrix include using an exogenous weight matrix, using an empirically convenient functional form for the weight matrix, selecting among traditional or well-known weight functions or conducting exhaustive tests to determine mathematically the most appropriate weight function based on the process of spatial dependence.

Deriving weights geographically guarantees that they are exogenous to the model being estimated (Anselin, 2001). There are well known ways to derive these weights such as using

an MLE according to Ord (1975) and Lee (2004), IV methods as per Anselin (1980) and Kelejian and Prucha (1998, 1999) and the GMM method as exemplified by Lee (2007) and Lee and Liu (2010). Using exogenous spatial weights ensures that there are no identification problems (Anselin, 2001). Traditional weight functions have been established via trial and error by past research. Some of these functional forms are based on geographic distance or shared boundaries between points in regions i and j. The choice of weight function that is mathematically based requires that the spatial process of autocorrelation is understood. The reciprocal of this spatial process should be used as the weight function. In this paper, the spatial process will be determined and used to calculate the weight function.

It is important to determine what model to choose that determines spatial effects of consumers for whole grain bread products and fruit juices using geocoded data and if a nonspatial model is sufficient. There are two main ways to model spatial dependence: spatial error models and spatial lag models. Spatial heterogeneity can be either heteroskedastic or structurally unstable. When choosing a model, the goal is to understand the underlying DGP and capture all the variance without leaving the model with omitted variable bias. One can try to model the relationship using a specific-to-general approach (the classic method), using the general-to-specific approach as prescribed by Hendry (1979), or by using the spatial model decision rule described by Anselin (2005).

There are four paradigms for discovering the best model given the data. One can follow the classical method which is a type of forward selection from the simplest spatial model specification to the most general. This was found by Florax et al. (2003) to be the best way to find the optimal spatial model as he found by conducting simulations. Begin with the estimation of the nonspatial OLS model. This estimation can be used as the baseline for all other model estimation. Estimate the spatial lag model and the spatial error model with LM test from Anselin (1988) and then, if needed, robust LM test from Anselin et al. (1996). If either or both robust LM statistics are significant, then estimate the spatial Durbin, spatial Durbin error, and Kelejian-Prucha models for significance using the chi-squared distribution. The tests are to determine the statistical significance of the main effects spatial coefficients together with the spatial autocorrelation coefficients of the lagged dependent variable Given

that these coefficients are statistically significant, the last model to be tested is the Manski model with spatial dependence in the lagged dependent variable, the main effects and the disturbance terms. The model with the highest log-likelihood value among all remaining models is the best model to represent the underlying spatial data generating process.

One can also use the method suggested by Hendry (1979) that is a type of backward selection from the most general to the simplest specification of spatial models. This method avoids omitted variable bias and allows a quick comparison between the OLS specification of non-spatial model and the most general spatial model incorporating all possible spatial effects. Start with the Manski model (Burridge, 2001), (Elhorst, 2010), which is the most general spatial dependence model. There are two lagged regressors and a spatial autocorrelated error term. The spatial weighting matrices can be the same or different. It is suggested to begin with the weighting matrices all equal to each other. Next, test the homoscedastic assumptions by setting the null hypothesis of the heteroskedastic location shocks equal to zero. This reduces the model to the spatial Durbin if the null hypothesis fails to be rejected.

The next test is to see if the spatial autocorrelation coefficient is statistically different from zero. If not, then both the Kelejian-Prucha and spatial Durbin models are eliminated. The spatial error and spatial Durbin error models remain and would need to be tested for the main effect and the autocorrelated error terms. If the coefficient on the spatial main effect is not statistically different from zero, then only the spatial error model remains. If the coefficient on the spatially lagged error term is statistically no different from zero, the resulting model is the nonspatial model, OLS.

Another way to select a model is by using the spatial regression decision method illustrated as text and by flow diagram in Anselin (2005). This method uses a set of model diagnostics beginning with an OLS estimation using Lagrange Multiplier (LM) tests. Depending on the outcome of the diagnostics, either the OLS is the chosen model or more models and tests are applied to the data. There are two applications of the LM tests, one to test for nuisance dependence and another to test for substantive spatial dependence. If only one of the two diagnostic test statistics is significant, then the model is estimated that corresponds to the spatial dependence identified.

If both of the diagnostic test statistics are significant, then more robust Lagrange multiplier test statistics for the error terms and spatial regressors are conducted to determine which spatial model should be performed next. Generally, only one test is significant and there is no recommendation on the flow chart for regressing a spatial Durbin or general spatial model that nest both spatial error and spatial lag models. The last method discussed here to use for model selection is to use the data generating process as illustrated by spatial diagram plots to decide which type of spatial dependence must be modeled. The fit statistics such as the log-likelihood values can be used to judge goodness-of-fit across all models estimated.

These paradigms are all assuming cross-sectional data. In an earlier paper by Elhorst (2001), he displays a flow chart of model selection using spatio-temporal data. The spatially and serially autocorrelated models estimated by OLS have biased but consistent estimates. However, spatial autocorrelation models have biased and inconsistent coefficients from OLS estimation. Thus, estimation cannot be accomplished satisfactorily by OLS, but can be estimated efficiently and consistently using maximum likelihood. Each of the variations of serially and spatially lagged models discussed by Elhorst (2001) can be estimated efficiently by maximum likelihood estimation. Estimation can be made more efficient by decreasing the number of parameters to be estimated.

Elhorst (2001) starts with a very general model that included serially lagged regressors, spatial lagged dependent and independent variables and spatially and serially lagged dependent and independent regressors where all lags are first-order. Various restrictions are imposed on the models such that the coefficient estimates for serial and spatial lags are a scalar factor of each other. This reduces the number of parameters to be estimated which increases the efficiency of the MLE coefficient estimator.

This study incorporated parts of these model selection paradigms where regression models from the specific to general were estimated after diagnostic tests were run on the OLS model. Beginning with SEM and SLM, the Mixed model was estimated and then the Durbin model was run. The chosen test for global spatial dependence was Moran's I. The LM tests and then the robust LM tests were used to decide between types of spatial models were the results were significant for more than one model.

3.3 Data

IRI Consumer Network data are received from the joint venture Nielsen/IRI National Consumer Panel (NCP) that is owned equally by Nielsen and IRI. NCP recruits members by advertisements where consumers are promised points toward gifts for milestone participation achievements and given in-home hand-held scanner devises to record their purchases from any store. Therefore, panelists are self-selected into the panel, but later reweighted to be nationally representative using projections from IRI.

Panelists must have had at least one day of positive ounces purchased of the given commodity per month for the entire 12 months to remain on the IRI yearly static panel of regularly participating panelists (Bronnenberg et al, 2008).

Coordinates for block group centroids from the Census Bureau were matched to the IRI households using state and county level FIPS codes. When estimating the spatial models, each household was uniquely identified by 12 digit Census geo identification number which indicated state + county + tract + block group. If the state and county level FIPS codes for a panelist household were missing or invalid, that member's observations were deleted from the dataset. The data were organized in the wide form where purchases in the current and past month were represented adjacent to each other and not stacked as with the long form of time-series datasets. The IRI datasets held many observations per month and the square weighting matrices were very large. Measures to save computer space in Stata MP 13.1 were needed to allow for spatial regressions to be performed.

The spatial weights matrix was exogenously determined using the inverse distance between block group centroids with latitudinal and longitudinal coordinates for the United States provided by the US Census. The distance band of centroids included as neighbors was determined for each commodity-year such that each remaining panelist would have at least one neighbor centroid. This was carried out to ensure full rank of the spatial weighting matrices. These matrices were row-standardized and thus, not symmetric. The same weights matrix was used in each of the spatial models. The mixed model used the same matrix twice. There was no identity problem for estimating ρ and λ when $\beta \neq 0$ (Bivand, 2011).

3.4 Model

Spatial autocorrelation models are most easily estimated as spatial lag models with spatial dependence restricted to a spatial autoregressive form (SAR). Spatial moving average models (SMA) are related models that can be estimated from this type of spatial influence.

Spatial error models (SEM) treat spatial dependence as non-meaningful (nuisance) in the sense that the spatial effect is between specified factors, but left undetermined in the error terms. There are consequences to ignoring spatial error. The spatial error model is basically autocorrelation for the model's residuals over the spatial dimension. So the spatial dependence is observed via the error terms of the fitted model. The spatial lag model is estimated as a spatial autoregressive model where the dependent variable is regressed on itself and the coefficient is the product of a spatial weight matrix and a strength-of-correlation factor. There are other spatial econometric models such as spatial probit models and spatial Poisson models where the dependent variable is discrete.

When estimating these SAR and SEM models, it is necessary to calculate a spatial weight matrix which functions as a spatial lag operator. These spatial econometric models do not add variables for distance measures between points of interest, but use weighting matrix pre-multiplying the spatially lagged dependent and/or independent regressors. When discussing time lags, there is only direction of influence; from the past to the most current time period of the data. In contrast, with spatial models, the influence is over areas and can occur from left or right and from above or below a given location. The result is that there are several directions of influence (or spatial dependence) on a given point and it is better to frame the spatial influence as the impact of location i on j where $i \neq j$ for all i, j $\in (1,2,\ldots,N)$. If we think of these locations and neighbors, then we can state that neighbor i has a given impact on the buying habits of neighbor j and on neighbor j's neighbor j+1, which can then form an N by N matrix of elements depicting the relational influence of these neighbors on each other where the diagonal elements of i's impact on i are all zeros.

The essence of each parametric spatial model is the ordinary least squares model. The OLS model is insufficient to represent the underlying data generating process because the units

are not independent of their neighbors and there is no variable in OLS models to represent this relationship. Thus, there is omitted variable bias and the error term captures the relationship that causes variation in the dependent variable. So the errors could be correlated spatially. These violations of Gauss-Markov assumptions of independent observations and homoscedastic variance of the errors lead to biased and inconsistent estimates from nonspatial models where OLS is a basic form. Different models try to represent the spatial interdependence either by spatially lagging the explanatory variables or by modeling the nuisance parameter.

The spatial autocorrelation model has a spatially weighted dependent variable added to an OLS model to capture the spatial effects that exist in the data. This spatial model differs from the first order autocorrelation model in that the main effects from a non-spatial model are allowed to capture effects other than spatially distributed interdependence. The first order autoregressive model creates a spatial lag of the dependent variable as its sole explanatory variable. Creating the spatial lag is performed by first defining and then quantifying the location component to economic or environmental data.

In determining the type of spatial dependence in the data, the data distribution was depicted on a map of the United States to check for spatial heterogeneity. If there was spatial dependence remaining after capturing the residuals from the geographically weighted regression (GWR) to control for spatial heterogeneity, then the type of spatial dependence was separate from the spatial heterogeneity. The source of spatial dependence could be checked regarding the dependent variable, independent regressors and the error term.

OLS was used as a baseline model for comparison to estimates from the spatial models. This nonspatial model assumes that the data do not vary by geographic region of the consumers. The spatial error model does not have spatial dependence of the regressors, but has errors with clusters by region because of unexplained variation of the dependent variable.

Due to the way county boundary lines were drawn separate from the economic attribution of interest and how spatial data are often provided via aggregation at the block

group or tract levels, there could be measurement error that leads to spatial dependence. As shown in Fig 2.1 of Anselin (1988, pg 12), the λ weighting parameter could determine the level of spatial correlation of Y_1 and Y_2. How regions are grouped regarding the availability of data could result in clustering for shopping trips by region that could be masked if grouped by tract versus block group. Some regions use MSA designations and other might use townships or villages to report geocoded data. There might be so much aggregation in some areas that the data variation between attributes averaged over those regions is dampened partially or completely.

Another type of spatial dependence is generated more substantially by consumer purchase variation dependent on location. People possibly change what they purchase and the frequency depending on geographic attributes like altitude, climate, existing infrastructure (cabs, mass transit, automobiles) and distance to shopping malls and outlet centers. For example, those in more mountainous locations might shop at large retail outlets and have higher expected travel distances than those in downtown locations of major urban cities. The density and variety of shops and apartments in a few square miles possibly affects buying locally and influences consumers' tolerance for driving long distances for groceries or other goods.

There can be variation in geographical landscape, but stability in the structure that relates the economic and demographic attributes to the geographic characteristics. The spatial stability represents constancy in how the environment affects consumers. In more densely populated urban areas, the average distances traveled to buy groceries is lower than in the more sparsely populated rural areas. If this spatial relationship between travel time to shop and population density changes sharply in magnitude or sign as one reached a certain boundary line in the US, the structural change (instability) in spatial dependence would be the spatial heterogeneity.

Spatial lag models (also called spatial autoregressive models) have spatial dependence between neighbors for the dependent variable. This is often modeled as spatial autoregression, but spatial autoregression is only one type of spatial dependence. In SAR and SEM, each model

has only one source of spatial dependence. Models with two sources of spatial dependence are the Kelejian and Prucha model (KPM), the spatial Durbin model (SDM) and the spatial Durbin Error model (SDEM). KPM has spatial dependence in the dependent variable and the error term. SDEM does not have a spatial lag term, but spatial dependence of the independent regressors and the error term. The dependent and independent regressors are spatially lagged in the SDM. This model is adjusted to model space-time dependence.

I estimated models starting with the SDM that includes an adjustment of a temporal lag and a spatially lagged temporal lag similar to the model estimated by Korniotis (2010) and largely derived from Elhorst (2001). The baseline model that I used was the OLS which is one of the simplest nonspatial models in order to determine how much variation in the dependent variable is determined by the spatial regressors if any.

The general model specified in figure 1 of the Elhorst (2001) paper as Model 0 is

(1) $Y_t = \tau Y_{t-1} + \delta W Y_t + \eta W Y_{t-1} + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 W X_t + \beta_4 W X_{t-1} + u_t,$

Each of the other models can be derived from this model above. The usual term for spatial correlation of the spatial lag term is ρ and not δ, so this will be used instead of δ.

The spatial lag model (also called the spatial autoregressive model in other papers) has the following structure:

(2) $Y_i = \rho W Y_i + \beta_1 X_i + u_i, \quad u_i \sim N(0, \sigma^2), \; cov(uu') = \sigma^2 I_n$

for households $\mathrm{i} = 1,2, \ldots \mathrm{n}, \mathrm{u} = [\mathrm{u}_1\ \mathrm{u}_2\ \mathrm{u}_3 \ldots \mathrm{u}_\mathrm{n}]'$

where Y_i are aggregated quantities purchased in ounces at the household level of one of the commodities in this study (whole grains, other grains, fruit juices or sugar-sweetened beverages), X_i are average prices paid ($/oz) across commodities per household and other demographic variables (race, Hispanic status, region of the country, county size and household size). The same specification of variables holds for equations (3) – (6).

Spatial error model

(3) $Y_i = \beta X_i + u_i,$

$u_i = \lambda W u_i + \varepsilon_i, \;\; \varepsilon_i \sim N(0, \sigma^2), cov(uu') = \sigma^2[(I_n - \lambda W)'(I_n - \lambda W)]^{-1}$

for households $\mathrm{i} = 1,2, \ldots \mathrm{n}$, $\mathrm{u} = [\mathrm{u}_1\ \mathrm{u}_2\ \mathrm{u}_3 \ldots \mathrm{u}_\mathrm{n}]'$

Spatial mixed model (including the spatial lag and spatial error models)

(4) $Y_i = \rho W Y_i + \beta_1 X_i + u_i,$

$u_i = \lambda W u_i + \varepsilon_i, \;\; \varepsilon_i \sim N(0, \sigma^2), cov(uu') = \sigma^2[(I_n - \lambda W)'(I_n - \lambda W)]^{-1}$

for households $\mathrm{i} = 1,2, \ldots \mathrm{n}$, $\mathrm{u} = [\mathrm{u}_1\ \mathrm{u}_2\ \mathrm{u}_3 \ldots \mathrm{u}_\mathrm{n}]'$

Spatial Durbin model

(5) $Y_i = \rho W Y_i + \beta_1 X_i + \beta_2 W X_i + u_i, \;\; u_i \sim N(0, \sigma^2), \;\; cov(uu') = \sigma^2 I_n$

for households $\mathrm{i} = 1,2, \ldots \mathrm{n}$, $\mathrm{u} = [\mathrm{u}_1\ \mathrm{u}_2\ \mathrm{u}_3 \ldots \mathrm{u}_\mathrm{n}]'$

Adjusted Spatial Durbin model (which is similar to the model by Korniotis (2010)) includes a time-lagged spatial lag for the dependent variable and the independent regressors to measure any diffusion time needed for the neighborhood effect to be realized. This model specification is show in equation (6):

(6) $Y_{it} = \tau Y_{i,t-1} + \delta W Y_{it} + \eta W Y_{i,t-1} + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 W X_t + \beta_4 W X_{t-1} + u_t,$

$u_i \sim N(0, \sigma^2)$

where $\eta W Y_{i,t-1}$ is the time-lagged spatial lag term capturing the diffusion effect of the neighborhood effect on consumption.

3.5 Methods

OLS estimation does not have unbiased and efficient properties when applied to data with underlying spatial processes. When estimating a model with spatial dependence with the regressors using OLS, the result is biased and inconsistent estimates. Spatial regressors that do not explain all of the spatial variation in the dependent variable result in either introducing spurious spatial correlation in the error term that was not present without those regressors or failing to explain the spatial dependence in the error term that was there previously.

The maximum likelihood estimator generates consistent and efficient estimates when the assumption of normally distributed error terms is satisfied. The joint log-likelihood function is

not simply a sum of the individual log-likelihood functions as with the time-series case or the case of cross-sectional data.

There are several tests to determine if a spatial interaction exists in the data. There are also tests to establish if the spatial effect is in the spatial independent variables, the spatial lagged dependent variable or the disturbance terms. Comparison of the OLS R^2 and parameter estimates to their counterparts of the spatial autoregressive mixed (Mixed) regression model helped decipher if the spatial component exists and how much additional variation in the dependent variable is explained by spatial modeling. Tests that use the contiguity spatial weighting matrix (row-standardized or not row-standardized) for determining spatial autocorrelation in the disturbance term are Moran's I statistic, the likelihood ratio test, the Wald statistic, the Lagrange Multiplier test for error terms and for lagged dependent variables where the asymptotic distributions for these statistics are used to determine statistical significance levels for drawing inferential statistical conclusions.

Spatial autocorrelation is the spatial dependence of location values in a designated area and can be either positive or negative. Global autocorrelation is the existence of overall clustering of values in a given area and local autocorrelation is subsector areas that combine to form this overall clustering pattern or detract from that general pattern (Rey and Anselin, 2010). There are many tests for global spatial autocorrelation such as the Gamma Index, join count statistics, Moran's I statistic, Matheron's variogram. Moran's I uses a specific attribute across a set of n spatial units to determine the existence of global autocorrelation. Matheron's classic variogram is based on the method of moments and contains a symmetric spatial design matrix (Genton, 2000).

A local spatial autocorrelation test statistic is Geary's C, which uses a spatial weight and a squared difference of the attribute, but this difference is between points i and j and not point i and the attribute's mean. Another local spatial autocorrelation test uses the Getis and Ord's G statistic which has a threshold distance to determine the spatial weight values that generate the statistic. Local Indicators of Spatial Association (LISA) has significance maps and cluster maps. There is a local version of the Moran's I statistic for local spatial autocorrelation tests.

It is important to have local spatial autocorrelation tests that are still robust in the presence of global spatial autocorrelation. This might require the ability to locate "hot spots" in the regions of the data that differ from the global or overall spatial autocorrelation. There might be difficulties finding these hot spots because of a strong global autocorrelation effect that masks other areas where the degree of autocorrelation differs from other regions (Ord and Getis, 2001). During the search for spatial patterns in a large dataset, there is most likely enough variation for a global pattern to emerge. The issue is how to find the local spatial patterns in the midst of the global pattern. The global spatial structure can be determined using variograms and other diagrams of cluster patterns such as correlograms. These plots are of spatial autocorrelation as a function of distance from a point of reference.

In order for the local spatial autocorrelation to be determined, a few assumptions must be met (Ord and Getis, 2001). The data must be normally distributed for exact results. The assumption of stationarity of the covariance of the spatial process must be identifiable on the variogram. When modeling the autocorrelation of the error terms, the assumption is that the estimation errors are ignorable. Geary's c statistic, Moran's I and Matheron's variogram are not able to find spatial autocorrelation past the global spatial structures in the data. LISA and Getis and Ord's G statistic are able to locate local patterns of autocorrelation, but not in the presence of global spatial autocorrelation. The G statistic is a spatial moving average measure. In Ord and Getis (2001), a similar statistic to the G statistic is developed to find local hot spots in the data that are local clusters of spatial dependence in the presence of global spatial dependence structures.

Maximum likelihood estimation will be used to produce estimates of the regression model. The coefficients that appear in more than one model will be compared to determine the level and type of spatial dependence. Also, comparison of R^2, log-likelihood and Akaike values will be used to determine the model specification that has the best fit with respect to the spatio-temporal data. The same models will be estimated using different years of IRI data.

Four spatial econometric models were regressed using both 2008 and 2012 IRI data for each commodity with price as the only regressor to determine how sensitive the results were

to additional variables for neighborhood effects, socio-demographic status, time lag and time lagged spatial lag of the dependent variable. Comparison of the models based on significance of the spatial correlation coefficients ρ and λ.

3.6 Results

The sugar-sweetened beverages model using 2008 data had statistically significant ρ and λ estimates for each of the models. Table 3-2 shows that λ of the spatial mixed model was not significant, but λ of the spatial error model was significant at the 1 percent α level. There was significant external habit formation with ρ estimated between 0.14 and 0.33 across each of the applicable models. Thus, there is a neighborhood effect from the spatial lag of the dependent variable where consumption of individual i is influenced by the consumption level of others living in the same area. This SLM is most fitting for the data given that the Durbin model had an insignificant spatially lagged independent variable and was essentially the same as the SLM. Also, the $\lambda = 0.139$ of the SEM was no longer statistically significant when accounting for the spatially lagged dependent variable. So there is an indication from the data that peers influence the consumption of each other by a factor of 0.141.

Table 3-1. Sugar-Sweetened Beverages 2008 Diagnostics

Fitted Model			
Total commodity quantity purchased (in ounces) = Average commodity price paid ($/oz)			
Weights Matrix			
Name: Sugar-sweetened Beverages 2008 IRI data			
Type: Distance-based (binary)			
Distance band: 0.0 < d <= 0.196 km			
Row-standardized: Yes			
OLS Diagnostics			
Test	Statistic	df	p-value
Spatial error:			
Moran's I	3.403	1	0.001
Lagrange multiplier	11.220	1	0.001
Robust Lagrange multiplier	1.015	1	0.314
Spatial lag:			
Lagrange multiplier	11.661	1	0.001
Robust Lagrange multiplier	1.456	1	0.228

Table 3-2 Sugar-sweetened Beverages Reference Table

	Durbin b/se	Mixed b/se	SEM b/se	SLM b/se
Current average price/oz	–926.694*** (73.00)	–928.146*** (110.55)	–924.231*** (110.81)	–926.846*** (73.09)
Spatially lagged current average price/oz	–702.288 (614.13)			
Constant	824.776*** (48.56)	629.420*** (112.56)	916.427*** (10.87)	792.847*** (38.31)
rho	0.138*** (0.04)	0.328** (0.13)		0.141*** (0.04)
sigma	817.260*** (19.08)	815.779*** (6.22)	817.315*** (6.00)	817.292*** (19.09)
lambda		–0.220 (0.16)	0.139*** (0.04)	

* p <0.10, ** p<0.05, *** p<0.01

Table 3-3. Sugar-sweetened Beverages 2012 Diagnostics

Fitted Model			
Total commodity quantity purchased (in ounces) = Average commodity price paid ($/oz)			
Weights Matrix			
Name: Sugar-sweetened Beverages 2012 IRI data			
Type: Distance-based (binary)			
Distance band: 0.0 < d <= 0.184			
Row-standardized: Yes			
OLS Diagnostics			
Test	Statistic	df	p-value
Spatial error:			
Moran's I	1.800	1	0.072
Lagrange multiplier	3.059	1	0.080
Robust Lagrange multiplier	4.555	1	0.033
Spatial lag:			
Lagrange multiplier	3.502	1	0.061
Robust Lagrange multiplier	4.998	1	0.025

Table 3-4 Sugar-sweetened Beverages 2012 Reference Table

	Durbin b/se	Mixed b/se	SEM b/se	SLM b/se
Current average price/oz	–676.497*** (91.55)	–660.831*** (105.64)	–684.610*** (106.46)	–687.068*** (92.27)
Spatially lagged current average price/oz	-1174.756 (441.74)			
Constant	829.638*** (43.54)	1220.869*** (84.01)	836.533*** (11.52)	773.313*** (36.88)
rho	0.067 (0.04)	–0.473*** (0.10)		0.078* (0.04)
sigma	765.981*** (22.59)	759.489*** (7.30)	766.253*** (7.00)	766.219*** (22.60)
lambda		0.461*** (0.07)	0.074* (0.04)	

p <0.10, ** p<0.05, *** p<0.01

Note: rho is the spatial correlation coefficient for the dependent variable, lambda is the spatial correlation coefficient for the error term, sigma is the estimated spherical variance variable.

In 2012, the sugar-sweetened beverages spatial models in Table 3-4 showed a spatial relationship between the prices of neighboring areas and the commodity quantities purchased in the reference area. This could show evidence of consumers' awareness of local prices that helps influence their purchases. The spatial autoregressive correlation coefficient from the Durbin model is not significant, but it is significant in the spatial mixed model and the spatial lag model. After accounting for the spatial correlation of the dependent variable, there is still spatial correlation of the error terms. When estimated using the spatial errors model, λ is still significant, but is lowered from 0.46 to 0.07. The Mixed model best fits the data in 2012 for sugar-sweetened beverages with a negative neighborhood effect of 0.473.

Table 3-5 Fruit Juices 2008 Diagnostics

Fitted Model			
Total commodity quantity purchased (in ounces) = Average commodity price paid ($/oz)			
Weights Matrix			
Name: Other Grains 2008 IRI data			
Type: Distance-based (binary)			
Distance band: 0.0 < d <= 0.187 km			
Row-standardized: Yes			
OLS Diagnostics			
Test	Statistic	df	p-value
Spatial error:			
Moran's I	1.946	1	0.052
Lagrange multiplier	3.580	1	0.058
Robust Lagrange multiplier	0.153	1	0.696
Spatial lag:			
Lagrange multiplier	4.161	1	0.041
Robust Lagrange multiplier	0.735	1	0.391

Table 3-6. Fruit Juices Spatial Regression Models 2008 Reference Table

	Durbin b/se	Mixed b/se	SEM b/se	SLM b/se
Current average price/oz	–1857.328*** (70.94)	–1859.841*** (94.89)	–1861.682*** (94.91)	–1859.528*** (71.02
Spatially lagged current average price/oz	–112.856 (377.97)			
Constant	389.479*** (29.85)	384.219*** (36.82)	410.433*** (6.47)	381.913*** (14.82)
rho	0.090* (0.05)	0.086 (0.12)		0.094** (0.04)
sigma	231.457*** (5.38)	231.462*** (02.24)	231.468*** (2.23)	231.456*** (5.38)
lambda		0.009 (0.13)	0.092* (0.05)	

p <0.10, ** p<0.05, *** p<0.01

Note: rho is the spatial correlation coefficient for the dependent variable, lambda is the spatial correlation coefficient for the error term, sigma is the estimated spherical variance variable.

Table 3-7 Fruit Juices 2012 Diagnostics

Fitted Model			
Total commodity quantity purchased (in ounces) = Average commodity price paid ($/oz)			
Weights Matrix			
Name: Other Grains 2012 IRI data			
Type: Distance-based (binary)			
Distance band: 0.0 < d <= 0.296			
Row-standardized: Yes			
OLS Diagnostics			
Test	Statistic	df	p-value
Spatial error:			
Moran's I	1.338	1	0.181
Lagrange multiplier	1.591	1	0.207
Robust Lagrange multiplier	0.210	1	0.647
Spatial lag:			
Lagrange multiplier	1.394	1	0.238
Robust Lagrange multiplier	0.013	1	0.909

Table 3-8 Fruits Juices 2012 Reference Table

	Durbin b/se	Mixed b/se	SEM b/se	SLM b/se
Current average price/oz	–1694.145*** (76.02)	–1679.802*** (111.69)	–1693.768*** (110.31)	–1691.748*** (75.63)
Spatially lagged current average price/oz	183.772 (445.83)			
Constant	339.679*** (38.02)	444.208*** (56.58)	372.217*** (8.95)	353.561*** (19.36)
rho	0.079 (0.07)	–0.281 (0.22)		0.071 (0.07)
sigma	205.264*** (7.74)	204.158*** (3.68)	205.265*** (3.42)	205.281*** (7.74)
lambda		0.325* (0.17)	0.079 (0.07)	

p <0.10, ** p<0.05, *** p<0.01

Note: rho is the spatial correlation coefficient for the dependent variable, lambda is the spatial correlation coefficient for the error term, sigma is the estimated spherical variance variable.

.

Table 3-9 Other Grains 2008 Diagnostics

Fitted Model			
Total commodity quantity purchased (in ounces) = Average commodity price paid ($/oz)			
Weights Matrix			
Name: Other Grains 2008 IRI data			
Type: Distance-based (binary)			
Distance band: 0.0 < d <= 0.289 km			
Row-standardized: Yes			
OLS Diagnostics			
Test	Statistic	df	p-value
Spatial error:			
Moran's I	5.077	1	0.000
Lagrange multiplier	25.014	1	0.000
Robust Lagrange multiplier	0.779	1	0.377
Spatial lag:			
Lagrange multiplier	24.235	1	0.000
Robust Lagrange multiplier	0.001	1	0.976

Table 3-10 Other Grains 2008 Reference Table

	Durbin b/se	Mixed b/se	SEM b/se	SLM b/se
Current average price/oz	-201.3*** (18.77)	-201.8*** (14.20)	–200.5*** (14.08)	–196.9*** (17.68)
Spatially lagged current average price/oz	61.714 (52.85)			
Constant	96.478*** (8.01)	150.495*** (18.71)	127.584*** (1.80)	103.799*** (5.30)
rho	0.231*** (0.05)	–0.216 (0.18)		0.223*** (0.05)
sigma	75.003*** (1.01)	72.916 (0.46)	73.003*** (0.45)	73.009*** (0.05)
lambda		0.406*** (0.13)	0.232*** (0.05)	

p <0.10, ** p<0.05, *** p<0.01

Note: rho is the spatial correlation coefficient for the dependent variable, lambda is the spatial correlation coefficient for the error term, sigma is the estimated spherical variance variable.

Table 3-11 Other Grains 2012 Diagnostics

Fitted Model			
Total commodity quantity purchased (in ounces) = Average commodity price paid ($/oz)			
Weights Matrix			
Name: Other Grains 2012 IRI data			
Type: Distance-based (binary)			
Distance band: 0.0 < d <= 0.224			
Row-standardized: Yes			
OLS Diagnostics			
Test	Statistic	df	p-value
Spatial error:			
Moran's I	4.153	1	0.000
Lagrange multiplier	16.706	1	0.000
Robust Lagrange multiplier	9.258	1	0.002
Spatial lag:			
Lagrange multiplier	12.726	1	0.000
Robust Lagrange multiplier	5.278	1	0.022

Table 3-12 Other Grains 2012 Reference Table

	Durbin b/se	Mixed b/se	SEM b/se	SLM b/se
Current average price/oz	–198.116*** (13.86)	–195.998*** (15.23)	–191.172*** (15.04)	–186.718*** (12.98)
Spatially lagged current average price/oz	125.425*** (47.68)			
Constant	88.040*** (6.97)	160.106*** (10.65)	117.911*** (2.00)	102.811*** (4.77)
rho	0.164*** (0.04)	–0.437*** (0.11)		0.153*** (0.05)
sigma	68.681*** (2.25)	68.334*** (0.52)	68.694*** (0.51)	68.710*** (2.25)
lambda		0.502*** (0.07)	0.172*** (0.05)	

p <0.10, ** p<0.05, *** p<0.01

Note: rho is the spatial correlation coefficient for the dependent variable, lambda is the spatial correlation coefficient for the error term, sigma is the estimated spherical variance variable.

Tables 2.3 and 2.4 show the results for fruit juices in 2008 and 2012. Using the 2008 data, the basic spatial models SLM and SEM showed significant correlation of the dependent variable and errors respectively. Moran's I statistic is statistically significant and confirms global spatial autocorrelation. Testing those spatial correlations together showed no spatial correlation of either element. Also, the Wald test for Mixed versus OLS failed to reflect the OLS model. The spatial Durbin model had an insignificant coefficient on the spatial lag of average monthly prices per ounce. The ρ value of 0.090 of the Durbin model was close to the 0.094 of the SLM. Therefore, the SLM best fit the 2008 fruit juices data.

Table 3-13 Whole Grains 2008 Diagnostics

Fitted Model			
Total commodity quantity purchased (in ounces) = Average commodity price paid ($/oz)			
Weights Matrix			
Name: Whole Grains 2008 IRI data			
Type: Distance-based (binary)			
Distance band: 0.0 < d <= 0.279			
Row-standardized: Yes			
OLS Diagnostics			
Test	Statistic	df	p-value
Spatial error:			
Moran's I	2.020	1	0.043
Lagrange multiplier	3.721	1	0.054
Robust Lagrange multiplier	27.819	1	0.000
Spatial lag:			
Lagrange multiplier	1.888	1	0.169
Robust Lagrange multiplier	25.986	1	0.000

Table 3-14 Whole Grains 2008 Reference Table

	Durbin b/se	Mixed b/se	SEM b/se	SLM b/se
Current average price/oz	–197.059*** (39.67)	–187.551*** (36.27)	–158.012*** (37.13)	–146.538*** (37.27)
Spatially lagged current average price/oz	468.068*** (101.99)			
Constant	42.198*** (6.97)	146.848*** (9.23)	95.836*** (4.60)	88.066*** (6.43)
rho	0.043*** (0.06)	–0.622*** (0.10)		0.0.084 (0.06)
sigma	52.521*** (2.29)	51.328*** (1.01)	52.917*** (0.97)	52.957*** (2.32)
lambda		0.570*** (0.06)	0.117*** (0.06)	

p <0.10, ** p<0.05, *** p<0.01

Note: rho is the spatial correlation coefficient for the dependent variable, lambda is the spatial correlation coefficient for the error term, sigma is the estimated spherical variance variable.

Table 3-15 Whole Grains 2012 Diagnostics

Fitted Model			
Total commodity quantity purchased (in ounces) = Average commodity price paid ($/oz)			
Weights Matrix			
Name: Whole Grains 2012 IRI data			
Type: Distance-based (binary)			
Distance band: 0.0 < d <= 0.373			
Row-standardized: Yes			
OLS Diagnostics			
Test	Statistic	df	p-value
Spatial error:			
Moran's I	1.339	1	0.181
Lagrange multiplier	1.509	1	0.219
Robust Lagrange multiplier	4.177	1	0.041
Spatial lag:			
Lagrange multiplier	0.890	1	0.345
Robust Lagrange multiplier	3.558	1	0.059

Table 3-16 Whole Grains 2012 Reference Table

	Durbin b/se	Mixed b/se	SEM b/se	SLM b/se
Current average price/oz	–201.026*** (47.05)	–195.025*** (42.92)	–183.291*** (43.08)	–178.261*** (45.54)
Spatially lagged current average price/oz	207.146** (93.84)			
Constant	68.748*** (13.92)	139.712*** (13.23)	97.510*** (5.46)	91.284*** (8.71)
rho	0.082 (0.04)	–0.542*** (0.16)		0.075 (0.08)
sigma	52.506*** (3.86)	51.361*** (1.34)	52.598*** (1.26)	52.621*** (3.86)
lambda		0.512*** (0.10)	0.096 (0.08)	

p <0.10, ** p<0.05, *** p<0.01

Note: rho is the spatial correlation coefficient for the dependent variable, lambda is the spatial correlation coefficient for the error term, sigma is the estimated spherical variance variable.

Other grains in 2008 had spatially correlated errors and dependent variable with model results and spatial dependence diagnostic results shown in Tables 3-9 and 3-10. The Mixed model had spatially correlated errors but not the dependent variable. The spatial Durbin model did not improve over the SLM due to the insignificant spatially lagged averaged prices coefficient. The model most statistically associated with the 2008 other grains data was the Mixed model which showed evidence for the nested SEM. In the Mixed model, both the SEM and SLM are nested, so the Mixed model can directly test both models to determine if both, one or the other, or neither is statistically significant. This Mixed model (with evidence for the nested spatial error model) was also selected by the Wald test when the OLS model was rejected with a $\chi^2_{(2)}$=36.76 (p = 0.000). Therefore, the other grains data in 2008 shows no indication of neighborhood effects. There was also no evidence of influence of neighboring

prices on consumption of households in block group i. The spatial Durbin model yields unbiased parameters regardless of the type of true data generating process (Glass, Kenjegalieva and Sickles, 2012). Even if the actual data generating process is SLM or SEM, Durbin will give unbiased parameter estimates. Therefore the Durbin model with $\rho = 0.164$ is the model that was selected to represent the spatial process for other grains in 2012.

Whole grains data in 2008 had spatial correlation in SEM, but not the SLM. The SEM model might have been suffering from omitted variable bias and the SLM might have had misspecification of the errors, which could have caused incorrect inferential analysis (LeSage and Pace, 2009). Moran's I statistic was significant as were the robust LM statistics which confirms that there was spatial correlation that might have been undetected in the SLM and SEM due to the reasons given above. When the models were combined, the results indicated both spatially correlated errors and dependent variable. This model still might have suffered from omitted variable bias because the spillover effect was negative for whole grains.

The models' results of Table 2.15 were similar for whole grains 2012 data, but now the SEM and SLM show insignificant autocorrelation, but the Mixed model did show spatial effects of the dependent variable and errors. The significant $\rho = -0.542$ is similar to the -0.622 from the 2008 Mixed model. The significant $\lambda=0.512$ is close to the 0.570 of the corresponding Mixed model of 2008. These whole grains data from 2008 and 2012 were sensitive to included variables and the negative spatial lag coefficient might indicate that the spatial lag term was representing other factors that have not been included in the model. The SEM of 2008 and 2012 might also have been sensitive to omitted variable bias. The models to select when deciding between spatial Durbin and model with spatial dependence in the dependent and error terms is the spatial Durbin because it produces unbiased estimates even when the true spatial data generating process is a spatial error or spatial lag model (LeSage and Pace, 2009). Therefore, the spatial models chosen for whole grains in 2008 and 2010 were the spatial Durbin models were the spatial dependence diminishes from 0.043 external habit formation to no statistically significant external habit formation in 2012.

Other grains and sugar-sweetened beverages have decreases in external habit formation respectively across years 2008 and 2012 before and after the publication of the 2010 Dietary

Guidelines for Americans. This decrease in neighborhood influences for these relatively unhealthy food and beverage commodities was not statistically significant, but was a move in the desired direction from the perspective of nutritionists and dietitians of the dietary guidelines.

2.7 Conclusions

There was evidence of neighborhood effects for sugar-sweetened beverages, fruit juices and whole grains for at least one of the comparison years. Sugar-sweetened beverages in 2008 showed significant positive neighborhood effects and then negative effects in 2012. Public awareness of the negative effects of consuming beverages with added sugars might have caused households to be more sensitive to their purchases of those items when viewing those purchases of others that they know in their communities. The USDA DGA could have increased the discussion of other healthy eating advocacy groups if they did not directly increase awareness of households regarding beverages with added sugars.

Fruit juices had a positive peer effect in 2008 and then no peer effect shown in 2012. Households could have been more aware of others' consumption of drinks with added sugars than with fruit juices, so the small positive neighborhood effect was not surprising to have diminished by 2012. Many venues and vending machines offer sodas and sports drinks, but not as many places offer 100 percent fruit juices or simply with more than 50 percent fruit juice. Increasing public knowledge of community consumption of healthy and unhealthy beverages could lead households to lower consumption of sodas and energy drinks and increase purchases of drinks with high percentages of fruit juices.

Surprisingly, whole grain purchases of households in 2008 and 2012 showed a strong peer effects factor ranging from -0.622 to -0.542. Controlling for neighborhood conditions and demographic characteristics did not increase the autoregressive correlation between neighbors to a positive effect. The negative impact from neighbors did decrease from 2008 to 2012, but was still strong. Households in one area could have high consumption of whole grains, but in the next community consider such practices to be arrogant or simply undesirable. This might explain the results for whole grains, but it is difficult to decipher the meaning of these peer effects as pertaining to the USDA DGA.

Chapter 4 Battle of the Consumer Demand Models

4.1. Introduction

The unique USDA's National Household Food Acquisition and Purchase Survey (FoodAPS) dataset collected in 2012 by Mathematica Policy Research (MPR) group contains food prices and quantities. In contrast to the NHANES dataset, which is an intake survey without price information, the Consumer Expenditure Survey which has food expenditures but neither price nor quantity, and the IRI dataset which has food price and quantity information, but was not a survey dataset.

The EASI demand system of Lewbel and Pendakur (2008) is a new technique that can be applied to the FoodAPS data to measure price and income sensitivity using demographic information and Engel curve flexibility. EASI is linear in prices only and polynomial in demographic characteristics and implicit utility and thus has a more flexible form than the AIDS model. Interaction terms can also be added to the EASI model as opposed to the AIDS model.

Specifically, the linear approximation to the EASI (LA/EASI) demand model will be applied to the FoodAPS dataset via iterated seemingly unrelated regression (ITSUR) estimation and compared to estimates of the linear approximation to the AIDS demand model. The estimates of these models will provide information that is essential for assessing the credibility and versatility of the LA/EASI demand system and the usefulness of the 2012 FoodAPS.

4.2. Literature Review

4.2.1. Almost Ideal Demand System

Deaton and Muellbauer (1980) introduced the Almost Ideal Demand System (AIDS). Its flexible functional form and aggregation properties have made it a very attractive demand model. Over the years, many articles have been published comparing the AIDS to other demand systems of the linear estimation system known as LES. The model is very popular and also has a linear version that can be applied very easily as it uses the Stone price index to estimate the real expenditures in the model.

4.2.2. Linear Approximation to the AIDS Model

The linear approximation to the AIDS (LA/AIDS) uses the Stone price index to replace the nonlinear price index of the nonlinear AIDS model. It retains all of the properties of the AIDS, and adheres to the adding up, homogeneity and symmetry restrictions of the AIDS. This model does have some ambiguity regarding estimation of price elasticities. It is not always known if it is better to estimate the full AIDS model to generate the coefficients to be used in the calculation of price elasticities, or if the LA/AIDS model would be sufficient for price elasticity estimation. Asche and Wessells (1997) demonstrated that the price and income elasticity formulas were the same if the prices and income were normalized to 1. Independent of normalization, the income elasticity formulas are the same for AIDS and LA/AIDS.

4.2.3. Linear Approximation to the Exact Affine Stone Index

Lewbel and Pendakur (2009) developed the Exact Affine Stone Index (EASI) demand system that is more flexible than the AIDS and other traditionally used demand systems. Exact price indices are used in the EASI demand system as opposed to the AIDS model, which results in more exact model estimates. The EASI demand system shares the flexible demand and nonlinear properties of the AIDS model, but unlike AIDS with linear Engel curves for real demand, EASI has Engel curves that can be fit to many shapes depending on the data. The linear approximation to the EASI retains many of the properties of the full nonlinear EASI and is easier to apply to the data. Therefore the LA/EASI will be used in this model to estimate price and income elasticities and generate Engel curves.

4.3. Models

The LA/AIDS model application is an incomplete system of six nonalcoholic beverage commodity groups j = 1, 2, 3, 4, 5, 6 where each share is represented as follows:

(1) $w_j = \alpha_j + \sum_{k=1}^{6} \gamma_{kj} \log p_k + \beta_j \log(x/P)$.

The intercepts are α_j, p_k are the prices for the commodity items, P is the Stone price index and x is nominal income.

P is the linear Stone price index expressed in logarithmic form as:

(2) $P = \sum_{j=1}^{6} w_j \log p_j$.

Adding up constraints $\sum_{j=1}^{n} \alpha_j = 1, \quad \sum_{j=1}^{n} \gamma_{jk}, \quad \sum_{j=1}^{n} \beta_j = 0$

(3) Zero-degree homogeneity $\sum_{k=1}^{n} \gamma_{jk} = 0$

Slutsky symmetry $\gamma_{jk} = \gamma_{kj}$

Marshallian LA/AIDS price elasticities formula for commodity j given a price change for commodity k is

(4) $e_{jk} = \frac{\gamma_{jk}}{w_j} - \frac{\beta_j w_k}{w_j} - \delta_{jk}$

Where δ_{jk} is the Kronecker delta, thus δ_{jk}=1 if i=j and δ_{jk}=0 if i≠j.

The expenditure elasticity is

(5) $\eta_j^x = 1 + \frac{\beta_j}{w_j}$

The linear approximation to the Exact Affine Stone Index (LA/EASI) will be used to estimate price and expenditure elasticities for nonalcoholic beverages. The dependent variables will be budget shares of milk, tea and coffee combined, fruit juices, juice drinks (very low percentages of fruit juice content), carbonated soft drinks and bottled water where each category is mutually exclusive.

This LA/EASI model is an incomplete system of six goods j=1, 2, 3, 4, 5, 6 where each share is represented as follows:

(6) $w_j = \sum_{r=0}^{R} b_{rj} y^r + \sum_{l=1}^{L} (C_{lj} z_l + D_{lj} z_l y) + \sum_{l=0}^{L} \sum_{k=1}^{J} A_{lkj} z_l p_k + \sum_{k=1}^{J} B_{kj} p_k y + \varepsilon_j$,

where R is the affine polynomial order in income y where $y = \log x - \sum_{j=1}^{6} \overline{w}_j \log p_j$ uses the Stone price index, x is total nominal expenditures, p_k is log-prices, z_l are the L different demographic variables and where three sets of interaction terms with prices, income or demographic characteristics are included (Lewbel and Pendakur, 2009).

(7) Adding up constraint $\sum_{j=1}^{n} A_{ljk} = \sum_{j=1}^{n} B_{jk} = 0 \; for \; l = 0, 1, \ldots L \;\; k = 1, \ldots, n.$

(8) Homogeneity $\sum_{j=1}^{n} C_{jl} = \sum_{j=1}^{n} D_{jl} = 0, \sum_{j=1}^{n} b_{j0} = 1, \sum_{j=1}^{n} b_{jr} = 0, for \; r \neq 0.$

(9) $A_{jk} = A_{kj} \; such \; that$ A matrix of A_{jk} elements is symmetric and $B_{jk} = B_{kj} \; such \; that$ B matrix of B_{jk} elements is symmetric satisfies the Slutsky symmetry constraint.

From the appendix of Rahkovsky and Snyder, 2015 the Hicksian elasticity can be derived from the semi-elasticity formula.

Hicksian semi-elasticity formula:

(10) $\frac{\partial w_j}{\partial lnp_k} = \sum_{l=0}^{L} A_{lkj} z_l + \sum_{k=1}^{J} B_{kj} y$,

(11) $$\frac{\partial w_j}{\partial lnp_k} = \frac{\partial\left[\frac{q_j^h p_k}{x^h}\right]}{\partial lnp_k} = \frac{\partial q_j^h}{\partial lnp_k}\frac{p_i}{x^h} + \frac{\partial p_j}{\partial lnp_k}\frac{q_j^h}{x^h} - \frac{q_j^h p_k}{(x^h)^2} - \frac{\partial x^h}{\partial lnp_k}$$

$$= \frac{\partial q_j^h}{\partial lnp_k}\frac{q_j^h}{q_j^h}\frac{p_i}{x^h} + \frac{\partial p_j}{\partial lnp_k}\frac{p_j}{p_j}\frac{q_j^h}{x^h} - \frac{q_j^h p_k}{(x^h)^2} - \frac{\partial x^h}{\partial p_k}\frac{p_j}{1}$$

$$= \frac{\partial \ln q_j^h}{\partial lnp_k} w_j^h + \frac{\partial \ln p_j}{\partial lnp_j}\frac{q_j^h p_j}{x^h} - \frac{q_j^h p_j}{x^h}\frac{q_k^h p_k}{x^h}$$

$$= \frac{\partial \ln q_j^h}{\partial lnp_k} w_j^h + 1_{jk} w_j^h - w_j^h w_k^h$$

Then solve for $\frac{\partial \ln q_j^h}{\partial lnp_k}$ to get Hicksian elasticity:

(12) $$e_{jk} = \frac{\partial \ln q_j^h}{\partial lnp_k} = \frac{\partial w_j}{w_j^h \partial lnp_k} + w_k^h - 1_{jk}$$

Where 1_{jk} is the Kronecker delta so 1_{jk}=1 if j=k and 1_{jk}=0 if j≠k.

The same method can be used to develop the income elasticity from the income semi-elasticity.

(13) $$\frac{\partial w_j}{\partial lnx^m} = \frac{\partial\left[\frac{q_j^h p_k}{x^h}\right]}{\partial lnx^m} = \frac{\partial q_j^m}{\partial x}\frac{p_j}{x^m} + \frac{\partial\left[\frac{1}{x^m}\right]}{\partial x^m}\frac{p_j q_j^m x^m}{1}$$

$$= \frac{\partial q_j^m}{\partial x}\frac{x^m}{q_j^m}\frac{p_j q_j^m}{x^m} - \frac{p_j q_j^m x^m}{(x^m)^2}$$

Solve for $\eta_j^x = \frac{\partial q_j^m}{\partial x}\frac{x^m}{q_j^m}$ to get:

(14) $$\eta_j^x = \frac{\partial w_j}{w_j \partial lnx^m} + 1$$

Applied to the LA/EASI model gives:

$$(15) \qquad e_m = \frac{\left[\sum_{r=0}^{R} b_{rj} y^{r-1} + \sum_{l=1}^{L} D_{lj} z_l + \sum_{k=1}^{J} B_{kj} p_k\right]}{w_j} + 1$$

Hicksian price and income elasticities from the LA/AIDS and LA/EASI models were generated using these formulas with the adding-up, homogeneity and symmetry constraints applied during estimation. These elasticities were compared across the two models to analyze any changes in the estimates due to inclusion of household characteristics such as age, education level, marital status and elderly status (age 60+ years old).

4.4 Data

FoodAPS is a nationally representative survey on the food acquisition and purchases of SNAP and non-SNAP participants collected by Mathematica Policy Research from April 2012 to January 2013. There were 4826 families surveyed. Of these all persons living in the household for that household's reference window were questioned. Some questions were directed at the principal respondent exclusively and others were asked of each member of the household. The respondents answered many questions regarding food consumed away from home and food purchases for consumption at home; these were recorded on separate blue and red survey instruments to make clear the division of at-home food consumption and food consumption away from home. The food away from home sources included hunting and fishing and also food pantries and family gatherings.

SNAP participants and low-income non-SNAP households under 130% of the federal poverty line were oversampled, but middle to high income earners were sampled as well, which made the survey nationally representative. The data were collected using a stratified random sampling technique with 50 primary sampling units (PSU) and 8 secondary sampling units (SSU) per PSU. It was necessary to use tertiary sampling units instead of SSU for some cases where it would be too expensive to create exhaustive lists of the SSU's to create a complete area listing of all the addresses in those SSU's.

The dependent variables were budget shares of nonalcoholic beverages. The six commodities used were milk, tea and coffee, fruit juices, juice drinks, carbonated soft drinks (CSD) and bottled water. The commodities were selected from the FoodAPS item level data.

Imputed prices were not used. Prices were computed per ounce and quantities purchased were also calculated in ounces. Descriptive statistics regarding the quantities purchased, the prices and the commodity shares are displayed in Tables 4-1, 4-2 and 4-3.

The unit of analysis is the household where data are reported by the household representative. This person was usually the principal food shopper for the household. The household was restricted to one family, but included members who would be in the household for the week the survey was being conducted. Nominal income was calculated as the total income for the household rather than per family unit. The demographic groups used in the linear EASI demand model were age, education level, elderly status and marital status. All estimates were generated using the FoodAPS household level weights. PROC MODEL would not allow strata and cluster variables; therefore, the ITSUR was performed without using those complex survey design variables.

The variables were created by using UPC descriptions for items on the FoodAPS products. These products were matched to IRI product descriptions such that the items had various levels of information regarding names of items, content and package size and what types of beverages were purchased. Some products had more information than others such as the percent whole wheat, types of fruit juice in blended juices, percentage of real fruit juice and if the juice was obtained from frozen concentrate.

The six commodities were formed specifically by using string text searches on the UPC descriptions and the text variables for the categories where these products were filed in the FoodAPS dataset. For example, the carbonated soft drinks (CSD) commodity was created using the string search "CARBONATED BEVERAGES" to select all UPC's attached to that category. The goal was to be as inclusive as possible to avoid having very small commodity shares that could be imprecisely estimated. However, it was important not to create such broad commodity groups that the results could not be compared to previous research on nonalcoholic beverage commodities and their included price and expenditure elasticities.

Table 4-1. Commodity Quantities Purchased using 2012 FoodAPS data

	Total Quantities Purchased (ounces)			
	Mean	s.e.	Minimum	Maximum
Milk	89.46	3.060	0.00	1216
Tea/coffee	18.13	2.715	0.00	1069
Fruit juices	29.54	1.952	0.00	1719
Juice drinks	29.86	1.860	0.00	960
Carbonated soft drinks	107.57	6.066	0.00	1960
Bottled water	66.62	5.358	0.00	4732

Note: Zero commodity quantity values represented no purchases of the given commodities

Table 4-2. Prices Paid/oz by Commodity Group using 2012 FoodAPS data

	Prices paid (per ounce)			
	Mean	s.e.	Minimum	Maximum
Milk	0.02	0.001	0.00	4.38
Tea/coffee	0.01	0.002	0.00	2.33
Fruit juices	0.01	0.001	0.00	0.56
Juice drinks	0.03	0.007	0.00	2.55
Carbonated soft drinks	0.02	0.002	0.00	3.16
Bottled water	0.01	0.001	0.00	0.36

Note: Zero price-paid values represented no purchases of the given commodities

Table 4-3. Nonalcoholic Beverages Budget Shares using 2012 FoodAPS data

	Nonalcoholic Beverages Budget Shares			
	Mean	s.e.	Minimum	Maximum
Milk	0.36	0.01	0.00	1.00
Tea/coffee	0.05	0.01	0.00	1.00
Fruit juices	0.12	0.01	0.00	1.00
Juice drinks	0.14	0.01	0.00	1.00
Carbonated soft drinks	0.25	0.01	0.00	1.00
Bottled water	0.08	0.01	0.00	1.00

Note: Commodity budget shares were equal to (total commodity expenditures)/(total expenditures per hhld) averaged across households

4.5 Methodology

In order to compare the models, the LA/EASI and LA/AIDS regressions were estimated using iterated seemingly unrelated regression and the results used to calculate own-price, cross-price and income elasticities. These estimates were compared across models to determine if the two methods produced statistically similar elasticity estimates. The models were also compared by Engel curve estimation to determine if using a more flexible form for the Engel curve produced statistically nonlinear curves. These curves were graphed using the coefficients of the linear and higher order real expenditures using the following expression:

$$w_j = \sum_{r=0}^{R} b_{rj} y^r + \sum_{l=1}^{L} (C_{lj} z_l + D_{lj} z_l y) + \varepsilon_j$$

where two reference persons were used for the benefit of comparison to demonstrate the features of the LA/EASI that are not possible using the LA/AIDS model. Prices in the Engel curves were held fixed at one, thus the logarithm of prices were held at zero. The error terms of the Engel curve were also held fixed at zero. Fixing the prices at 1 caused real income to be equal to nominal income in the Engel curve estimation. This made it easier to compare the Engel curve results of the logarithm of real income more intuitive given that on a daily basis

individuals spend income at nominal values. Elasticities were calculated using the LA/AIDS and LA/EASI models. LA/AIDS was used to calculate estimates for model coefficients which were then used in equations (4) and (5) to estimate Marshallian own-price, cross-price and expenditure elasticities. The LA/AIDS price elasticities were converted to Hicksian price elasticities using the Slutsky equation for elasticities

(16) $e_{jk}^{*} = e_{jk} + \eta_j^x w_k$, for goods j and k where η_j^x is the expenditure elasticity, e_{jk}^{*} is the Hicksian elasticity, e_{jk} is the Marshallian elasticity and w_k is the commodity budget share.

The LA/EASI Hicksian price elasticities were converted to Marshallian elasticities using the same Slutsky relation

(17) $e_{jk} = e_{jk}^{*} - \eta_j^x w_k$, for goods j and k where η_j^x is the expenditure elasticity, e_{jk}^{*} is the Hicksian elasticity, e_{jk} is the Marshallian elasticity and w_k is the commodity budget share.

4.6 Results

The Hicksian price elasticities for nonalcoholic beverages are relatively close to each other from the two linear approximations to the AIDS and EASI demand models and have the expected negative signs. Although they share the same signs, there are large differences in the cross-price estimates from the two models. Milk cross-price elasticities with tea/coffee and then with bottled water differ across models by ± 50 percent. The tea/coffee commodity cross-price elasticities with each commodity group differ by 10-30 percent. This is also the result for the fruit juices category with elasticity differences up to 50 percent. The juice drink commodity group has some Hicksian price elasticity differences between the two demand systems greater than 50 percent. Carbonated soft drinks elasticities are close relative to the two model systems, but differ the most for cross-price estimates with water followed by tea/coffee.

Cross-price elasticities that include juice drinks or carbonated soft drinks or fruit juice commodities are the most sensitive to model specification. Bottled water, milk and tea/coffee

cross-price elasticities with each other are also sensitive to the type of model specification, but not as much as with the other mentioned groups. Some of the cross-price elasticity estimates are larger from the linear AIDS model while other are larger with the linear EASI model estimates. Therefore, there was not a clear direction for the estimate differences.

Elderly status of 60+ years old had a significant positive effect on purchased budget shares of tea/coffee and water and a significant negative relationship with juice drinks shares. Marital status significantly increased budget shares of tea/coffee and decreased shares of fruit juices. Education significantly increased bottled water and tea/coffee budget shares and lowered milk and juice drinks shares. Surprisingly, higher levels of education showed statistically significant positive correlation with tea/coffee consumption, but negative correlation with milk consumption.

In Table 3-8, The LA/AIDS statistically significant income elasticity was 1.00 for carbonated soft drinks (CSD), milk was 0.99, and for juice drinks was 0.88, where the latter two commodities had income inelastic demand. All three commodities were necessities. The income elasticities for the other commodity shares were 1.05 for fruit juices, 1.12 for bottled water and 1.09 for tea/coffee. Each of these groups had income elastic demand and were therefore categorized as luxury beverages with bottled water as the most income sensitive beverage category.

Also in Table 3-8, the LA/EASI model generated statistically significant income elasticities for each of the six nonalcoholic beverage commodities. Tea/coffee, juice drinks and carbonated soft drinks had inelastic demand where carbonated soft drinks had the most inelastic demand and was the most necessary good. The income elastic commodities were milk, fruit juices and bottled water where the last commodity was the most price sensitive as with the LA/AIDS model.

Increases in the log of real income significantly increased purchases of tea/coffee and juice drinks with both models. Income increases decreased the shares purchased of milk, fruit juices, CSD and bottled water from the LA/EASI model. The LA/AIDS model also showed a

significant positive relationships between shares purchased and income for tea/coffee and a significant negative relationship for milk. There were positive, significant relationships between budget shares and log of real income generated from LA/AIDS for fruit juices, CSD and bottled water, but opposite signs from those estimated using LA/EASI.

Engel curve estimation for each of the commodity shares using the LA/AIDS model simply involved looking at the coefficients from the real expenditures variables. When analyzing the real expenditures coefficients, it could be seen that the commodity was treated as an inferior good or as a normal good for the entire length of the curve given that all LA/AIDS Engel curves were assumed to be linear and estimated as such. For example, milk and juice drinks were estimated to be necessities given that their Engel curve slopes were negative. Fruit juice, tea/coffee, bottled water and CSD were determined to be luxury goods from their positively sloped income expansion paths. Some results differed at various points along the Engel curves estimated using the LA/EASI demand system.

The Marshallian price elasticities were compared across models where the null hypothesis assumed the LA/AIDS elasticities were the true elasticities. This assumption was justified in that test as the LA/AIDS is the traditional model for demand system estimation and the LA/EASI is the newly developed model. The results are shown in Table 3-7. The column headed by juice drinks shows that the LA/EASI elasticities were not statistically significantly different from the LA/AIDS elasticities. The same was true for the fruit juices column. The milk column only had milk and tea/coffee cross-price elasticity that differed statistically across models. The tea/coffee price elasticity of -3.20 differed statistically from the -3.33 of the LA/EASI model. Every elasticity in the carbonated soft drinks column differed across the two models except for the carbonated soft drink and tea/coffee cross-price elasticity which was the same in both models. The bottled water elasticity was statistically different and so were the cross-price elasticities for bottled water with carbonated soft drinks and fruit juices. All of the Marshallian price elasticities were statistically significant except the LA/EASI bottled water and milk elasticity = 0.10.

The Marshallian expenditure elasticities for LA/AIDS were assumed to be the true elasticities in the null hypothesis for the t-tests across the two demand systems. The only statistically different expenditure elasticity was for carbonated soft drinks which was unit elastic in the LA/AIDS model and inelastic from the LA/EASI demand system. Allowing the interaction between the demographic variables and prices could have caused the estimate to be decreased using the LA/EASI model. The result that the other five commodities were not statistically different could be an indication that the LA/EASI is able to produce estimates similar to the estimates from a frequently used demand system that has been considered reliable for years by many researchers.

Table 4-4. Hicksian Price Elasticities LA/AIDS model

Hicksian Price Elasticities LA/AIDS model

	milk	tea/coffee	fruit juice	juice drinks	carbonated soft drinks	bottled water
milk	-1.08***	0.88***	0.60***	0.59***	0.53***	0.53***
tea/coffee	0.11***	-3.28***	0.20***	0.21***	0.12***	0.22***
fruit juice	0.22***	0.62***	-1.84***	0.30***	0.24***	0.37***
juice drinks	0.22***	0.63***	0.31***	-1.89***	0.24***	0.33***
carbonated soft drinks	0.39***	0.73***	0.49***	0.47***	-1.29***	0.50***
bottled water	0.15***	0.43***	0.24***	0.22***	0.16***	-2.01***

Note: * p<0.10, ** p<0.05, *** p<0.01

Table 4-5. Hicksian Price Elasticities LA/EASI model

Hicksian Price Elasticities LA/EASI model

	milk	tea/coffee	fruit juice	juice drinks	carbonated soft drinks	bottled water
milk	-1.06***	0.83***	0.59***	0.62***	0.52***	0.61***
tea/coffee	0.10***	-3.15***	0.20***	0.20***	0.11***	0.21***
fruit juice	0.21***	0.62***	-1.82***	0.33***	0.24***	0.35***
juice drinks	0.22***	0.59***	0.31***	-1.87***	0.24***	0.33***
carbonated soft drinks	0.38***	0.69***	0.48***	0.50***	-1.27***	0.67***
bottled water	0.15***	0.42***	0.23***	0.23***	0.16***	-1.14***

Note: * p<0.10, ** p<0.05, *** p<0.01

Table 4-6. Marshallian Price Elasticities using LA/AIDS model with 2012 FoodAPS data

Marshallian Price Elasticities LA/AIDS model

	milk	tea/coffee	fruit juice	juice drinks	carbonated soft drinks	bottled water
milk	-1.43***	0.49***	0.22***	0.28***	0.17***	0.13***
tea/coffee	0.06***	-3.33***	0.16***	0.17***	0.08***	0.17***
juice drinks	0.10***	0.50***	0.18***	-2.00***	0.12***	0.19***
carbonated soft drinks	0.13***	0.44***	0.22***	0.24***	-1.55***	0.21***
bottled water	0.06***	0.34***	0.15***	0.14***	0.07***	-2.11***

Note: * p<0.10, ** p<0.05, *** p<0.01

Table 4-7. Marshallian Price Elasticities using LA/EASI model with 2012 FoodAPS data

Marshallian Price Elasticities LA/EASI model

	milk	tea/coffee	fruit juice	juice drinks	carbonated soft drinks	bottled water
milk	-1.44***	0.41***	0.26***	0.32***	0.25***†	0.16**
tea/coffee	0.05***†	-3.20***†	0.16***	0.16***	0.08***	0.15***
fruit juice	0.08***	0.46***	-1.97***	0.22***	0.14***†	0.19***
juice drinks	0.09***	0.44***	0.20***	-1.97***	0.14***†	0.18***
carbonated soft drinks	0.11***	0.38***	0.24***	0.28***	-1.47***†	0.33***
bottled water	0.06***	0.32***	0.15***	0.16***	0.09***†	-1.25***†

Note: * p<0.10, ** p<0.05, *** p<0.01, † Statistically significantly different from the Marshallian LA/AIDS estimates for at least p<0.10

Table 4-8. Comparing Marshallian Expenditure Elasticities for LA/AIDS and LA/EASI

	Marshallian Income Elasticities	
	LA/AIDS	LA/EASI
milk	0.99***	1.07***
tea/coffee	1.09***	0.91***
fruit juices	1.05***	1.22***
juice drinks	0.88***	0.83***
carbonated soft drinks	1.00***	0.76***†
bottled water	1.12***	1.43***

Note: * p<0.10, ** p<0.05, *** p<0.01, † Statistically significantly different from the Marshallian LA/AIDS estimates for at least p<0.10

Engel curves were used to illustrate changes in the budget shares for changes in monthly real income for two reference groups. Lewbel and Pendakur (2009) and Zhen et al. (2013) also used a reference person in each study to generate Engel curves. The first reference person in this study was a married, 30 year old high school to college graduate as the highest level of education (referred to as Ref1). The second reference person had the same educational attainment and marital status, but was 60 years of age (referred to as Ref2). These reference persons were chosen to hold education and marital status constant, but change elderly status to observe the impact on budget share changes by income across elderly status. The monthly log average income range for the reference person was $2.86 to $9.10. This range for the second reference person was $0.00 to $9.38, so a longer income range for the second reference person compared to the first. The range of tea/coffee shares in Figure 4-1 for Ref1 is 0.06 to 0.17 and 0.10 to 0.31 for Ref2 in Figure 4-2. There is some overlap in the budget share ranges and the range is wider for Ref2 than for Ref1.

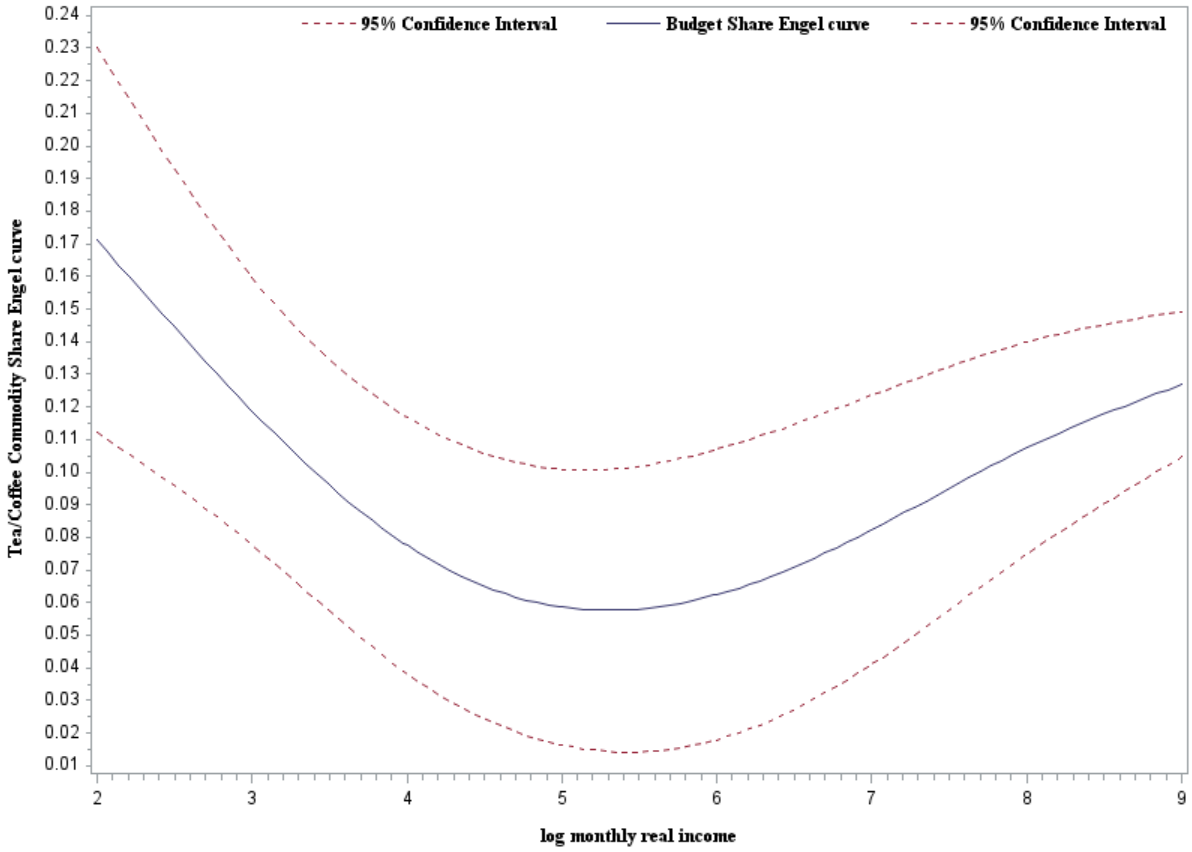

Figure 4-1. LA/EASI Estimation of Tea/coffee Engel Curve for Ref 1

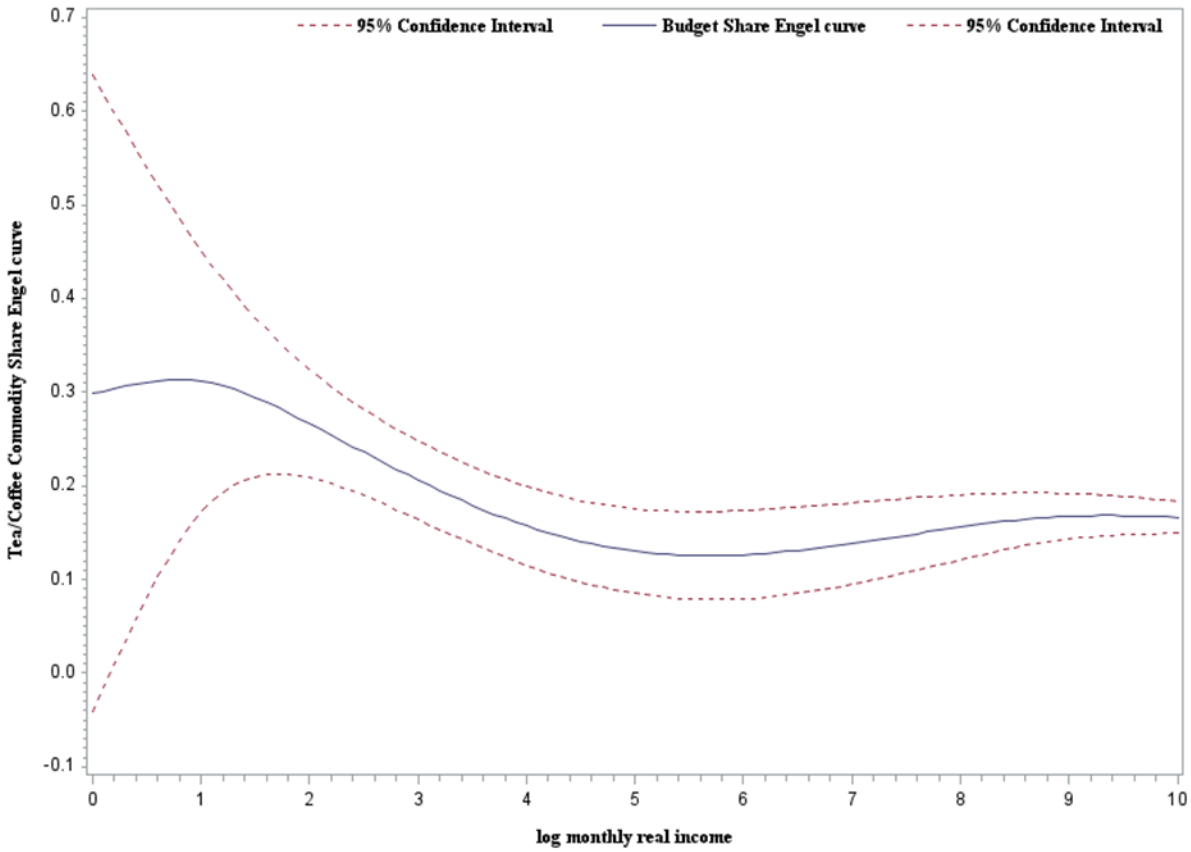

Figure 4-2 LA/EASI Estimation of Tea/coffee Engel Curve for Ref 2

The juice drinks Engel curve budget shares in Figures 4-3 and 4-4 range from 0.17 to 0.24 for Ref1 and for Ref2 is from 0.10 to 0.40. Again, the budget share range of the Engel curve is wider for Ref2 than for Ref1. The carbonated soft drinks commodity has a budget share range from 0.21 to 0.29 in Figure 4-5 compared to 0.20 to 0.88 for Ref2 in Figure 4-6. The range where the monthly log income range for Ref2 matches Ref1 had a budget share range between 0.20 and 0.24, which was narrower than the budget share ranges for Ref1. The fruit juices budget share ranges were 0.10 to 0.21 and 0.20 to 0.22 for Ref1 and Ref2 respectively in Figures 4-7 and 4-8 where the income ranges were the same across reference persons.

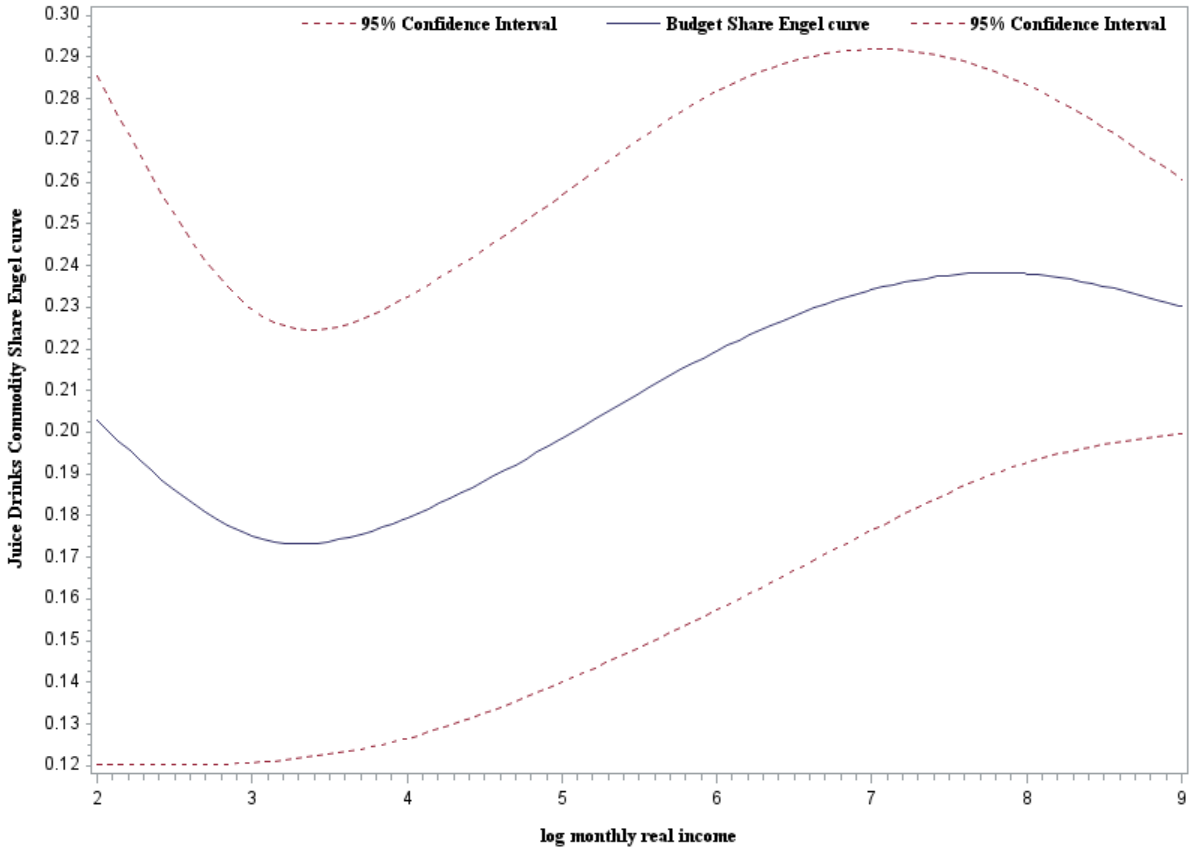

Figure 4-3. LA/EASI Estimation of Juice Drink Engel Curve for Ref 1

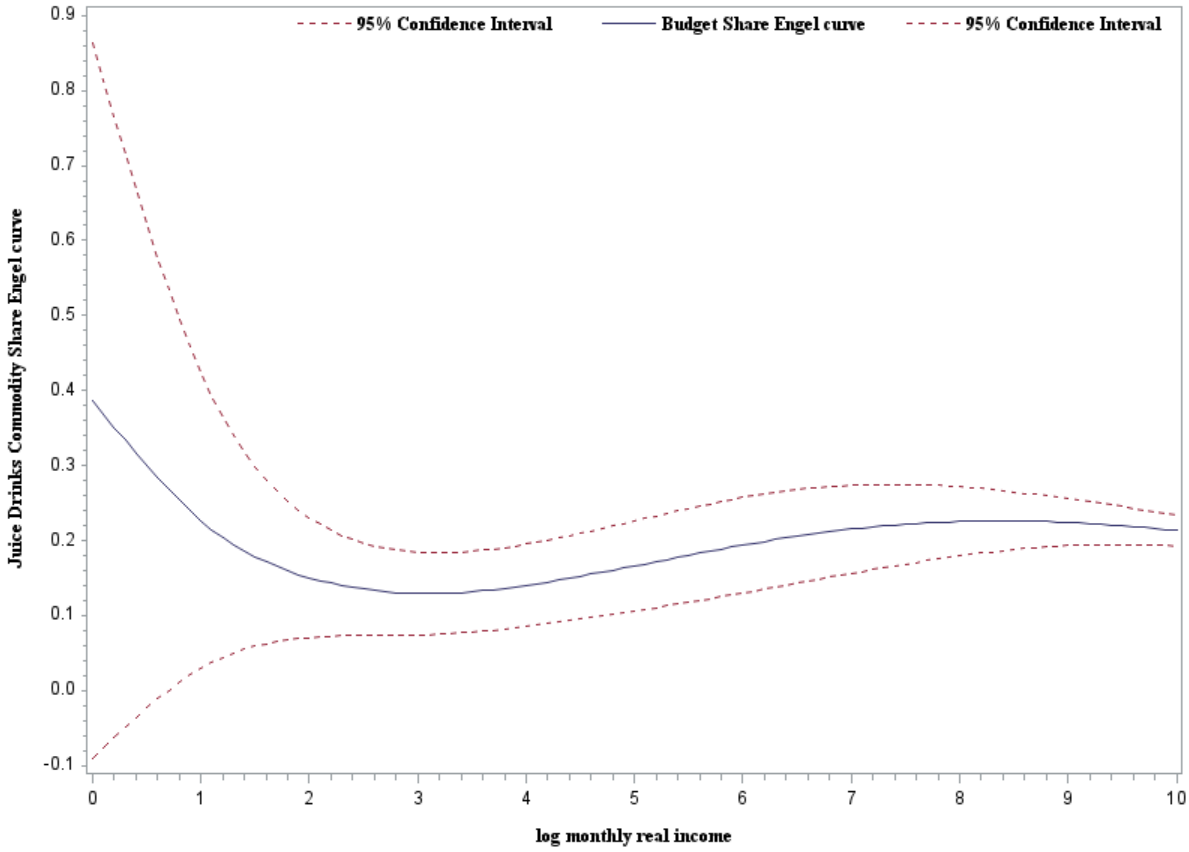

Figure 4-4 LA/EASI Estimation of Juice Drink Engel Curve for Ref 2

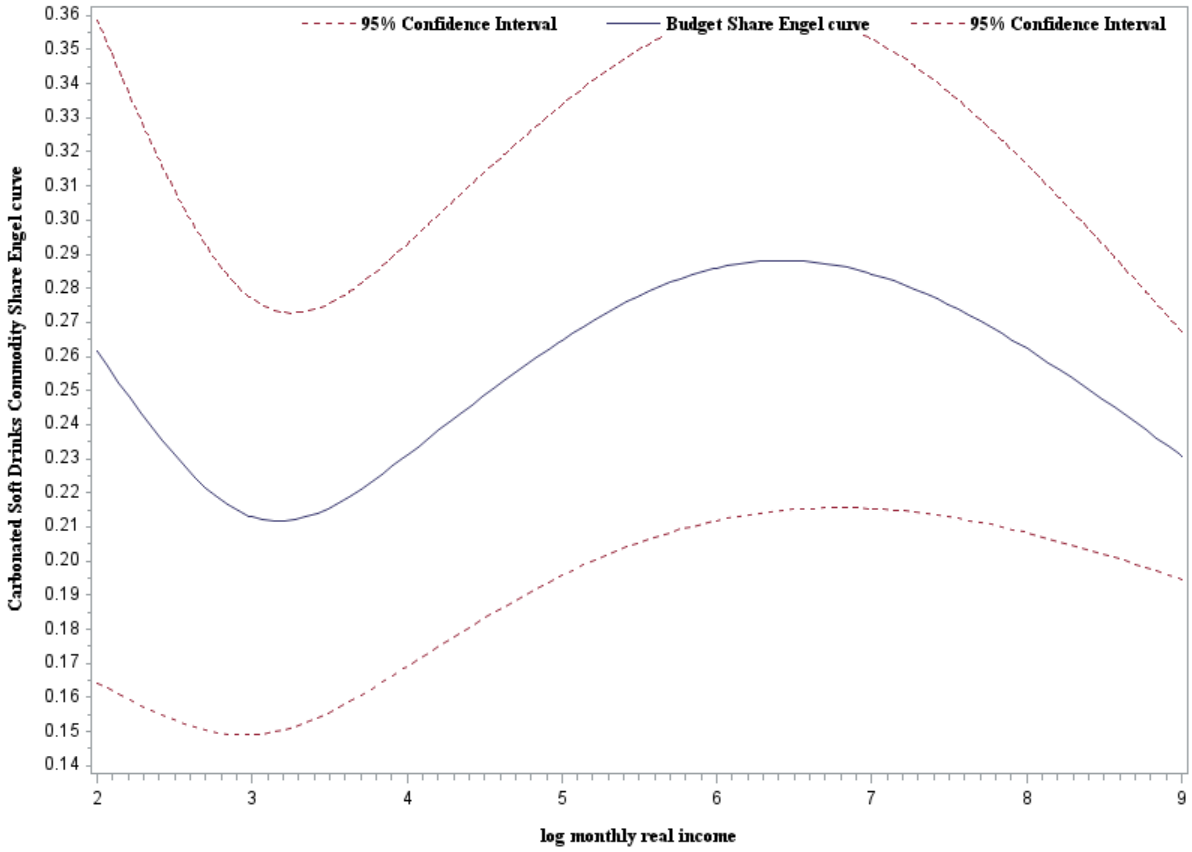

Figure 4-5. LA/EASI Estimation of Carbonated Soft Drinks Engel Curve for Ref 1

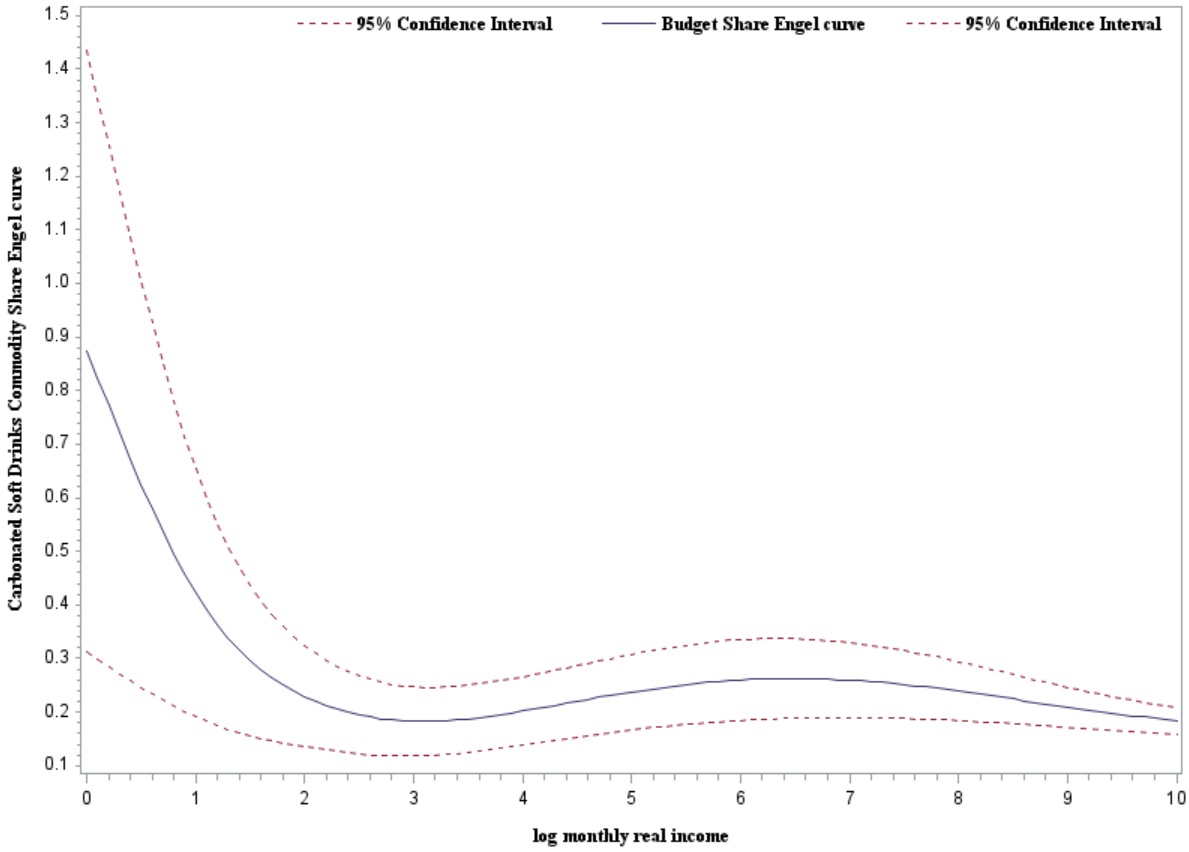

Figure 4-6 LA/EASI Estimation of Carbonated Soft Drinks Engel Curve for Ref 2

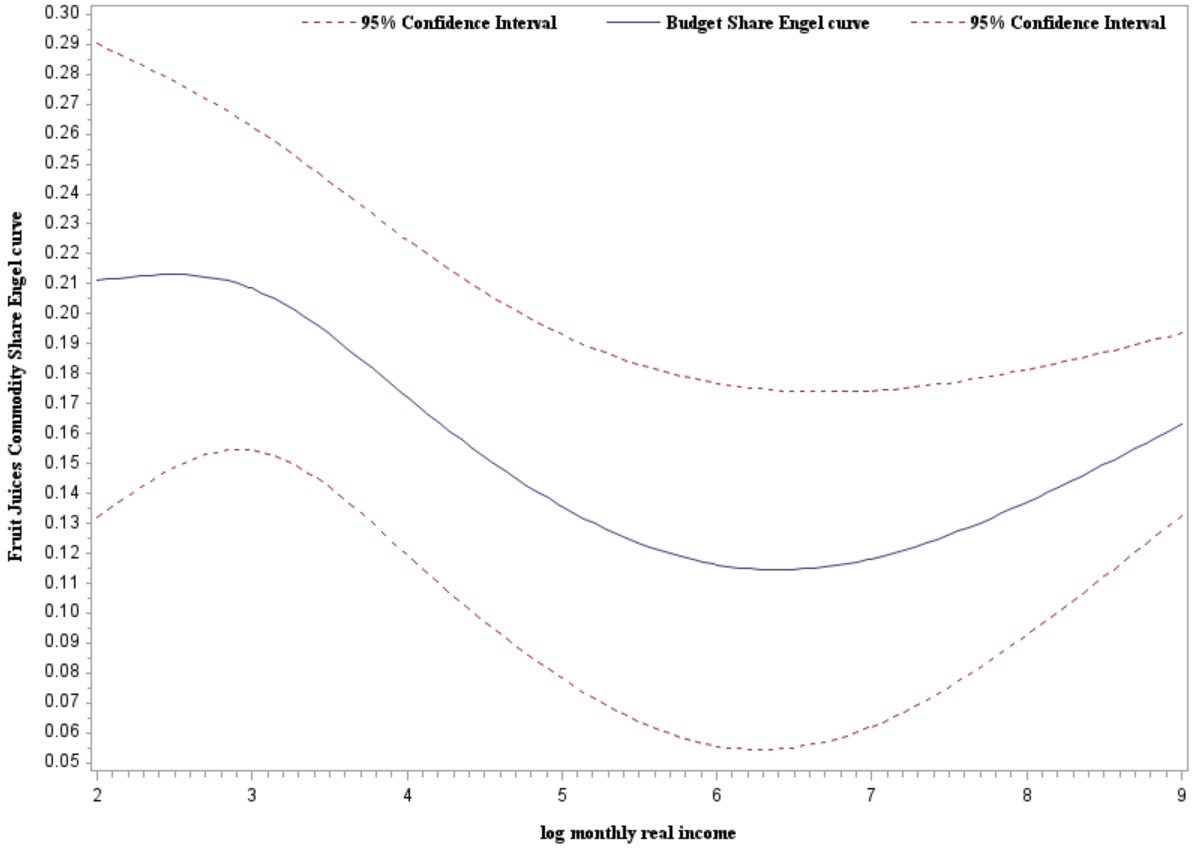

Figure 4-7 LA/EASI Estimation of Fruit Juices Engel Curve for Ref 1

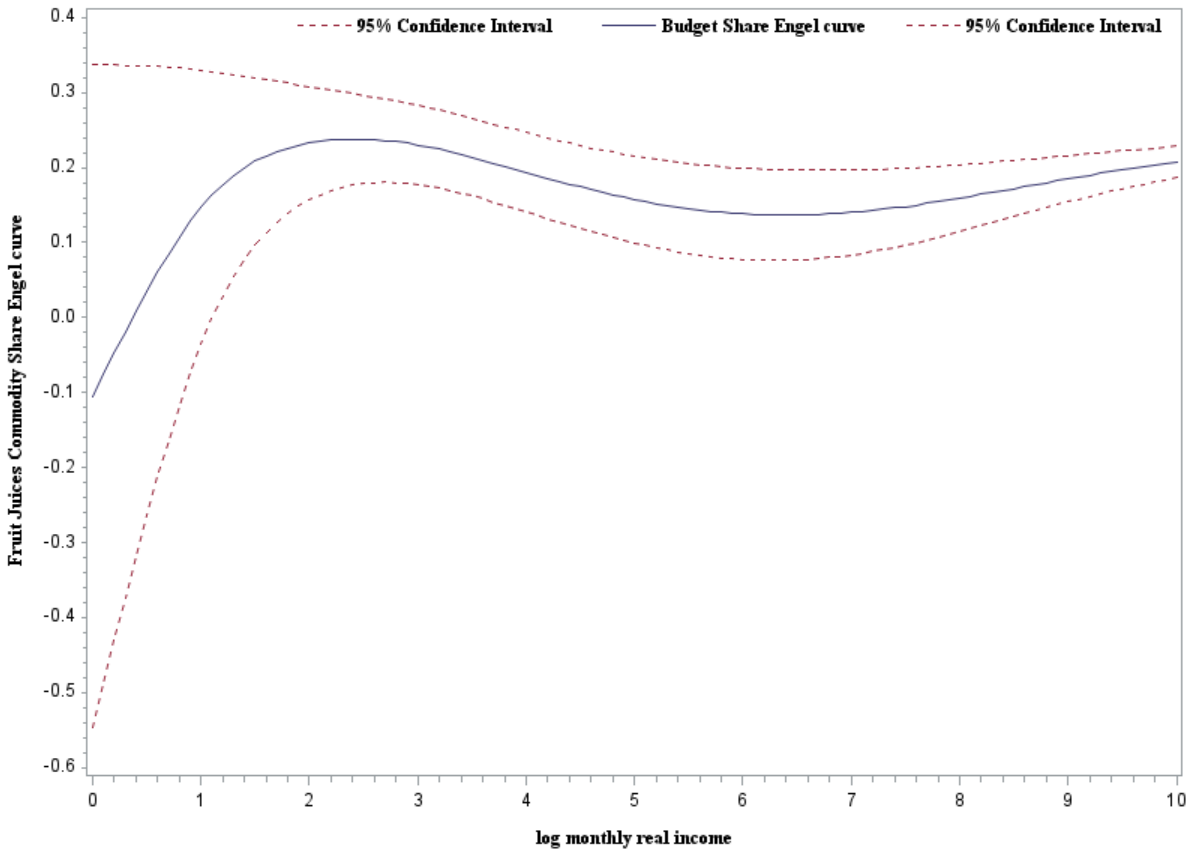

Figure 4-8 LA/EASI Estimation of Fruit Juices Engel Curve for Ref 2

Milk budget shares for Ref1's Engel curve ranged from 0.10 to 0.19 shown in Figure 4-9. The range for Ref2 in Figure 4-10 where income ranges matched for both reference persons was 0.10 to 0.20 in Figure 3-10, which was almost the same as for Ref1. Thus, elderly status did not affect milk demand as a share of total expenditures as incomes increased. Bottled water budget shares in Figures 4-11 and 4-12 were from 0.09 to 0.20 for Ref1 and 0.10 to 0.26 for Ref2 where the two reference persons had monthly log incomes ranging from $2 to $9. The elderly status had a higher upper bound budget share due to a rising Engel curve for Ref2 where the curve falls for Ref1's Engel curve. So the younger reference person treats bottled water as a luxury but then as a necessity as incomes rise, but the elderly reference person treats the commodity as a luxury good even as incomes rise. The older person might not be as accustomed to purchasing bottled water, might prefer tap water to bottled water or might not

believe that it makes sense to purchase a commodity that can be accessed at the tap or at fountains among other possible taste and preferences reasons. It was useful, therefore, to observe these differences in the Engel curves for different cohorts made possible by the LA/EASI demand model. At higher income levels for the elderly reference person, the Engel curves were almost horizontal lines but slightly downward sloping except for fruit juices and bottled water. This indicates normal goods for all commodities but fruit juices and bottled water with income elasticities close to 1 for the elderly reference person.

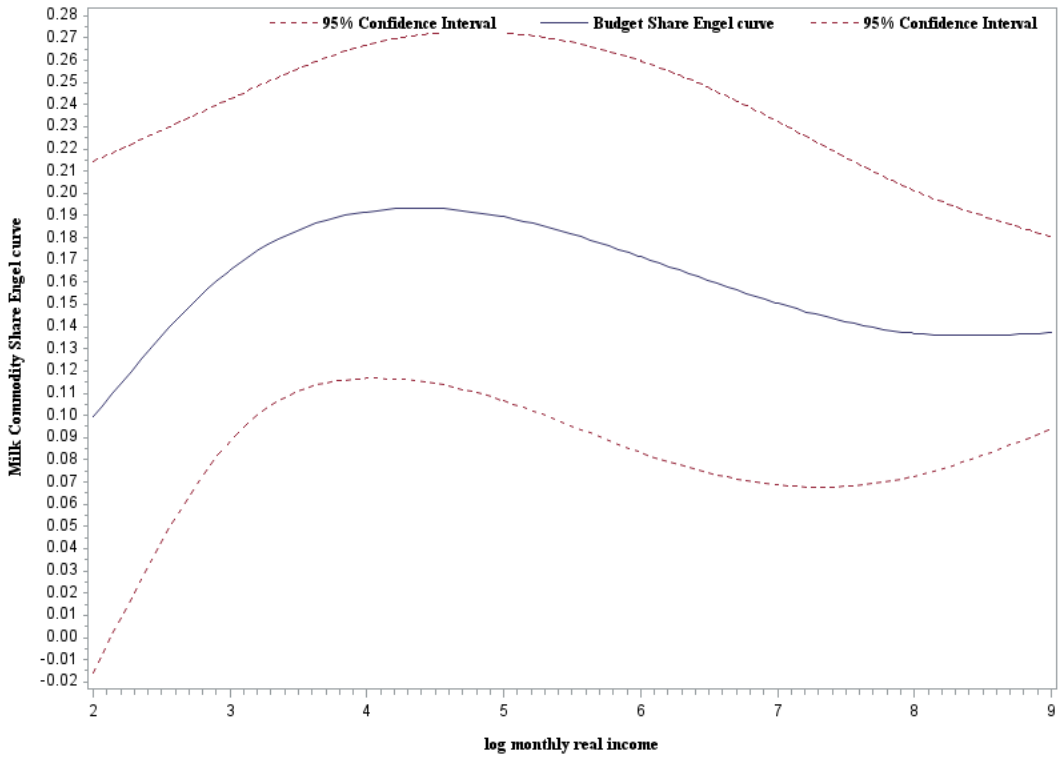

Figure 4-9. LA/EASI Estimation of Milk Engel Curve for Ref 1

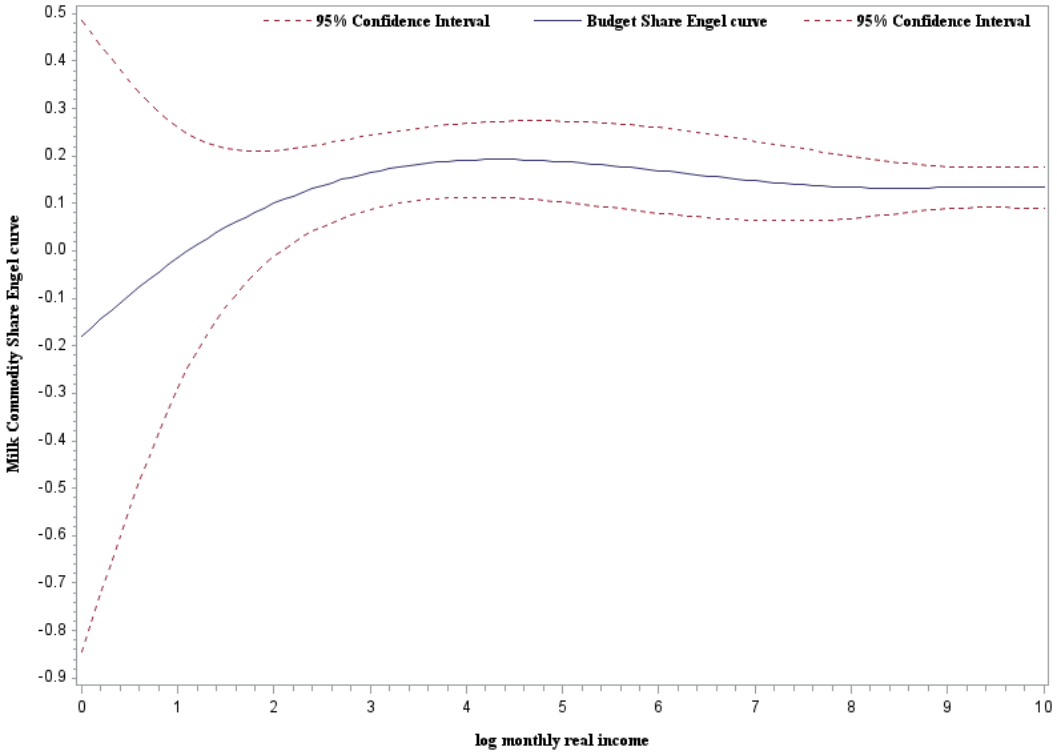

Figure 4-10 LA/EASI Estimation of Milk Engel Curve for Ref 2

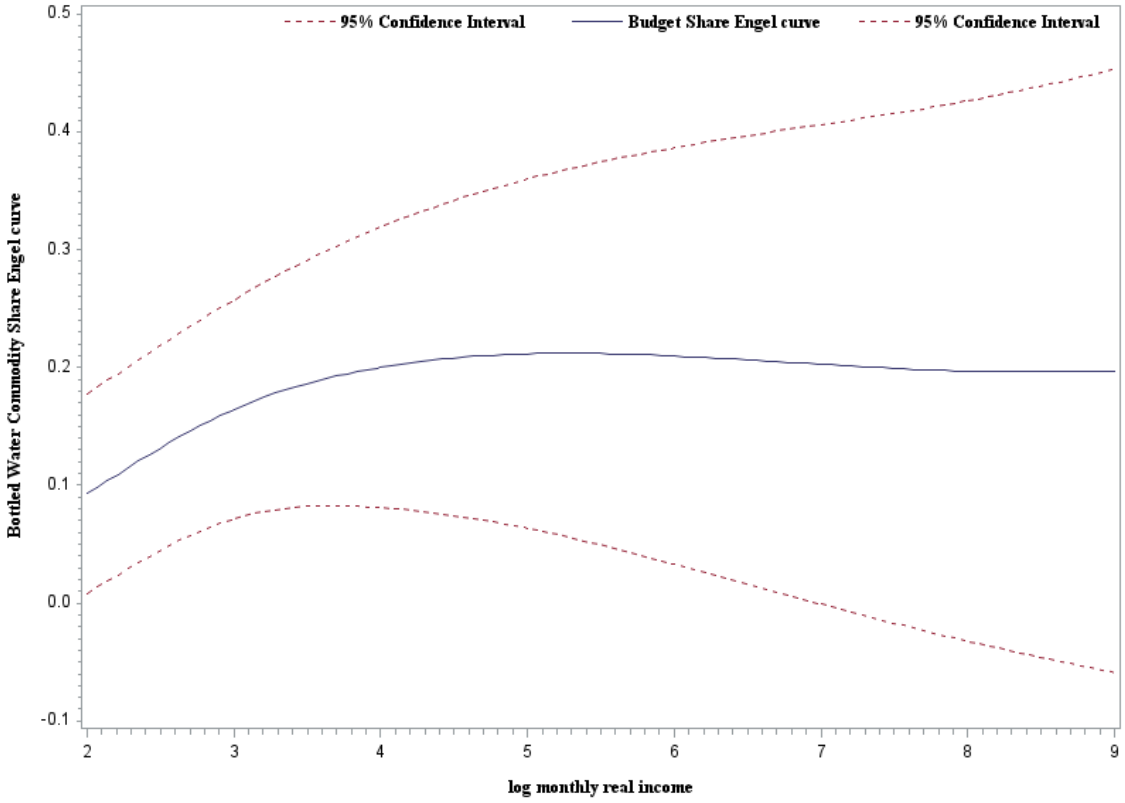

Figure 4-11. LA/EASI Estimation of Bottled Water Engel Curve for Ref 1

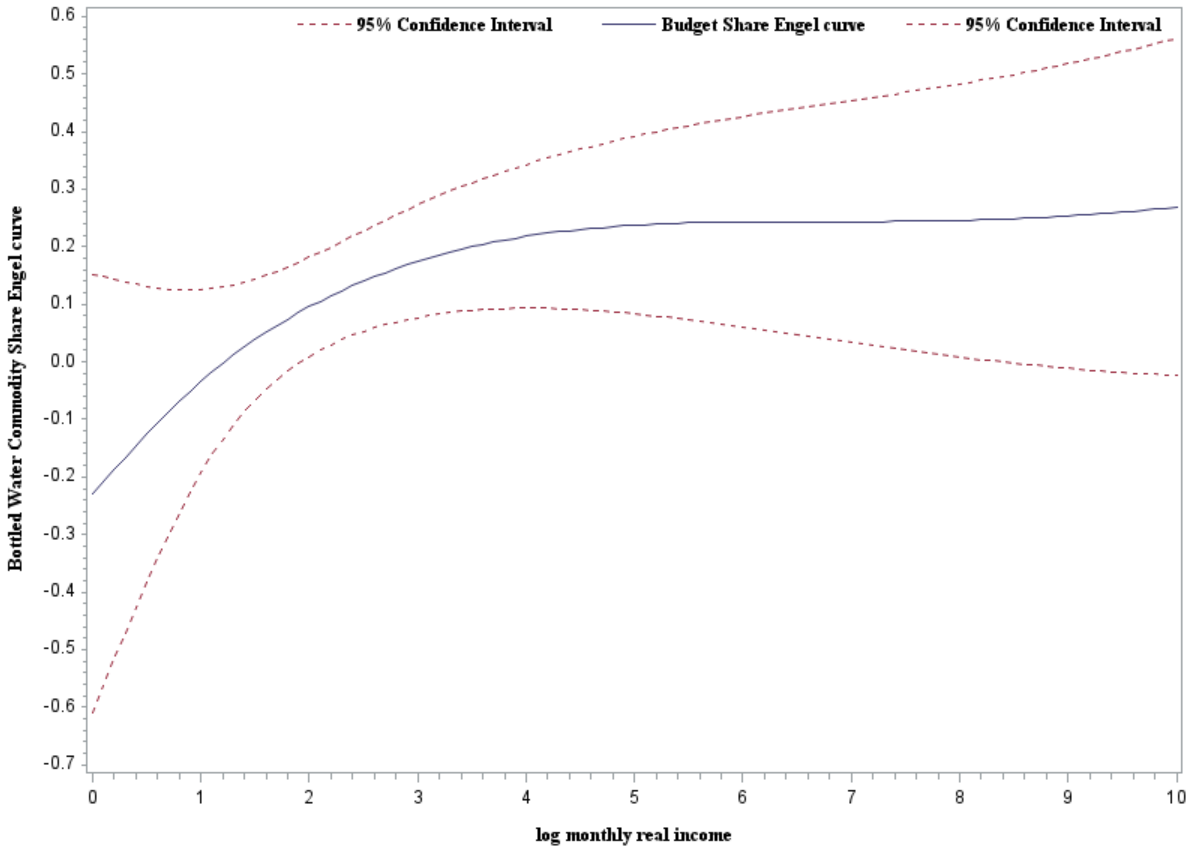

Figure 4-12 LA/EASI Estimation of Bottled Water Engel Curve for Ref 2

Estimation of Marshallian price elasticities allowed for comparisons to other research using the LA/EASI models. Zhen et al. (2013) used a censored LA/EASI model to estimate median Marshallian price elasticities with all incomes included. Four of their commodity groups were very close in definition to commodities in this paper. Regular CSD, 100% juice, juice drinks and bottled water were compared to the companion groups in this study. Regular CSD of the Zhen et al. (2013) had a price elasticity of -1.035 compared to this study's LA/EASI Marshallian CSD price elasticity estimate of -1.47 (shown in Table 4-8). 100 percent juice from their study had a price elasticity value of -1.506 compared to this study's fruit juice elasticity estimate of -1.97. Their price elasticity estimate for juice drinks was -1.192 compared to the -1.97 estimate from this study.

In both studies, these three commodities had price elastic demand, but the current study has more elastic price elasticities than the Zhen et al. (2013) study. The fourth commodity, bottled water, had a value of -1.703 elasticity of demand compared to -1.25 price elasticity in this study. Their study showed a higher elastic demand than in the current research for bottled water. The differences in price estimates could be based in their censored model where the current study did not use a censored model and did not censor any households from the FoodAPS dataset. The cross-price elasticities for those commodities differed across the two studies with substitutes as the outcome for those four comparison commodities for the current study, but complements for bottled water and regular CSD in the Zhen et al. (2013) paper.

The LA/AIDS model is nested in the LA/EASI model. Therefore an F test was performed where the null hypothesis was that the LA/AIDS was the better model for the nonalcoholic beverage demand system for the dataset compared to the LA/EASI. In other words, the additional parameters of the LA/EASI would not contribute in explaining the variation in the budget share dependent variables. So under the null hypothesis, the LA/EASI would not be an improvement on the LA/AIDS model. The critical F- statistic is $F_{0.01,185,2620} = 1$. The actual F value from the estimation of the LA/AIDS as the restricted case of the LA/EASI as the complete model is 1.18 > 1, so $F_{actual} > F_{0.01,185,2620}$. This is evidence to reject the null hypothesis and conclude that LA/EASI is an improvement on the LA/AIDS model for explaining consumer demand for nonalcoholic beverages given the six commodities chosen in this study as applied to the 2012 FoodAPS data.

4.7 Conclusions

Conducting estimation of price and income sensitivity of consumers for nonalcoholic beverages using two estimation methods allowed for comparison of price and income elasticities based on the different assumptions of the models. The linear approximation to the AIDS assumed linear Engel curves; therefore, only one type of income elasticity of demand was possible per commodity. On the other hand, the LA/AIDS model did not indicate the ranges of inelastic and elastic regions of income sensitivity.

In addition, it was not possible to allow for the representative consumer to vary by specific socio-demographic characteristics with the LA/AIDS model form. If the consumer represented different racial and economic conditions, there was a need to use LA/AIDS with demographic translation to model household characteristics as described by Pollak and Wales (1981) where the original LA/AIDS is replaced by parameters that depend on the demographic variables. Interpretation of the common values in the results as reflecting the demand of the "equivalent adults" might not relate well with the average consumer. Also, there are very specific conditions under which the use of translation is valid. Lastly, the use of translation is not as thorough in including household characteristics as with the LA/EASI method that allows interaction of demographic characteristics with prices and real expenditures.

Engel curves had statistically different forms between the two models for tea/coffee, fruit juices, carbonated soft drinks and juice drinks. The Engel curves were each linear for those commodities when using the LA/AIDS model. Fruit juices, carbonated soft drinks and bottled water had elastic income demand from LA/AIDS estimation which just slightly indicated those commodities as luxury beverage items and not necessities. Using LA/EASI, each of those commodities had inelastic income demand and therefore were viewed as necessities.

Income elasticities were determined at a point, so they could be valid if considering elasticities at the given averages of the variable distributions for income, price and quantities purchased. The Engel curves can show more information regarding different income sensitivity as the income elasticities can vary over a nonlinear curve. Using LA/EASI, fruit juices had elastic income demand before the inflection point of the curve, then inelastic income demand until the maximum point of the curve. Lastly, the LA/EASI model had negative income elasticity beyond the maximum total beverage expenditure of the curve. This showed that income sensitivity changed depending on the amount of income spent per week on beverages. This disclosed much more detail regarding reactions of households to income changes than the constant upward-sloping Engel curve generated by the LA/AIDS model that was always income elastic.

Carbonated soft drinks had an elastic demand for one section of the Engel curve beginning at zero dollars spent on beverages. After the inflection point, the income elasticity was inelastic, indicating that carbonated soft drinks were a necessity for those who spent larger amounts of income weekly for food-at-home beverages. Unlike the case with the Engel curve from the LA/AIDS model that remains at elastic income demand, the LA/EASI Engel curve shows that there is a point where the elasticity becomes inelastic and then eventually negative with larger amounts of beverage expenditures. Therefore, it is essential to have a demand system that can include demographic characteristics and not just via scaling factors such as with translation for the LA/AIDS model.

It seems that both demand systems are sensitive to very low budget shares. Bottled water and tea/coffee had shares of 0.08 and 0.05 respectively, each with standard errors of 0.01. LA/AIDS estimated 1.28 as the bottled water income elasticity and 1.63 as the income elasticity for tea/coffee. These were the largest income elasticities from that model and for the lowest budget shares. It might be reasonable to assume that tea/coffee and bottled water are not necessities generally, but income elasticities for alcoholic beverages tend to be around 1.90 to almost 3.00. It is not common for bottled water, coffee and tea to be considered high-end beverages. It would be more reasonable to think of tea, coffee and bottled water as necessities for those who purchase them.

However, the estimates for income elasticity for tea/coffee and bottled water were just as surprising from the LA/EASI demand system. Tea/coffee was estimated to be inferior with -0.04 income elasticity and bottled water was estimated at 0.46. Both of those figures are lower than the typical estimates for those commodity groups. Tea/coffee is more likely a necessity than an inferior commodity and bottled water is probably more elastic than 0.46 even if it is a necessity for very mobile individuals who like to stay hydrated. Estimates of income elasticity were much closer across the two models for the category groupings with higher budget shares.

REFERENCES

Anselin, L. (1980). Estimation methods for spatial autoregressive structures. Regional Science Dissertation & Monograph Series, Program in Urban and Regional Studies, Cornell University, (8).Anselin, L. (1988). Spatial econometrics: methods and models (Vol. 4). Springer Science & Business Media.

Anselin, L. (2001). Spatial econometrics. *A companion to theoretical econometrics*, *310330*.

Anselin, L. (2005). Spatial Regression Analysis in R–A Workbook. *Urbana*, *51*, 61801.

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional science and urban economics, 26*(1), 77-104.

Argo, J. J., Dahl, D. W., & Morales, A. C. (2008). Positive consumer contagion: Responses to attractive others in a retail context. *Journal of Marketing Research*, *45*(6), 690-701.

Asche, F., & Wessells, C.R. (1997). On Price Indices in the Almost Ideal Demand System. *American Journal of Agricultural Economics79*, 1182-1185.

Becker, G., M. Grossman, and K. M. Murphy, "An Empirical Analysis of Cigarette Addiction," American Economic Review, (1994), 396–418.

Bernheim, B. D. (1994). A theory of conformity. *Journal of political Economy*, 841-877.

Bivand, R. (2011). After 'Raising the Bar': Applied Maximum Likelihood Estimation of Families of Models in Spatial Econometrics. *NHH Dept. of Economics Discussion Paper* 22 .

Bronnenberg, B. J., Kruger, M. W., & Mela, C. F. (2008). Database paper-The IRI marketing data set. *Marketing Science*, *27*(4), 745-748.

Burridge, P. (2011). A research agenda on general-to-specific spatial model search. *Investigaciones regionales*, (21), 71-90.

Elhorst, J. P. (2001). Dynamic models in space and time. *Geographical Analysis*, *33*(2), 119-140.

Fletcher, J. M., Frisvold, D., & Tefft, N. (2010). Can soft drink taxes reduce population weight?. *Contemporary Economic Policy, 28*(1), 23-35.

Florax, R. J., Folmer, H., & Rey, S. J. (2003). Specification searches in spatial econometrics: the relevance of Hendry's methodology. *Regional Science and Urban Economics*, *33*(5), 557-579.

Genton, M. G. (2000). The correlation structure of Matheron's classical variogram estimator under elliptically contoured distributions. *Mathematical geology*, *32*(1), 127-137.

Glass, A. J., Kenjegalieva, K., & Sickles, R. (2012). The economic case for the spatial error model with an application to state vehicle usage in the U.S. working paper.

Guthrie, J., & Newman, C. (2013). Eating Better at School: Can New Policies Improve Children's Food Choices?. United States Department of Agriculture. Retrieved from www. ers. usda. gov.

Hendry, D.F. Predictive failure and econometric modelling in macroeconomics: the transactions demand for money, in: P. Ormerod (Ed.), Economic Modelling, Heinemann, London (1979)

Kelejian, H. H., & Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, *17*(1), 99-121.

Kelejian, H. H., & Prucha, I. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International economic review*,*40*(2), 509-533.

Kelejian, H.H. & Prucha, I.R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*,*157*, 53-67.

Korniotis, G. M. (2010). Estimating panel models with internal and external habit formation. *Journal of Business & Economic Statistics*, *28*(1), 145-158.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics* 443-477.

Leibenstein, H. (1950). Bandwagon, snob, and Veblen effects in the theory of consumers' demand. *The Quarterly Journal of Economics*, 183-207.

Pace, R. K., & LeSage, J. P. (2009). A sampling approach to estimate the log determinant used in spatial likelihood problems. *Journal of Geographical Systems*, *11*(3), 209-225.

Lewbel, A. (2006). Engel curves. *The New Palgrave Dictionary of Economics*,*2*.

Lewbel, A. and K. Pendakur (2009). Tricks with Hicks: The EASI Demand System, *American Economic Review, 99* (3), 827- 863.

Mancino, L., & Kuchler, F. (2012). Demand for whole-grain bread before and after the release of Dietary Guidelines. *Applied economic perspectives and policy*, *34*(1), 76-101.

Nickell, S.(1981). Biases in dynamic models with fixed effects, *Econometrica: Journal of the Econometric Society* 1417-1426.

Ord, J. K., & Getis, A. (2001). Testing for local spatial autocorrelation in the presence of global autocorrelation. *Journal of Regional Science*, *41*(3), 411-432.

Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, *70*(349), 120-126.

Pollak, R. A., & Wales, T.J. (1981).Demographic variables in demand analysis. *Econometrica: Journal of the Econometric Society* 1533-1551.

Rahkovsky, Ilya & Snyder, S.. *Food Choices and Store Proximity*, ERR-195, U.S. Department of Agriculture, Economic Research Service, September 2015.

Rey, S. J., & Anselin, L. (2010). PySAL: A Python library of spatial analytical methods. In *Handbook of applied spatial analysis* (pp. 175-193). Springer Berlin Heidelberg.

Todd, J.E., Mancino,L. Leibtag, E. & Tripodo, C. 2010. Methodology Behind the Quarterly Food-at-Home Price Database, Technical Bulletin No. 1926, U.S. Department of Agriculture, Economic Research Service.

Yang, S., & Allenby, G. M. (2003). Modeling interdependent consumer preferences. *Journal of Marketing Research*, *40*(3), 282-294.

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. *American journal of agricultural economics*, aat049.

Appendix A

Table A 1. Sugar Sweetened Beverages 2008 IRI data

Variable	*Estimate*
1-month lead	0.15
2-month lead	0.07
1-month lag	0.24
2-month lag	0.17
Widowed	-34.39
Divorced/Separated	-7.20
Single	-33.25
August	-108.30
September	68.10
October	-13.67
November	-45.92
December	29.23
Household Size	39.27
Household Income	-8.17
Black	-54.61
Asian	-8.04
Other	-4.26
Hispanic	9.78
Relatively urban	-6.25
Relatively rural	14.09
Very rural	22.12
Northeast	33.24
South	16.53
West	25.04
Current avg price	-554.52
Constant	325.50
N	55032
R^2	0.49
δ	0.65
δ^2	0.39
θ_1	0.24
θ_2	0.17
E_d	-0.03

** p<0.10, ** p<0.05, *** p<0.01*

Table A 2. Sugar Sweetened Beverages 2012 IRI data

Variable	*Estimate*
1-month lead	0.03
2-month lead	0.00
1-month lag	0.10
2-month lag	0.11
Widowed	37.45
Divorced/Separated	14.25
Single	2.54
August	-64.80
September	25.08
October	-4.46
November	-25.55
December	0.42
Household Size	86.97
Household Income	174.55
Black	204.71
Asian	240.22
Other	267.78
Hispanic	392.16
Relatively urban	561.86
Relatively rural	-137.78
Very rural	-125.77
Northeast	-10.24
South	83.35
West	7.30
Current avg price	31.23
Constant	64.80
N	76.68
R^2	76.69
δ	37.32
δ^2	-534.43
θ_1	297.59
θ_2	40032.00
E_d	0.28

* p<0.10, ** p<0.05, *** p<0.01

Table A 3 Fruit Juices 2008 IRI data

Variable	*Estimate*
1-month lead	0.21**
2-month lead	0.16**
1-month lag	0.07
2-month lag	0.18**
Widowed	-7.76
Divorced/Separated	-2.21
Single	-6.06
August	-10.52*
September	-6.38
October	-20.37***
November	-22.21***
December	-6.57
2-person household	11.28**
3-person household	15.49**
4-person household	28.64***
5-person household	49.70***
6-person household	57.80***
7-person household	73.54
8-person household	70.56**
Household Income	1.13*
Black	17.63***
Asian	8.64
Other	11.3
Hispanic	-6.15
Relatively urban	-5.95*
Relatively rural	-8.67*
Very rural	-3.33
Northeast	-5.75
South	-2.55
West	4.13
Current avg price	-789.90***
Constant	160.59***
N	31026
R^2	0.51
δ	3.14
δ^2	0.88
θ_1	0.07
θ_2	0.18
E_d	-0.16

** p<0.10, ** p<0.05, *** p<0.01*

Table A 4 Fruit Juices 2012 IRI data

Variable	*Estimate*
1-month lead	0.14
2-month lead	0.10
1-month lag	0.07
2-month lag	0.14
Widowed	-12.97
Divorced/Separated	8.80
Single	13.17
August	-10.55
September	-10.98
October	-15.45*
November	-7.83
December	-5.89
2-person household	11.65
3-person household	19.11**
4-person household	47.05***
5-person household	48.57**
6-person household	27.73
7-person household	38.10*
8-person household	5.27
Household Income	1.49
Black	25.41
Asian	10.44
Other	7.67
Hispanic	-19.14
Relatively urban	-2.79
Relatively rural	-6.30
Very rural	-6.21
Northeast	-7.70
South	-8.11
West	0.87
Current avg price	-852.29***
Constant	211.32**
N	11496
R^2	0.45
δ	1.94
δ^2	0.73
θ_1	0.74
θ_2	0.07
E_d	0.14

* p<0.10, ** p<0.05, *** p<0.01

Table A 5 Sugar-Sweetened Beverages Myopic Model 2008 IRI data

Variable	*Estimate*
1-month lag	0.34***
2-month lag	0.27***
Widowed	-19.94
Divorced/Separated	6.44
Single	-11.54
August	-93.41***
September	67.86***
October	-26.75
November	-76.04***
December	-23.45
2-person household	78.43***
3-person household	147.47***
4-person household	167.85***
5-person household	198.02***
6-person household	201.77***
7-person household	201.24***
8-person household	346.21***
Household Income	-9.40***
Black	-63.12***
Asian	10.35
Other	-7.45
Hispanic	14.4
Relatively urban	-5.11
Relatively rural	13.04
Very rural	21.02
Northeast	35.98***
South	19.82*
West	32.44**
Current avg price	605.05***
Constant	359.19***
N	55032
R^2	0.45
θ_1	0.34***
θ_2	0.27***
E_d	-0.03***

Table A 6 Sugar-Sweetened Beverages Myopic Model 2012 IRI data

Variable	*Estimate*
1-month lag	0.11***
2-month lag	0.11***
Widowed	1.10
Divorced/Separated	-19.77
Single	-24.69
August	-61.81***
September	27.42*
October	-3.16
November	-23.11
December	0.40
2-person household	102.33***
3-person household	198.26***
4-person household	237.43***
5-person household	268.41***
6-person household	301.25***
7-person household	420.35***
8-person household	596.93***
Household Income	-17.72***
Black	-138.32***
Asian	-113.80***
Other	-20.22
Hispanic	83.38**
Relatively urban	-1.97
Relatively rural	15.69
Very rural	40.53***
Northeast	73.29***
South	74.53***
West	30.28**
Current avg price	-527.17***
Constant	463.49***
N	40032
R^2	0.26
θ_1	0.11**
θ_2	0.11**
E_d	-0.03**

Table A 7 Fruit Juices Myopic Model 2008 IRI data

Variable	*Estimate*
1-month lag	0.29***
2-month lag	0.29***
Widowed	-10.80**
Divorced/Separated	-3.48
Single	-7.24
August	-14.62**
September	-11.84**
October	-25.12***
November	-23.52***
December	-8.04
2-person household	11.33**
3-person household	17.55**
4-person household	27.48***
5-person household	43.56***
6-person household	55.04**
7-person household	65.42
8-person household	70.18**
Household Income	0.88
Black	16.89***
Asian	13.16
Other	18.57
Hispanic	-1.15
Relatively urban	-5.78
Relatively rural	-11.56**
Very rural	-4.01
Northeast	-7.28*
South	-1.77
West	5.01
Current avg price	-807.14***
Constant	164.12***
N	31026
R^2	0.47
θ_1	0.29***
θ_2	0.29***
E_d	-0.16***

Table A 8 Fruit Juices Myopic Model 2012 IRI data

Variable	*Estimate*
1-month lag	0.20**
2-month lag	0.21**
Widowed	-15.52*
Divorced/Separated	10.09
Single	13.93
August	-12.22
September	-11.82
October	-13.69
November	-4.85
December	-6.14
2-person household	12.66*
3-person household	18.37**
4-person household	48.36***
5-person household	49.18**
6-person household	30.09
7-person household	40.26*
8-person household	18.13
Household Income	1.33
Black	27.19
Asian	17.77
Other	8.00
Hispanic	-15.73
Relatively urban	-2.83
Relatively rural	-7.62
Very rural	-7.00
Northeast	-5.80
South	-7.88
West	0.96
Current avg price	-906.58***
Constant	216.96**
N	11496
R^2	0.42
θ_1	0.20**
θ_2	0.21**
E_d	-0.24**

Table A 9 Other Grains 2008 IRI data

Variable	*Estimate*
1-month lead	0.15***
2-month lead	0.13***
3-month lead	0.13***
4-month lead	0.10***
1-month lag	0.14***
2-month lag	0.12***
3-month lag	0.12***
4-month lag	0.07***
Widowed	-0.41
Divorced/Separated	-0.38
Single	0.16
October	-9.46***
November	-6.84***
December	-1.84*
2-person household	1.44
3-person household	2.59**
4-person household	2.78*
5-person household	4.22**
6-person household	2.37
7-person household	-0.51
8-person household	14.78**
Household Income	-0.14
Black	-1.48
Asian	-2.11
Other	1.06
Hispanic	1.5
Relatively urban	0.84
Relatively rural	1.75**
Very rural	1.94***
Northeast	-1.35**
South	-1.44**
West	-2.06**
Current avg price	-41.30***
Constant	13.04***
N	54116
R^2	0.56
δ	1.02***
δ^2	1.05***

Table A9 (continued)

Variable	*Estimate*
δ^3	1.13***
δ^4	1.45***
θ_1	0.14***
θ_2	0.12***
θ_3	0.12***
θ_4	0.07***
E_d	-0.04***

* p<0.10, ** p<0.05, *** p<0.01

Table A 10 Other Grains 2012 IRI data

Variable	*Estimate*
1-month lead	0.14***
2-month lead	0.13***
3-month lead	0.11***
4-month lead	0.09***
1-month lag	0.15***
2-month lag	0.12***
3-month lag	0.11***
4-month lag	0.09***
Widowed	1.32
Divorced/Separated	-0.18
Single	-0.33
October	-2.26**
November	-1.39
December	-3.26***
2-person household	3.18***
3-person household	3.59***
4-person household	3.62**
5-person household	2.82
6-person household	8.52***
7-person household	1.78
8-person household	5.63
Household Income	-0.16
Black	-0.70
Asian	-1.70
Other	2.51*
Hispanic	5.63***
Relatively urban	-0.99*
Relatively rural	-0.4
Very rural	-0.33
Northeast	-0.14
South	-0.44
West	-0.31
Current avg price	-52.35***
Constant	3.39
N	42308
R^2	0.5
δ	0.99***
δ^2	1.11***

Table A 10 (continued)

Variable	*Estimate*
δ^3	0.99***
δ^4	0.98***
θ_1	0.15***
θ_2	0.12***
θ_3	0.11***
θ_4	0.09***
E_d	-0.06***

* p<0.10, ** p<0.05, *** p<0.01