Access versus Choice: testing the "food desert" construct in Champaign, IL

by

Rhiannon L. Jerch[a]
Katherine R. Baylis[b]
Sahan T. M. Dissanayake[c]

[a] MS Candidate in Agricultural & Resource Economics, University of Illinois at Urbana-Champaign
rjerch2@illinois.edu. Tel. +1 217-333-1810
[b] Assistant Professor of Agricultural & Resource Economics at University of Illinois at Urbana-Champaign
baylis@illinois.edu. Tel. +1 217-244-6653
[c] Assistant Professor of Economics at Colby College
Sahan.Dissanayake@colby.edu. Tel. +1 207-859-5241

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4-6, 2013.

Abstract: How does one's food environment affect food purchase decisions? Food access has received significant political and academic attention, particularly under the Obama administration. Existing literature on food access and "food deserts" mainly focuses on geographic distance or the retail of fresh fruits & vegetables versus fast food within a neighborhood to determine and identify inequitable access. In this paper I attempt to develop an endogenous measure of food access by asking how geographic placement of food retail affects food expenditure, particularly of fruits & vegetables. I use novel data on 886 households matched to food prices from a census of geocoded food retailers in Champaign County to approach this question from two perspectives. I first estimate the household's share of grocery expenditures allocated to fresh, frozen, and canned fruits & vegetables versus other grocery items. I then use data on a person's residence and geocoded data on food retail locations in Champaign County to test for relationships between retailer proximity, and the share of expenditure on fruits & vegetables. The next perspective uses a choice experiment to measure the tradeoff among store characteristics that determine where a consumer shops. The demand estimation reveals how much fruits & vegetables a person is actually consuming, while the choice experiment reveals whether that individual is constrained in their consumption by their existing characteristic set of stores. I find that while proximity to a grocery store is positively correlated with healthier food consumption, policy response should focus on improving store quality and product quality to induce behavioral change. I further find policy response should be cognizant of endogenous locational sorting which may require alternative means to improve health other than changing the food geography.

Chapter One: Introduction

1.1 Motivation & Research Question

Outside of a Save-a-Lot store in Chicago's Auburn Gresham neighborhood in Spring of 2011, mayor Rahm Emanuel praised a joint initiative between the city and the retailer to stamp out food deserts: "The notion that people do not have access to fresh fruits and vegetables is unconscionable, and you cannot have that today" (*Time Out Chicago*). The United States Department of Agriculture (USDA) defines "food desert" as a low-income census tract where a significant share of residents live more than one mile from a large grocery store. Roughly one in five Chicagoans lives in a food desert, according to a 2006 study by Mari Gallagher Research & Consulting Group.[1] One such area includes Auburn Gresham, a heavily African-American neighborhood on Chicago's south side. After Emanuel's speech, resident Taylor McDowell stated in an interview with *Time Out Chicago*: "...[T]his is not a food desert: We have a Jewel down the road, another Save-A-Lot on 79th and Halsted, Aldi around the corner on State. People shouldn't feel sorry for us. This is a black middle-class neighborhood. It's not like I'm suffering."

Today, federal, state and municipal governments are rushing to vanquish "food deserts" across the country. In February of 2010, the Obama Administration proposed a $400 million Healthy Food Financing Initiative to promote retailers of healthful foods to move to underserved urban and rural communities. Since 2001, California, Nevada, New Mexico, Texas, Oklahoma, Louisiana, Illinois, Chicago, New York, Pennsylvania, D.C., and Maryland have enacted legislation aimed at promoting healthier food retail. Such legislation has included financial incentives to attract healthier food retail outlets to underserved areas, improving healthy food offerings in existing stores through subsidized remodeling and restocking, and relaxing zoning requirements to make it easier for grocery stores to locate in underserved areas (Centers for Disease Control and Prevention). Such policy interventions rest on the assumption that the mile-or-more expanse to a grocery store prohibits area residents from consuming produce and healthy foods. This paper creates a framework to investigate whether residents of these so-called "food deserts" are truly constrained by their food geography.

Classical consumer theory purports that individuals are mobile and possess perfect information about choice alternatives, including residential location. As the Tiebout model suggests, people move and are free to choose their communities (Tiebout 1956). Still, there is evidence that a dearth of grocery stores is more likely to occur in poorer areas (Berg and Murdoch 2007; Ver Ploeg et al. 2009; Morland et al. 2002, Liese et al. 2007). Residents of poorer neighborhoods may face challenges to food access on several fronts: convenience stores and quick-service restaurants tend to outnumber grocery stores (Larson, Story, & Nelson 2009; Morland et al. 2002); fresher nutrient-rich foods tend to be more expensive than calorie-dense, nutrient- poor foods; (Drewnowski, 2003; Drenowski, Darmon, & Briend 2004; Monsivais & Drewnowski 2007), and those retailers that provide lower-cost

[1] 500,000 people cited in the study divided by the 2000 Census Chicago population of 2.9 million.

health options fail to establish stores in low-income areas due to security costs or higher profit risk. While prohibitive supply-side costs can deter grocery stores from locating in low-income areas, so can demand-side factors, such as consumer preferences. A food desert might arise in a geographic area with insufficient supply of nutritious food, yet a food desert might also arise in an area with insufficient demand (Bitler & Halder 2011).

In either case, the extra distance required to purchase fresher foods increases the implicit costs of these foods. The result is that some people face greater chances of obesity and diet-related disease while others may regularly shop for their food at convenience stores. Recent research has found a positive correlation between supermarket access and diet and health (Wilde, Gundersen, Baylis, & Llobrera 2012; Larson et al. 2009; Chen, Florax & Snyder 2010) as well as a negative correlation between convenience store access or fast food access, and diet and health (Wilde et al. 2012; Giskes, van Lenthe, Avendano-Pabon, & Brug 2011; Gibson 2011; Currie, Della Vigna, Moretti & Pathania 2009). The policy response to food deserts has largely ignored the fact that a causal relationship between access and diet has yet to be proven. Researchers in the United Kingdom, for example, found that cultural influences represented by a person's age or gender are significant determinants of healthy food consumption whereas distance to a grocery store, produce prices, or socio-economic status are not (Pearson, Russell, Campbell, and Barker 2005). If policy response aims to promote dietary health, the appropriate action depends on solving an identification problem: whether poor diet and obesity result from supply-side factors, or from an individual's inherent preferences.

The food policy and public health literature has called significant attention to the proper definition and identification of a "food desert" (California Center for Public Health Advocacy 2007; Mari Gallagher 2006; Ver Ploeg et al. 2009). The USDA "Food Access Research Atlas" identifies over 8,800 census tracts, harboring 29.7 million people, which qualify as food deserts.[2] By incentivizing grocery stores to locate in these areas, policy makers, including the Obama administration and Rahm Emanuel, have reacted under the assumption that insufficient access to grocery stores deters people from eating fresh produce. Further, they assume access will have an opposing effect, and induce residents living nearby a grocery store to consume more fresh produce. These policy makers assume that food consumption is exogenously related to one's food environment. In other words, the "food desert" construct ignores the Tiebout principle that people may choose where they live. Consumption behavior may be exogenously driven, but may also be a product of existing consumer predilections and tendencies that are independent of their exogenous environment. In other words, endogenous locational sorting may drive those who enjoy fresh produce to live nearby a grocery store, such that the apparently positive effects of grocery store proximity on diet are spuriously related. This paper explores potential weaknesses of the "food desert" construct by asking first whether food geography can constrain consumption of fresh produce; and second, whether the source of such consumption constraints may be exogenously or endogenously driven.

[2] <http://www.ers.usda.gov/data-products/food-access-research-atlas/documentation.aspx#.UUvLCBk1cYJ>

In 2008 and 2009, Detroit received significant national attention from the academics, the media, and the federal government for its dearth of grocery stores. Detroit resident Noah Stephens initiated a photography survey project to document the produce of Detroit's supermarkets. Stating his motivation for the project, Stephens conceded: "The food choices of many Detroiters may be constrained by affordability, transportation, safety, food quality. Or people may just like potato chips. The heart of this project is the question of access versus choice" (Huffington Post 2012). Like Stephens' *Kickstarter* project, this paper seeks to address and quantify access versus choice.

1.2 Hypothesis

In this paper I attempt to develop an endogenous measure of food access by asking how consumer preferences interact with geographic location of food retailers to effect food consumption, particularly of fruits & vegetables. I use novel data on 886 households matched to food prices from a census of food retailers in Champaign County to approach this question from two perspectives. I first estimate household shares of grocery expenditures allocated to fresh, frozen, and canned fruits & vegetables and other grocery items. I then use data on household location and geocoded data on food retail locations in Champaign & Piatt counties in central Illinois to test for correlation between retailer proximity and household grocery expenditure share of fruits & vegetables. The second perspective uses a choice experiment to measure preferences for grocery store characteristics that determine where a consumer shops for food. I then compare the grocery expenditures and preferences between individuals who live near a grocery store, and those that live further from a grocery store. If these two subsets have differing consumption behavior, but similar preferences, this will suggest food geography constrains consumption. Such a result would support current policy's assumption that exogenous, treatment effects of adding grocery stores to "food deserts" can induce fresh produce consumption. On the other hand, if both consumption behaviors as well as preferences are systematically different between individuals living nearby a grocery store and individuals living far from a grocery store, this will suggest food geography has an indeterminate effect on food choice, and instead endogenous sorting is the primary cause of divergent consumption behavior. Alternative policy responses such as subsidizing fresh produce, or nutrition education may be more effective at combating poor food choice than simply adding a grocery store.

Fruit & vegetable consumption is the main focus of this thesis for two reasons. First, nutritional science unanimously agrees that eating fruits & vegetables is a primary component of dietary health and is directly associated with reduced diabetes and obesity risk (Harvard School of Public Health). Second, consumers who are food-access constrained are more likely to compromise consumption of fresh foods compared to non-perishable foods because either the fresh foods are expensive relative to non-perishable foods, unavailable entirely, or poor quality (Drewnowski, 2003; Drenowski, Darmon, & Briend 2005; Monsivais & Drewnowski 2007; Bitler and Haider 2011). Fresh fruits & vegetables are costly to supply because stocking fresh produce necessitates a refrigerated

and timely supply chain as well as expensive in-store refrigeration, which both serve to increase the sale price of fresh produce. This is particularly true for if the retailer supplies a small quantity and is unable to harness cost savings through economies of scale (Bitler and Haider 2011; Ver Ploeg et al. 2009). The effective cost of fresh produce, thus, increases in absolute terms, but also in terms of opportunity cost. A consumer who must travel far to grocery shop will want to maximize their grocery expenditure and purchase foods that will last until their next grocery shopping trip. Thus perishable foods are effectively more expensive to an infrequent grocery shopper because the shelf life opportunity cost of the fresh food is high relative to the non-perishable foods. The shelf-life opportunity cost issue exacerbates with poor quality produce, which may be both displeasing the consumer, and can spoil quickly. Geography can therefore constrain a consumer's effective budget, time availability, and quality preferences. Consequently, consumers will be more likely to trade off fresh fruits & vegetables for non-fresh fruits & vegetables in order to reduce costs in a restrictive food environment.

I find that survey respondents affiliated with the University of Illinois exhibit inelastic demand for fresh produce the closer they live to a grocery store. Additionally, respondents who live within one mile of a grocery store have lower propensity to substitute fresh for non-fresh produce compared to respondents who live over one mile from a grocery store. While endogenous sorting explains some of this divergent consumption behavior, a significant proportion of respondents who live over one mile from a grocery store would likely exhibit more stable consumption of fresh produce if higher quality specialty stores were made accessible to them.

This study extends the current literature on food choice by accounting for the effect of the food environment on an individual's food purchases, in addition to the standard consumer preference considerations. This thesis contributes specifically to existing empirical demand work by utilizing first-hand data on food purchases and observed prices to estimate fresh produce demand in Champaign & Piatt counties, Illinois. Further, this study contributes to food access policy by empirically testing for consumption constraints. Current food access policy relies on geographic measures, such as number of large grocery stores per square mile, to identify "food deserts", or areas deemed inhospitable to healthful food consumption. This paper allows individual preferences, as opposed to hetero-normative definitions of acceptable food landscapes, to motivate identification of areas truly underserved. This research can improve targeting areas in genuine need of policy intervention through quantifying gaps in consumers' preferences and their actual consumption bundles.

In the next chapter, I review theoretical and empirical literature concerning food access, hedonic analyses, and food demand estimation. In the third chapter, I present a theoretical basis for evaluating consumer preferences and consumer behavior through a choice experiment and demand system. In the fourth chapter, I discuss my data and provide summary statistics. The fifth chapter discusses my results and conclusions, and the final chapter summarizes and concludes my findings.

Chapter Two: Review of Theoretical and Empirical Research

2.1 Introduction

This thesis investigates how the food environment affects food choices in Champaign & Piatt counties, IL. Specifically, I examine how the array of foods available to you affects, if at all, what you purchase. I employ both demand system estimation as well as a choice experiment to answer this question with particular concern for fruit & vegetable consumption. Analysis of microeconomic food demand and consumption behavior at the disaggregated level possesses an inherent simultaneity problem; demand is affected simultaneously by the choice of where to shop - a discrete choice - and what to consume - a continuous choice. When purchasing groceries, the consumer cannot make one decision without implicitly making the other decision. Previous literature has looked at both pieces in isolation, but few scholars have considered the discrete choice and the continuous choice as simultaneous determinants of food consumption. In order to understand how, and to what extent, the food environment affects healthy food consumption, I investigate both the choice of where to grocery shop, as well as the choice of what foods to consume once at the grocery store. I compare these choices between survey respondents living within one mile of a grocery store, and respondents living over one mile from a grocery store following the USDA geographic criteria for residence of a "food desert." I will hereinafter refer to these respondent groups as "less than 1 mile" residents and "over 1 mile" residents, respectively.

Two questions drive the methodology of this paper:

1) How are the preferences that drive where people grocery shop - the discrete choice - different among "less than 1 mile" residents and "over 1 mile" residents?

2) How does consumption of fresh fruits & vegetables, non-fresh fruits & vegetables, and other foods – the continuous choice - differ among "less than 1 mile" residents and "over 1 mile" residents?

In the following sections, I review literature concerning theoretical and empirical applications of choice preferences and consumption behavior that will provide a basis for my subsequent analyses.

2.2 Non-market Valuation

2.2.1 Introduction

In order to understand preferences that drive where people grocery shop, I consider what people are willing to pay for their preferred grocery shopping experience. Specifically, I consider how individuals might value improved access to groceries. Food "access" possesses several dimensions in addition to proximity. A corner store next door to your house might carry apples, but if these apples are bruised, tasteless, and expensive, you will be unlikely to purchase them. It follows that I can consider "improved access" to include product appearance, quality, and price, in addition to proximity. Further, assume two supermarkets selling apples of the same quality and price exist next door to one another. If Store A has substantially better customer service, shorter check-out lines, or sources its

produce in a more sustainable manner than its competitor, Store B, you may be more inclined to shop at Store A. Harold Hotelling argued that even if Store B cut its apple prices, many consumers will still prefer Store A because preferences for in-store service, merchandise mix, product quality, methods of doing business, and political persuasions of the seller or vendor can all shield Store B from harsh price competition with Store A (Brown 1989). An apple purchased at any grocery store has several characteristics that dictate whether or not you purchase that apple, besides those specific to the apple itself. Even for completely standardized products, absolute homogeneity does not exist if we account for the location and characteristics of the retailer. Thus, in deciding what groceries to purchase, we inherently consider not only store location and the time required to travel to the store, but also product quality, product price, and the marketing chain in which the groceries are sold, among other things (Hanemann 2006). The result is that an apple purchased at one store location is not the same as an identical apple purchased at a different store because the time it took to acquire the apple, its price, the store ambience, and the marketing channel from which it came can all differ. In this thesis, I investigate if and how people alter their produce purchases given variations in these characteristics.

2.2.2 Lancaster's "New Theory"

Lancaster postulated that utility, or a preference for one good over another, is derived from characteristics that goods possess rather than the goods themselves (Lancaster, 1966). Kelvin Lancaster's "new theory" provides the theoretical basis for estimating changes in utility that result from attribute changes of a good. His theory holds that the goods' characteristics are taken as given, and the consumer is free to vary only the quantity they consume of a particular good. Lancaster postulated that the more advanced an economy, the more characteristics available relative to goods, so that in the most advanced economies, consumers could obtain any combination of characteristics they desired, constrained only by their budget.

Product characteristics cannot be purchased individually be the consumer. For example, a consumer cannot buy high quality produce separately from a five-minute travel time and manufacture their own grocery store. Grocery stores comprise "lumpy" goods because shoppers have a finite, discrete opportunity set of grocery stores to patronize, or "consume." In other words, the array of grocery store characteristics available to a consumer will not be an exhaustive set of possible store attribute combinations. Assuming that travelling to multiple stores is costly, consumers must decide which of these characteristics is most important, and patronize the store that best fits their criteria. By isolating the consumer's internal tradeoffs of grocery store characteristics, such as travel time, produce quality, price, and marketing channel, it is possible to assign a numeric value or "price" to these characteristics through non-market valuation. Non-market valuation measures the satisfaction from these store characteristics as the monetary amount that a consumer would be just willing to exchange for the characteristic if it were possible to make such an exchange (Hanemann 2006). Non-market valuation enables us to understand how consumers value different store characteristics, and in turn, predict the store people are most likely to patronize

assuming any given store type is accessible to the consumer. This paper aims to identify how quantity of fruit & vegetable consumption is affected by varied access to an individual's preferred grocery store.

2.2.3 Contingent Valuation & Choice Experiments

Preferences for particular good characteristics can readily be observed through the monetary value individuals place on the consumption or use of a good or service. Contingent valuation is a form of non-market valuation that utilizes interviews or surveys to elicit individuals' monetary value for particular goods (Hanneman 2006). If products are considered to be composed of attributes, such as the Lancaster theory suggests, a form of contingent valuation known as a choice experiment can be used to determine which product attributes are important to overall preferences for that product and which combinations of attribute levels are most preferred. Using rankings or rating-scale evaluation judgments of products, choice experiments decompose such rankings into components based on qualitative attributes of the products (Kuhfeld 2010).

Choice experiments are a popular contingent valuation technique in marketing, environmental, and public policy research (Kuhfeld 2010; Louviere, Hensher & Swait 2000; Alpizar, Carlsson, & Martinsson 2001). A choice experiment provides a framework in which one can establish a set of behavioral rules under which it is reasonable to represent an individual's choice-making decision when considering a set of alternatives. Choice experiments estimate welfare changes resulting from changes in a good's attributes. In the case of this paper, the choice experiment estimates changes in consumer welfare for changes in grocery store attributes, where "attributes" are defined as the characteristics of a store that cause a person to choose one alternative over another. Attributes are the source of utility (Hensher, Rose, and Green 2005).

My thesis applies Lancaster's concept of characteristic-derived utility to the decision of where to grocery shop for residents of Champaign & Piatt counties. I hypothesize that unless consumers can obtain their preferred combination of grocery store characteristics, they are constrained in their ability to purchase the exact bundle of food products they desire because the decision of what groceries to purchase is wrapped up in a simultaneous decision of where to grocery shop. I test whether consumers are constrained in their consumption of produce through a choice experiment framework where grocery stores are considered to be composed of attributes including size, price, service offerings, produce quality, and travel time. For example, one hypothetical grocery store may possess large floor space, a pharmacy and florist, and a wide variety of non-food items, sell high quality fresh produce, require five minutes of travel time to access, and charge a five percent premium for groceries. While, another hypothetical grocery store may possess small floor space, sell very few non-food items, offer a limited array of fresh produce, require 15 minutes of travel time to access, and charge a 15 percent discount for groceries. The choice experiment is used to determine which attributes are important to food product preference, and which combinations of attributes are most preferred among respondents (Kuhfeld 2010).

In reality, most consumers face multiple constraints that restrict them from choosing their preferred alternative. An individual may not be able to afford a car that they want because of expense. Similarly, residents of Champaign County, though they prefer to grocery shop at Trader Joes, are inhibited by the 2.5-hour travel time needed to get to the nearest store.[3] Yet, through the choice experiment, I can allow such constraints to be non-binding in order to elicit information about a person's preferences.

2.2.4 Grocery Store Preferences: Previous Empirical Work

Literature on consumer preferences for supermarkets have explored the effects of store size, travel distance, customer service, and quality on grocery shopping behavior. Food access policy has historically used quantifiable metrics such as store size or travel distance to identify areas where residents cannot easily consume fresh produce. However, researchers Blitstein, Snider, and Evans (2011) found that grocery shoppers don't only care about cost and proximity, but also need choice and quality if they are going to buy fresh produce. Their study utilized survey data from 495 respondents in a low-income neighborhood of Chicago, IL to examine how residents' perceptions of their food shopping environment was related to their dietary intake of fresh produce. The researchers found that among their urban, minority low-income sample, those who felt their produce vendors provided high quality produce and selection consumed higher quantities of produce.

Daniel Brown (1978) found in a panel study of 101 Chicago suburban women that travel distance to a grocery store had a stronger effect than store size in determining store patronage. Brown also found that grocery store size attracts patronage up to a point, after which increasing size deters grocery shopping. People want variety, but not too much. Gómez, McLaughlin, and Wittink (2004) found that a decrease in product quality (as opposed to an increase), and changes in customer service and value explain 75 percent of the variation in customer satisfaction among consumers surveyed in 250 east coast stores. Thus, product quality and customer service are important determinants of utility with respect to grocery shopping.

Huang and Oppewal (2006) conducted a choice experiment among 152 supermarket shoppers in England to determine which attributes are most instrumental in causing people to purchase their groceries online rather than in-store. Retailer attributes included the online shopping delivery charge, time availability of the respondent, travel time to the grocery store, and purpose of the grocery purchases. Travel time has a stronger effect on the decision to purchase groceries online than does the delivery charge, which indicates people place high value on their personal time. The authors further find that the purpose of the shopping trip has little effect on respondent's choice outcome. People will shop for ingredients for a dinner party at the same place they will shop for their routine groceries.

[3] Of 886 responses, 60 noted they would shop at Trader Joe's if one existed in Champaign County.

Analyses of grocery store preferences appear frequently in the marketing literature, as well as the urban planning literature. Arnold, Oum, and Tigert (1983) investigated the "determinant attributes" of store choice, which they defined following Myers and Alpert (1968) as the "attributes projected by the product's image which lead to the choice of that product (p. 150) [.]" Using a multinomial logit model, the authors identified robust determinants of grocery store patronage through interview-based survey data across six geographic markets and six years. Locational convenience and low prices were the most important determinant attributes across geographic markets and over time. In addition to store location, Tang, Bell and Ho (2001) also found that familiarity with the store, service quality, and breadth and depth of products are key drivers of shopping utility. The authors parsed a utility model of fixed and variable costs and benefits to describe overall grocery shopping utility. Using data from 500 households across five stores over two years, their utility model correctly predicted 83 percent of the sample's store choice. Carrasco (2008) perceived determinants of grocery store choice as a destination choice problem in transportation planning. He used a revealed preference approach and travel behavior survey data on 32,000 Swiss households from 2005 to investigate what factors best predict where individuals choose to grocery shop. The author's discrete choice model included variables for the following store characteristics: store quality, price range, store size, opening hours, distance from home, distance needed to deviate from route for non-home based trips, and retail density near the grocery store. He found households have a clear preference for larger stores. He also found sample households are averse to travelling far from home to grocery shop, and greatly dislike deviating from a set route in order to purchase groceries. Higher income households had a stronger aversion to travel distance, indicative of the value these households assign to time. Also, households preferred relatively low retail density near their grocery store, likely due to timesavings from less congestion on the shopping errand.

A 2011 revealed preference study by Hsieh and Stiegert (2011) found that households that spend a high proportion of their grocery bill on "dry" grocery items are more likely to shop at value-oriented stores such as supercenters, and are less likely to shop at high-end, natural food supermarket chains. Their study gives evidence that travel distance from a consumer's household to the store is an important determinant of the store type in which a consumer decides to shop. My thesis research aims to take one step further to investigate whether distance also affects the foods consumers ultimately purchase.

2.3 Literature on Food Demand

2.3.1 Introduction & Theoretical Considerations

"Food deserts" are problematic to public health in so far as they inhibit consumption of fresh produce. If food geography inhibits or induces consumption of fresh produce, its effect will manifest through distinctive consumption behavior between "less than 1 mile" residents and "over 1 mile" residents.

The basic objective of consumer behavior theory is to explain how a rational individual makes decisions on what to consume when faced with a constraint. Under the assumption that a consumer can make a complete and consistent order of preferences for particular goods within a finite choice set, the choices the consumer will make are limited by their income or budget (Bentham 1789). The consumer must find an optimal bundle of goods within the finite choice set so as to maximize their utility but keep expenditures at or below their budget constraint (Teklu and Johnson 1986). A budget constraint is one example of factors that shape, or "constrain" the possibility frontier within which preferences can be honored. One may equally think of additional constraints such as time and location that limit the combinations of goods that an individual can consume. For example, current policy addressing food deserts asserts that residents face both budgetary and geographic constraints in their grocery consumption decisions (Ver Ploeg et al. 2009). Current policy is particularly concerned with how such constraints inhibit demand for fresh, healthy foods. Because demand for fruits & vegetables is integrated with demand for other types of foods if both are considered a subset of the same commodity group, I investigate relationships between food geography and fruit & vegetable demand using a system of demand. The system measures fruit & vegetable demand simultaneously with all other food demand.

Figure 1 illustrates individual consumer demand through a graphical representation. The y-axis is a ratio of quantity of two goods, and the x-axis is a price ratio of the two goods. The horizontal line is a budget constraint representing all possible combinations of the two goods that the consumer can purchase subject to a finite amount of money they can spend. If the consumer's budget constraint changes, the amount of the two goods they consume may change. Further, if the price of one good changes, the ratio in which they consume both goods can also change. Thus, variation in an individual's budget can change their consumption levels, and changes in prices of each good can change the ratio in which both goods are consumed. Otherwise stated, changes in prices can cause the consumer to "trade off" one good for the other.

Quantification of trade off behavior is measured through elasticity of income and the elasticity of substitution (or cross-price elasticity), respectively. Income elasticity measures changes in quantity purchased given a change in the consumer's budget, where cross-price elasticity measures changes in consumption of a particular good given changes in price of a different good. Cross-price elasticities are generally negative for complementary goods and positive for substitute goods. Own-price elasticity, also referred to simply as the price elasticity of demand,

measures the percent change in consumption of a good given a percent change in price for the same good. Price elasticities are generally negative, except in the case of Giffen goods (Nicholson & Snyder 2008, p159). Price elasticities are necessary in order to understand consumer behavioral responses to variation in product price. Price elasticity of demand is determined by several factors including availability of substitutes, income, consumer preferences, and household budget share for a particular product, among others (Nicholson and Snyder 2008, p. 150).

Demand is inelastic when the relative change in quantity purchased is less than the relative change in price, so that the absolute value of the own price elasticity is between 0 and 1. Demand is elastic when the relative change in purchased quantity is greater than the relative change in price, so that the absolute value is greater than 1. Price elasticity of demand, itself, is dimensionless, but provides a reference for comparing purchasing behavior for price changes across different goods. Goods that are commonly price inelastic include necessity or every-day-use items with few substitutes such as public transportation, cigarettes, staple food items like eggs and rice, or gasoline. Price elastic goods have relatively many substitutes. Price elastic goods can include "luxury" goods, not necessary for every-day-use such as cinema tickets, imported French wine, or vacation travel. Because price elasticity lacks dimension and provides meaning only in relative terms, the degree of elasticity for a particular good depends on an individual's theorized consumption basket (Nicholson and Snyder 2008 pg. 26).

I estimate both uncompensated, or "Marshallian", and compensated, or "Hicksian" demand elasticities. Where compensated demand models estimate only substitution between products, uncompensated demand models account for both substitution effects and the effective income change resulting from a good's price change (Nicholson & Snyder 2008, p. 151). I further uses both income elasticity as well as own-price elasticity to compare behavioral differences in grocery consumption between "1 mile or less" and "Over 1 mile" residents. Income elasticities will reveal whether proximity to a grocery store is correlated with greater price sensitivity of fresh produce. Own-price and cross-price elasticities will reveal whether food geography is correlated with greater substitution of fresh for non-fresh produce.

2.3.2 Food Demand Functional Forms

The Almost Ideal Demand System (AIDS) pervades the empirical food demand estimation literature. The model was first developed and applied to consumer demand empirics by Deaton and Muellbauer in 1980. The AIDS model, unlike simpler demand models based on the linear expenditure system generalization of Cobb-Douglas (e.g., Rotterdam model, translog models, etc.) does not require that consumers at all points along the income distribution allocate their budgets identically (Nelson 2013; Tiffin, Balcombe, Salois, and Kehlbacher 2011). The resulting system is flexible, thus better situated to model empirical data. Further, the resulting system satisfies the neoclassical utility theory axioms of choice exactly, including the adding up, homogeneity, and symmetry

restrictions (Deaton and Muellbauer 1980a). The model gives an arbitrary first-order Taylor series approximation to any demand system, and conditions of the model guarantee the existence of theoretically consistent aggregate demand (Deaton and Muellbauer 1980a; Nelson 2013), an advantage over the Rotterdam model perpetuated by Henri Theil and Anton Barten in the 1960's. The AIDS model allows the theoretical adding-up, homogeneity, and symmetry restrictions to be tested statistically rather than imposed, as with translog linear demand models (Deaton and Muellbauer 1980a). Additionally, estimation of the Rotterdam model requires time series variation in the data, either through panel data or times series on aggregate consumption (Nelson 2013). Because data for this thesis comprises a cross-section, use of the Rotterdam model would be inappropriate.

Determination of demand system matrix rank provides the precursor to appropriate model specification. Rank 1, rank 2, or rank 3 imply budget shares and total expenditure exhibit either an independent, linear, or quadratic relationship, respectively. A rank 2 system, for example, has budget shares that are linear in log total expenditure and are composed of two independent price functions. Gorman (1981) demonstrated that any exactly aggregable demand system could possess at most three independent income (or expenditure) terms in order to satisfy Slutsky symmetry[4], such that the highest possible rank is rank 3.

The rank 2 AIDS form is generally consistent with empirical food expenditure data. For example, Katchova and Chern (2004) tested several demand models using Chinese household expenditure data on grains, vegetables, fruits, pork, and poultry and concluded the AIDS model produced the most reasonable set of expenditure elasticity estimates based on their a priori expectations compared to the rank 3 model. Wood, Nelson, and Nogeuira (2009) estimated consumer welfare loss resulting from increased staple food prices among poor households in Mexico. Their Engle curve analysis based on food expenditures and budget shares across six food groups identified rank 2 as appropriate for estimating Mexican food demand among low-income households. Huang and Lin (2000) derived a modified version of the AIDS model to estimate food demand using 4,500 American households. The authors assumed linearity in income, thus did not test model fit using quadratic income terms. They investigated elasticity differences among three income brackets. The majority of their demand elasticities were statistically significant and met a priori expectations on size and magnitude. Tiffin et al. (2011) examined how changes in food prices and food expenditure affect consumption of various nutrients among UK households. Their extensive literature review of European household food demand estimation concluded that the AIDS model was most robust model. Eales and Unnevehr (1994) found the AIDS model has a functional form consistent with observed US meat demand. Erdil (2006) compared estimates for agricultural products in OECD countries between the AIDS, CBS (Central Bureau of Statistics of the Netherlands), and Rotterdam models using a nested model selection procedure. For 12 of the 25

[4] The Slutsky equation (also known as the Slutsky identity, or Slutsky symmetric) demonstrates that change in the demand for a good caused by its price change can be decomposed into two effects, the substitution effect and the income effect. Slutsky symmetry must be satisfied in order to relate changes in uncompensated demand with changes in compensated demand and empirically estimate demand elasticities. (Nicolson, Snyder 2008, 156.)

OECD countries, the AIDS model provided the best representation of demand for agricultural products. Lee, Brown, and Seale (1994) tested several models combining features of the Rotterdam and AIDS models and found the AIDS-based models better described Taiwanese consumer behavior.

Conventional demand system estimation techniques that specify budget shares as the dependent variable, including the AIDS model, have received recent criticism regarding use of unit values as a proxy for price (Gibson and Rozelle 2005; McKelvey 2011). When price data for particular goods is not available, which is often the case; price data is inferred using expenditure and quantity survey data to generate unit values. One concern is that individuals may respond to price changes by substituting on the quality margin rather than quantity, which would understate the change in unit value. For example, a consumer may respond to a price increase of fresh organic green beans by substituting the fresh beans for frozen, conventionally grown green beans. When the price of fresh beans increases, the change in unit value for the category containing green beans will understate the original price increase, all else equal. The imperfect proxy for price can be exacerbated if the food category is particularly broad, if several quality differentials are available to the surveyed consumer, or if the household sample contains significantly heterogeneous consumers. Wealthier households, for example, may exhibit high income- elasticity of demand for particular foods because of frequent substitution on the quality margin (Moro, Daniele, and Sckokai 2000; Abdulai 2002). McKelvey (2011) demonstrated that quality substitution is substantively large when comparing demand estimation results between unit value-based estimation and actual price data.

I use a sampling of actual price data for each food category to avoid any issues arising from use of unit values. Additionally, I disaggregate food categories using quality as a major characteristic, thus mitigating major quality differences stemming from aggregation. For example, I separate fresh produce from frozen and canned fruits or vegetables. Separation of these food categories follows Deaton (1988) whereby commodity groups are deemed to be weakly separable. Weak separability is a necessary condition for two-stage budgeting in which the consumer first decides how much to spend on food, clothing, and other broad product groups, then allocates consumption of particular goods within those broad groups. Consumption of goods within each group is modeled as a sub-utility function, and a sub-maximization problem. The weak separability assumption is important for demand estimation because it places severe restrictions on the degree of substitutability between goods in different groups. While a consumer may substitute between fresh green beans and frozen green beans, we assume they do not substitute between fresh green beans and shoes. Further, by assuming that commodity groups are weakly separable, considerable structure is imposed on the nature of quality responses to price changes (McKelvey 2011).

2.3.3 Review of Fruit & Vegetable Demand & Consumption Behavior

Consumer theory accepts that food demand in industrialized countries is largely inelastic, or price-insensitive. (Herrmann and Roeder 1998). The USDA *Commodity and Food Elasticities* database reports average income

elasticity for fruit & vegetables in the US of 0.22 and an average own price elasticity of (-.26), both relatively inelastic indicators.[5] You, Epperson, and, Huang (2011) estimated a demand system for 11 fruits and 10 vegetables using aggregate national data from 1960 through 1993 and found fresh fruits and vegetables respond significantly to changes in their own prices but less so to changes in income, suggesting that produce price has greater influence than household income in determining fruit and vegetable demand among US consumers. Income elasticities of demand for fresh vegetables & fresh fruits were relatively inelastic at 0.15 and (-0.13), respectively, though neither were statistically significant. Cross-price elasticity for processed vegetables was 0.05 while cross-price elasticity for processed fruit was 0.02, though neither was statistically significantly different from zero.

Andreyeva, Long, and Brownell (2010) completed an inventory of all empirical US-based food demand work over the past 70 years and found demand of fruit & vegetables have mean own-price elasticities of 0.70 and 0.58, respectively. Consumers are historically most price sensitive to purchases of food away from home (elasticity of 0.81) while they are least price sensitive to egg purchases (elasticity of 0.27). Estimated parameters for vegetables are historically consistent, suggesting that consumer price sensitivity with respect to vegetable purchases has changed very little in the past 70 years. Fruit demand, on the other hand exhibits a greater range of elasticity parameter estimates. Durham and Eales (2010) inventoried 16 studies on food demand in the US and found that while studies based on market-level data rendered relatively price inelastic fruit demand, those using detailed retail-level data exhibited more elastic fruit demand. In a pedagogical article on food demand, Herrmann and Roeder (1998) support the need for disaggregated consumption data in estimating consumer food demand. They argue that because food retailers in industrialized nations compete with one another on a local level by adopting active pricing strategies, the point-of-sale consumption will be price elastic compared to the aggregate-level market demand.

The household production process (Becker 1965) implies that higher income households will be less price-sensitive due to their relatively high searching costs compared to low-income households. However, to be responsive to prices, consumers have to be aware of the distribution of prices, which requires considerable time and psychological effort. Households with vehicle access and storage space, thus higher income households, may have greater ability to take advantage of price deals as well as to stockpile (Hoch, Kim, Montgomery, and Rossi 1995). Based on this literature as well as the aforementioned findings of Durham and Eales (2010) and Herrmann and Roeder (1998), I can expect that my disaggregated sample data on fruit & vegetable purchases will exhibit relatively high elasticity even for higher-income households. Further, because my consumption data represents a cross-section, elasticities will tend to be more elastic.

[5] USDA ERS, Commodity & Food Elasticities database. July 2012. http://www.ers.usda.gov/data-products/commodity-and-food-elasticities/demand-elasticities-from-literature.aspx. 28 Feb 13.

Durham and Eales (2010) gathered price and sales data for several fruits from two supermarkets in Portland, Oregon and found all fruits were own-price elastic excepting bananas, but not significantly so. Their sample consisted of individuals across a wide income distribution, which supports the generality of their findings. A 1995 study by Hoch et al. (1995) of Dominick's Finer Foods, a major Chicago-based supermarket chain, found food and beverage categories have own-price elasticities above unity. These authors analyzed 160 weeks of scanner data for 18 product categories (including twelve food categories) at 83 store locations. These results suggest price-elastic consumer food purchasing behavior at the retail level.

2.3.4 Review of Fruit & Vegetable Demand and Income

Numerous studies have analyzed income constraints on fruit & vegetable consumption. While low-income households generally spend a lower share of their budget on fruits & vegetables than higher income-households, the literature is unclear as to how sensitive fruit & vegetable expenditures are to changes in income or prices. Raper, Wanzala, and Nayga (2002) compared demand elasticities for nine food groups between low-income and higher-income households. Their data consisted of 3,577 household food expenditure diaries and demographics from the 1992 Bureau of Labor Statistics Consumer Expenditure Survey (BLS CES). Fruit & vegetable purchases were equally sensitive to income changes, but more elastic to own-price changes for low-income households compared to higher-income households. Blisard, Stewart, and Jolliffe (2004) examined fruit & vegetable demand behavior across 5,000 households using the BLS CES from 1991 and 2000. These authors compared the expenditure distributions on fruits & vegetables between low-income households' (below 130 percent of the poverty line) and higher income households using a stochastic dominance test and found that per capita fruit & vegetable expenditures of higher income households are always greater than those of the lower income households. Their finding was robust for all fresh and processed fruits & vegetables. Further, low-income households were more likely to spend an additional dollar of income on food groups other than fruits & vegetables or to other non-food products compared to higher-income households. While quality differentials on the price margin of the fruits & vegetables could explain the expenditure differences between low and higher-income households, the authors note a 1995 study by Krebs-Smith that found lower income households consume smaller quantities of fruits & vegetables.

Studies by Wilde, McNamara, and Ranney (1999 and 2000) found that household recipients of food stamps are more likely to consume meats, added sugars, and fats than they would without the program, and receipt of food stamps are not associated with higher levels of fruit & vegetable consumption. These studies suggest lower income households consume less fruits & vegetables in aggregate, but are also less likely to substitute fruits & vegetables for non-fruits & vegetable foods if the relative prices of fruits & vegetables fall, or if they are subsidized through food stamps. Thus, for low-income households, fruit & vegetables appear to have very low price elasticity of demand.

2.3.5 Review of Fruit & Vegetable Demand and Food Geography

While income has a significant effect on food consumption behavior, food access policy aims to influence consumption behavior among low-income households through the food environment. Previous research suggests proximity, in addition to income, is correlated with fruit & vegetable consumption. Weatherspoon et al. (2013) found food desert residents of Detroit will purchase fresh produce if it is made available at affordable prices. Using retail-level receipt data from a non-profit produce grocer, they found income elasticities to be significantly more elastic for residents of a Detroit food desert than for average American consumers. In other words, income, used as a proxy for expenditure, played a significant role in determining the purchasing behavior of consumers. Their study found purchases of fruits would increase by over one percent for a one percent increase in income, suggesting that increasing household income, or lowering fresh fruit prices can play a substantial role in increasing fresh fruit consumption of food desert residents. Their study further suggests that residents of a Detroit food desert exhibit significantly different fruit purchasing behavior than the average American.

Ver Ploeg et al. (2009) examined food purchases among food stamp recipients from June 1996 through January 1997 to measure relationships between supermarket access and household purchases of perishable and non-perishable fruits, vegetables, and milk. Results of their Tobit censored regressions revealed households that did not shop mainly at a supermarket purchased significantly less perishable vegetables and fruits compared to households that shopped frequently at supermarkets. The study, however, lacked geographic data on food retail access. The researchers relied on respondent's stated shopping frequency and the type of store at which they most often shopped to proxy for food retail access.

Results from the 1996-1997 National Food Stamp Program Survey by Rose and Richards (2003) showed that food stamp recipient households living greater than five miles from their principal food store consumed 63 grams of fruit and 36 grams of vegetables less per day than those living within one mile of their principal food store. These measurements translate into approximately 0.8 servings of fruit and 0.5 servings of vegetables less per day for food stamp recipients living over five miles from their main food store.

Researchers in the UK completed one of the earliest "before and after" studies on food access and health in Leads, England in 2002. In 2000, the British supercenter Tesco opened a 70,000 square foot retail store in the highly deprived local authority housing estate area, Seacroft on the edge of Leeds. Wrigley, Warm, Margetts, and Whelan (2002) reviewed resident survey data on food consumption before and after the opening of Tesco. Respondents with the lowest initial levels of fruit & vegetable consumption (less than one portion per day) increased their fruit & vegetable consumption from an average of 4.1 portions per week to 9.8 portions per week. A broader analysis on respondents who initially consumed less than two portions of fruits & vegetables per day showed similar

results: average fruit & vegetable consumption for this group increased from 9.1 portions per week to 12.3 portions per week. The UK study lends convincing evidence to the impact access can have on improving diet. US-based research, however, remains inconclusive.

2.3.6 Contributions

The methods developed in this thesis introduce preferences into the identification of constraining food geographies. I specifically contribute to the current literature on food policy and access through the following: First, I utilize first-hand survey data on a comprehensive set of fresh food purchases and observed store-specific price data to infer fresh produce demand and consumption behavior. Second, while previous food access literature largely focuses on consumption differences between low and high income households, I attempt to identify the effect of food geography on consumption independent of income levels or significant demographic variation. The respondent sample under review is comprised of a homogenous cross-section of upper middle-class, educated households thus I am confident that any behavioral variation correlated to food geography is either a result of the exogenous environment itself, or personal preferences unrelated to income constraints or highly variant demographics. Third, by combining a stated preference analysis with a demand system, I seek to detach individual preferences from observed consumption behavior in order to isolate the extent to which proximity can constrain consumption of fresh produce. Thereby, I introduce a methodology that can address causality issues pervasive in pre-existing literature. Finally, previous research on food access has largely ignored produce quality as an important determinant of consumption behavior. I explicitly account for produce quality variation to test how lack of access to high quality produce can inhibit fresh produce consumption in Champaign County.

Chapter Three: Methodology & Theoretical Framework

3.1 Introduction

This thesis asks whether food geography constrains consumption of fresh produce. To empirically test my hypothesis I explore the relationship between food geography and behavior using two methods: The first method analyses current food consumption behavior and asks how price sensitivity to fresh produce as well as the propensity to substitute between fresh produce and other foods relates to proximity to grocery stores. The second method tests how consumer willingness-to-pay (WTP) for fresh produce changes based on accessibility of the produce, the level of its quality, and the type of store selling the produce, and how these changes are informed by proximity to food stores.

Using these methods, I explore whether food deserts have constraining effects under the common exogenous USDA classification as well as a stricter "endogenous" classification. Under the exogenous classification, I compare results between "1 mile or less" respondents and "Over 1 mile" respondents in order to test whether food geography has any effect on food consumption. Under an endogenous classification, I again compare results between two subsamples, but I refine these subsamples to increase the likelihood of observing behavioral differences caused by the food environment. The first subsample is comprised of respondents who live within one mile of a grocery store who also reveal explicitly that they prefer their current grocery shopping scenario, the "Constrained" group. The second subsample, by contrast, is comprised of respondents who not only live over one mile of a grocery, but also reveal explicitly that they do not prefer their current grocery shopping scenario, the "Ideal" group. The endogenous classification refines the USDA exogenous classification to a smaller sample of respondents who reveal that they are either satisfied or unsatisfied with their current food geography. The dissatisfied respondents are clearly constrained by their food geography because they meet the USDA definition of "food desert" residents with respect to their geographic proximity, and they explicitly state their desire for a new situation. The endogenous classification restricts any statistical noise that may arise by removing those respondents who appear ambivalent about their current food geography. I expect the endogenous classification to reveal differences in consumption behavior if food geography, indeed, can effect food consumption.

As an additional robustness check on constraining effects of food geography, I explore the behavior of respondents who live within one mile of a grocery store, but choose to grocery shop at a store far away from their residence. I refer to this group of respondents as those who "Opt to Travel." Such respondents represent a population of individuals who might display identical behavior and preferences regardless of their proximity to a grocery store. I compare behavior and preferences of these respondents who willingly travel, to the "Constrained" respondents who must travel over one mile to access a grocery store in order to test how necessitated travel as opposed to willing travel can affect consumption behavior. The following sections present the methodology used to implement these analyses.

3.2 Demand Model: The Almost Ideal Demand System

A system of demand for grocery products provides a method for comparing consumption behavior between consumer groups. The food demand model for this thesis relies on a system-wide estimation, where I approach fruit & vegetable demand as simultaneously dependent upon demand for all other groceries. The theory of multistage-stage budgeting and the related concept of weak separability allow me to estimate grocery demand as a function of household grocery expenditures alone (Deaton and Muellbauer 1980b).

Before determining the structure of the food demand system, one must determine how budget shares are affected by income. Demand system specification is grounded in identifying the relationship between total expenditure and budget shares for each good in the demand system. When plotted on a graph, the relationships between total expenditure and budget shares are known as Engel curves. I analyzed Engel curves using linear, quadratic, and locally weighted non-parametric regressions of food budget shares on the log of total weekly grocery expenditures to distinguish between linear (rank 2) and quadratic (rank 3) functional forms of the demand system.

For purposes of specifying the appropriate function form, I aggregated food expenditures into fresh fruit, fresh vegetables, frozen & canned fruit, frozen & canned vegetables, and other foods to evaluate the impact of grocery store access on non-perishable produce consumption. Ver Ploeg et al. (2009) similarly disaggregate fruits and vegetables into canned and non-canned to test whether households with limited access to supermarkets tend to spend proportionally less of their food budget on perishable foods. The linear and quadratic Engel Curves for these five food categories are presented in Figure 2: Linear and quadratic Engel Curves of food budget shares on log expenditureFigure 2. Based on the results of the regression lines, there does not appear to be significant deviation from linear behavior, suggesting demand for all food commodities can be represented by rank 2, or linear demand. I disaggregated the non-perishable fruit & vegetable categories into frozen fruit, canned fruit, and frozen vegetables, and canned vegetables, for a total of seven share equations in the final demand system. Engel curves for each of these disaggregated groups behave identically when separated as when combined. However, I separate the non-perishable groups in order to test for differential substitution behavior between fresh and frozen, and fresh and canned fruits & vegetables.

Linear demand estimation follows the AIDS specification of Deaton & Muellbauer (1980a):

$$w_n = \alpha_n + \sum_j \gamma_{nj} logp_j + \beta_n \log(\tfrac{x}{P}) \quad (1)$$

where the price index, P in log form is given by:

$$logP = \alpha_0 + \sum_k \alpha_k logp_k + \tfrac{1}{2}\sum_j \sum_k \gamma_{kj} logp_k logp_j \quad (2)$$

Budget shares are denoted by w_n where n is the index of food categories 1 through 7. Total food expenditure is denoted by x and price of food categories is denoted by p_j where j is a free index on each food category 1 through 7. The α_n, β_n and γ_{nj} terms are parameters that satisfy the following linear demand restrictions:

$$\text{Adding-up } \sum_n \alpha_n = 1 \quad \text{and} \quad \sum_k \gamma_{kj} = \sum_j \beta_j = 0$$

$$\text{Homogeneity } \sum_j \gamma_{kj} = 0$$

$$\text{Symmetry } \gamma_{jk} = \gamma_{kj}$$

The adding-up restrictions force total expenditure to be equal to the sum of individual expenditures for each food category. Homogeneity of degree zero means that if all prices and income both double, demand does not change, ceteris paribus. The symmetry condition states that the cross-substitution effect between good k and good j must be the same as the cross-substitution effect between good j and good k. I have set the α_0 term equal to 3, the minimum log expenditure observed in the sample data (Banks, Blundell, and Lewbell 1997).

The demand system parameters α_n, β_n and γ_{nj} were estimated in Stata 11[6] by applying nonlinear seemingly unrelated regression to the system of share equations. The adding-up restrictions result in singularity of the error covariance matrix, thus I dropped the seventh "Other Foods" share equation for estimation purposes to avoid covariance matrix singularity. The estimator is a feasible nonlinear generalized least squares estimator executed using the Stata command nslur.

Elasticity measures result from the share equations and parameter estimates as follows:

Income elasticity:

$$\eta_n = {(\beta_n + w_n)}/{w_n} \tag{3}$$

Marshallian "uncompensated" cross price elasticity:

$$\varepsilon_{nj} = {(\gamma_{nj} - \beta_n \alpha_j - \beta_n(\sum_j \gamma_j ln p_j))}/{w_n} \tag{4}$$

Hicksian "compensated" cross-price elasticity:

$$s_{nj} = \varepsilon_{nj} + (\eta_n * w_j) \tag{5}$$

Elasticities were calculated at sample medians. Because Marshallian cross-price elasticity accounts for substitution effects as well as income effects resulting from a good's price change, I use the Marshallian elasticity to analyze substitution behavior resulting from price and cost variation of fresh produce.

[6] StataCorp. 2009. *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.

Prices p for each food category were calculated based off of an average representative basket of products as follows: For the fresh fruit & vegetable categories as well as frozen and canned fruit & vegetable categories, I calculated the average amount of fruits & vegetables purchased per week across respondents (e.g. one pound of bananas, two pounds of apples, one pound of carrots, one bag of frozen green beans). Respondents provided their grocery purchase history in terms of the week or month prior to the survey depending on which time frame was easier for them. To calculate average weekly fruit & vegetable purchases, I normalized all responses into weekly quantity terms. I then calculated the price of these average quantities of fresh fruits, fresh vegetables, frozen and canned fruit, and frozen and canned vegetables at each major grocery store in Champaign County using observed price data from these stores. Each respondent provided information on the allocation of their grocery expenditures across all food retailers (e.g. respondent A spent 10 percent of their food expenditures at Strawberry Fields, 50 percent at Schnucks, and 40 percent at Meijer). Based on these allocations, I calculated a price index for each food category for each respondent. Respondents of my survey were asked specific questions on quantity purchased only for fruits & vegetables. Thus, the quantity data I used to generate a price index for the "other foods" category is based on average weekly food consumption among adults 19 to 50 years of age from the USDA's 2001-2002 National Health and Nutrition Examination Survey (NHANES). NHANES data provides average weekly quantities for grains including snacks, cereals, and breads; beverages; dairy products; proteins; and other food items such as condiments, oils & fats, candy, frozen entrees, etc. (Carlson, Lino, Juan, Hanson, and Basiotis 2007). Prices for these foods are based on observed price data for each major grocery store in Champaign County. The resulting data set of respondents who provided information on their grocery purchases consisted of 886 observations for all variables in equation (1).

3.2.1 Expected Results of Demand Model Analysis

One mechanism by which lack of access to grocery stores could inhibit people from consuming fresh produce is through limiting the frequency of shopping trips to a distant store. If this mechanism holds, I expect respondents who live far from a grocery store to exhibit both income- and price-sensitivity to fresh produce because their opportunity cost of purchasing perishable foods will be higher relative to respondents who live near to a grocery store and can grocery shop more frequently. Income- and price-sensitivity will manifest through a relatively elastic income- and price- elasticity of demand. Conversely, I expect these respondents to have relatively inelastic income- and price- elasticities for non-perishable foods if non-perishables provide a basis to their diet. This hypothesis follows the findings of Weatherspoon et al. (2013) who found "food desert" residents had a high income elasticity of demand for fresh fruit; as well as Ver Ploeg et al. (2009) who found households who grocery shop infrequently spend a greater share of their grocery bill on non-perishable foods.

Fresh produce share categories may exhibit higher income- or price- elasticities because the respondent prefers non-perishable goods to fresh produce, and treats the fresh produce consumption as low priority, residual

consumption. However, if a respondent would prefer to purchase fresh produce, but is constrained by their distance to a grocery store, I expect that they will substitute canned or frozen fruits & vegetables for fresh produce to a greater extent than respondents who live near to a grocery store. In other words, the cross-price elasticity of demand between fresh vegetables (fresh fruit) and canned or frozen vegetables (canned or frozen fruit) will be negative and elastic, indicating that if the price of fresh produce increases, the demand for non-fresh produce increases. This action would imply that the extra distance to a grocery store forces the respondent to compromise their desired consumption of fresh produce for non-fresh produce.

I anticipate that income-, price-, and cross-price elasticities of demand to be similar in magnitude and sign for the "Constrained" and "Over 1 mile" groups. The income-, price-, and cross-price elasticities of demand I expect to be comparatively inelastic for both the "1 mile or less" and "Ideal" respondent groups. The demand system will illustrate how food geography, alone, may compromise fresh fruit & vegetable consumption. I anticipate this effect will be similar for all respondents living far from a grocery store relative to those living near to a grocery store. However, results of the choice experiment will provide some indication as to how each of these groups "feel" about those compromises, which I expect will demonstrate that the "Constrained" group are significantly less pleased with the compromise than are the "Over 1 mile" group.

3.3 Choice Experiment

3.3.1 Theoretical Framework

Grocery shopping involves several decisions including where to shop, what to buy, and how much to spend. Because grocery stores are 'lumpy goods,' these decisions are constrained not only by budget, but also by the food retail options available, particularly in "food deserts" where the array of food stores is limited. To illustrate, a consumer living in downtown Urbana with $20 to spend can choose to shop for apples at Wal-Mart, which is 15 minutes away, Schnucks, which is eight minutes away, or the Food Co-op, which is three minutes away. Figure 3 provides a graphical illustration. Each of these stores offers the consumer a unique combination of the quantity of apples they can buy with $20 based on the apple price, and time spent traveling to the store, or, viewed differently, time saved. The consumer must decide where to shop based off of their inherent value of time and desire for monetary savings. Simultaneously, the consumer must make a tradeoff between the quantity of apples they can buy and the quality of the apples. For example, assume the Food Co-op has the highest quality apples, but the most expensive, whereas Wal-Mart has the lowest quality apple, but the least expensive. The consumer can purchase more apples at Wal-Mart, but must forgo quality for higher quantity if they choose to shop there. See Figure 4 as an example. The consumer must also consider the travel time required to get to the Food-Coop, which may be a short trip versus Wal-Mart, which may be a longer trip on the expressway. Further, the consumer must decide on the type of store they shop at, taking into consideration size, ease of navigating the store, the amount of non-food items and services available, and other factors such as ownership structure. The Food Co-op has the

most expensive produce and is also the smallest store with few services and non-food items, while Wal-Mart is the largest option with a wide array of non-food items and services. See Figure 5 as an example. In each of these decisions, the consumer faces a discrete choice among the food retail opportunity set. "Food desert" residents are theoretically forced to trade off distance for quality, price, store amenities, and several other factors to an extreme. I investigate whether the obligation to make such trade offs has any effect on produce consumption.

Such tradeoffs are impossible to observe from individual behavior because the consumer's choice of grocery store is predicated on their internal, unobservable rankings of importance among distance, quality, price, and several other factors. Consequently, I rely on a stated preference approach to parse consumers' partiality across these factors.

Stated preference methodologies enable economists to estimate welfare changes that are usually hypothetical and cannot be observed in revealed behavior. This thesis utilizes a type of stated preference method known as a choice experiment (CE) to evaluate people's willingness to pay for various grocery store attributes. A choice experiment provides the framework to simulate how survey respondents choose among a set of grocery store alternatives. In a typical choice experiment the decision maker, or respondent, faces a series of questions, each offering the respondent a choice between several alternatives with differing characteristics. The respondent decides which of the alternatives they most prefer. The respondent answers a series of such questions, whereby the attributes of each alternative differ slightly according to the experimental design. By varying the attribute levels for each round, the researcher can estimate the respondent's willingness to substitute between attributes. The consumer's choice is then modeled in probabilistic terms (Greene (a)). Monetary cost can be used as one of the attributes to calculate the implicit price or willingness to pay (WTP) for a change in a particular attribute. Such calculations also enable the researcher to calculate respondent's WTP to move from their current situation to a given choice scenario. The framework of a choice experiment needs to be realistic enough to explain an individual's previous choices, and should be capable of assessing the support for alternatives not presented in the choice set. The choice set generation should have mutually exclusive, exhaustive, and finite alternatives. Consequently, the experiment designer must ensure the range of alternatives and influences affecting the decision to choose one alternative over another, such as demographic variables, are accounted for (Hensher, Rose, and Greene 2005; Greene (b)). The designer also sets levels within each attribute such that the levels may imply meaningful changes in utility for changes in the attribute (Alpizar, Carlsson, and Martinsson 2001).

3.3.2 Survey Design

I developed a survey to measure the value of food access and identify preferential and environmental factors determining fruit & vegetable consumption. The survey asked respondents several questions pertaining to their grocery shopping activities from the month prior to the survey, roughly December 2012 to early January 2013, as

well as their typical food consumption and purchasing routines. The choice experiment component presented respondents with six sets of hypothetical grocery stores and asked respondents to indicate their preferred shopping scenario based on variation in these grocery store attributes: store type, travel distance to the store, variety and quality of produce, and price of one week's worth of groceries from that store. Respondents could also opt to choose neither hypothetical grocery store, and instead choose their current grocery store as their preferred shopping scenario. At the start of the choice experiment, respondents were instructed to select their preferred grocery store based off of where they would routinely shop for groceries. Figure 6 provides the varied levels of each of the four grocery store attributes. Figure 7 provides an example of one of the choice questions.

The travel time and produce attributes were motivated by my intent to measure a trade off between travel time, produce quality, and variety. I left the definition of "quality" in the produce quality and variety attribute undefined intentionally to ensure that quality perception remained individual-specific across respondents. Outside of the choice experiment questions, respondents provided information on their own definition of produce quality by ranking the relative importance of taste, freshness, growing method, seasonality, and shelf life. I varied the travel time attribute enough to ensure that respondents would perceive an increase from five minutes to 15 minutes as a significant increase in searching costs, which Hoch et al. (1995) define as a function of the physical distance and driving time between stores. Other attributes I identified through informal interviews with residents of Champaign County and a focus group involving graduate students of the University of Illinois, as well as examining literature on grocery store preferences and determinants of store choice (see Section 2.2.4). I tested the online survey with graduate students of the University of Illinois on multiple occasions and modified the final survey in response to test-takers' feedback. I designed the online survey through the survey platform Qualtrics.[7] The Division of Management Information at the University of Illinois distributed the survey via email to 5,000 faculty and staff of the University of Illinois at Urbana-Champaign. Out of 5,000 individuals who received the survey, 991 provided answers resulting in a 19.8 percent response rate.

To decide the exact ordering and combination of attributes and levels presented to each respondent, I followed standard practice to create a fractional factorial experimental design (Kuhfeld 2010; Louviere et al. 2000; Hensher et al. 2005). Such a design ensures that variations of choices presented to respondents are uncorrelated. I used Kuhfeld's SAS macro (Kuhfeld 2010) to generate a design that achieves a 100 percent D-efficiency and can be implemented with a subset of 36 profiles.[8] I blocked the 36 choice profiles into three sets, resulting in three unique surveys. Each survey asked respondents six choice questions where each choice question consisted of two choice

[7] Qualtrics 2013. Provo, Utah. Version 37,892. <www.qualtrics.com>

[8] A 100% D-efficiency indicates a design is balanced and orthogonal. A design is balanced when each level occurs in the design equally often within each attribute. A design is orthogonal when every pair of levels occurs equally often across all pairs of attributes. Stated more generally, a design is orthogonal a when the frequencies for level pairs are proportional or equal. (Kuhfeld 2010 p. 58)

profiles plus a status quo option. The status quo option represented the respondent's current principal grocery store. The SAS code used to generate my design as well as the SAS output are presented in Appendix A. Of 991 survey respondents, 863 subjects completed the choice experiment, each of which answered six choice sets with three alternatives. In choice experiments, the unit of observation is each unique choice alternative as opposed to each respondent. Thus the surveys provided a total of 863 x 6 x 3 = 15,534 observations for the econometric analysis (Kuhfeld 2010 p. 292).

3.3.3 Econometric Methods

Econometricians often rely upon structural models of behavior called random utility maximization models to model and interpret discrete choice. Random utility maximization theory asserts that various characteristics of a good endow individuals with utility along with a random component, which captures unobserved differences. The utility U_j associated with alternative j is comprised of a systematic, representative utility V_j and a stochastic, unobservable element of utility ε_j as follows:

$$U_j = V_j(x_j, \beta) + \varepsilon_j \tag{6}$$

where x_j is a vector of observable attributes of alternative j and β is a vector of parameters based on observable attributes of the decision-maker. Assumptions on distribution of the unobserved portion of utility determine what type of discrete choice model is used. The multinomial (MNL) and conditional logit (CL) models are widely used for discrete dependent variable analysis because their resulting choice probabilities take on a convenient form (Heiss 2002).

Random utility maximization theory assumes that consumers choose the alternative that endows them with the highest utility (Greene (b)). Suppose we may assign a utility level U_{ij} to each alternative j = 1, . . . , J for each decision-maker i = 1, . . ., I. The utility levels are determined by characteristics of the choice alternatives as well as characteristics of the decision-maker, some of which are identifiable and some of which are unknown. The MNL model is a special case of a CL model in which all explanatory variables are assumed to be individual-specific. In a CL model, a choice among alternatives is treated as a function of the characteristics of the alternatives as opposed to a function of characteristics of the individual decision-makers, thus each β term will be identical across respondents (Kuhfeld 2010). If we consider each choice situation for each respondent, the utility gained by person i for alternative j is given by

$$U_{ij} = \beta X_{ij} + \varepsilon_{ij} \tag{7}$$

where X_{ij} is a vector of attributes of alternative j relative to individual i and β, a vector of parameters, acts as a scalar relating the relationship between individual i and the alternative X_{ij} to the individual's utility for the alternative. The parameters of vector β are not alternative specific, thus are not subscripted by j.

The MNL model, however, allows the effect of each independent variable (i.e. attributes of the decision-maker) to vary across choices. The basic MNL utility equation of person i for alternative j is given by:

$$U_{ij} = z_i\gamma_j + \varepsilon_{ij} \tag{8}$$

The vector γ_j of alternative-specific parameters relate the characteristics of a respondent to the respondent's utility for the jth choice. The parameter z_i is a matrix of individual characteristics that are independent of the alternatives available. While X_{ij} from the CL model is a vector of alternative-specific characteristics, $z_i\gamma_j$ from the MNL model is a vector of individual-specific characteristics relating the characteristics of a respondent to the respondent's utility for the jth choice.

The mixed multinomial logit model (MMNL) combines the MNL and CL models so that the effect of individual characteristics as well as choice characteristics on the probability of choosing an alternative may both be identified simultaneously. The MMNL model assumes that the unobserved portions of utility may follow any specified distribution thus relaxes the CL assumption that respondents are homogeneous with regards to their preferences (Dissanayake 2011). The probability P_{ij} that an individual i will choose alternative j from the choice set of m total alternatives does not have a closed form solution, and must be evaluated numerically through iterative simulations. Utility gained by person i for alternative j takes the form

$$U_{ij} = z_i\gamma_j + \beta X_{ij} + \varepsilon_{ij} \tag{9}$$

where z_i is a vector of individual-specific independent variables relevant for individual i and γ_j is a vector of alternative-specific parameters. Together, these parameters relate the characteristics of a respondent to the respondent's utility for the jth alternative such that the effect of the independent variables will vary across all choices (Green (b)). The MNL model assumes that the unobserved portions of utility ε_{ij} are independently and identically distributed. Thus, the unobserved portion of utility for one alternative is independent of the unobserved portion of utility for other alternatives. McFadden (1974) demonstrated that under these assumptions, the probability P_{ij} that an individual i will choose alternative j from the choice set of m total alternatives has a closed form solution given by

$$P_{ij} = \exp(U_{ij}) \Big/ \sum_{k=1}^{m} \exp(U_{ik}) = \exp(z_i\gamma_j + x_{ij}\beta) \Big/ \sum_{k=1}^{m} \exp(z_i\gamma_k + x_{ik}\beta) \tag{10}$$

The unknown parameter β can be estimated by the maximum likelihood procedure whereby one finds the vector $\hat{\beta}$ that maximizes the joint probability of observing the observed choice outcomes of all consumers in the sample (Heiss 2002; Arnold et al. 1983).

The coefficient estimates for the CL/MNL or MMNL models alone are difficult to interpret directly. Instead, the ratio between the non-monetary parameters and the cost parameter provide the marginal WTP for a change in each attribute η. The average marginal WTP is given by

$$MWTP_{\eta} = -\frac{\beta_{\eta}}{\beta_{bill}} \quad (11)$$

3.3.4 Econometric Specification

I use an MMNL model to estimate parameters from the choice experiment data because the MMNL is capable of modeling preference heterogeneity across respondents and accounts for possible correlations between individuals' survey responses. I employ two specifications. The first MMNL specification employs only main effects. The "main effects" refer to the four grocery store attributes in my choice experiment, including cost of a week's worth of groceries, travel time required to get to the store, store type, and the produce selection. The econometric specification is given by:

$$U_{ij} = \alpha_j + \beta_1 X_{iBill} + \beta_2 X_{iTravel} + \beta_3 X_{iSpecialty} + \beta_4 X_{iGrocery} + \beta_5 X_{iHQProduce} + \beta_6 X_{iAQProduce} + \varepsilon_{ij} \quad (12)$$

The second specification employs additional interaction terms given by:

$$U_{ij} = \alpha_j + \beta_1 X_{iBill} + \beta_2 X_{iTravel} + \beta_3 X_{iSpecialty} + \beta_4 X_{iGrocery} + \beta_5 X_{iHQProduce} + \beta_6 X_{iAQProduce} + \beta_7 X_{iBill} * \gamma_{i\,CurrentTravel} + \beta_8 X_{iBill} * \gamma_{i\,Distance} + \beta_9 X_{iHQProduce} * \gamma_{i\,Income} + \beta_{10} X_{iAQProduce} * \gamma_{i\,Income} + \varepsilon_{ij} \quad (13)$$

where α_j is an alternative specific constant (ASC). The ASC term captures the average effect on utility of all factors that are not included in the model (Greene (b)). The ASC term is a dummy variable equal to 1 if the respondent selected their current situation status quo option, and equal to 0 if the respondent chose a hypothetical option A or B. To account for the potential tendency for people to choose their current store I include an ASC to capture attributes affiliated with the status quo choice that I did not explicitly account for in the experimental design. Tang et al. (2001) found in a sample of 500 households that knowledge of store layout was a significant source of utility of grocery shopping. Each X term corresponds with one of four grocery store attributes included in the choice experiment: weekly grocery bill (*Bill*), travel time to the store (*Travel*), store type (*Grocery* and *Specialty*), and produce quality and availability (*HQProduce* and *AQProduce*). The *Grocery* and *Specialty* terms are dummy variables that will both equal 0 if the respondent chose a supercenter. Similarly, *AQProduce* ("average quality produce") and *HQProduce* ("high quality produce") are dummy variables that will both equal 0 if the respondent chose a store with low quality produce. The model based on equation (13) includes four γ terms that interact the grocery bill and produce quality attributes with person-specific variables including income level (*Income*), proximity to nearest grocery store (*Distance*), and reported travel time to their principal grocery store (*CurrentTravel*). The error term $\varepsilon_{ij}$is independent and identically distributed extreme value type I with a zero mean by construction due to the inclusion of the ASC variable (Greene (b)).

Results of the MMNL yield estimates of the median and standard deviations of each of the coefficients' distributions. If any of the attribute coefficients are significant and positive, this implies that the respondent has higher marginal utility for increases in these attribute levels. Negative attribute coefficients imply disutility. I expect the *Bill* and *Travel* attributes to be negative, for example, because respondents presumably get disutility from paying extra for groceries or traveling further, ceteris paribus. The interaction term $\beta_7 X_{iBill} * \gamma_{i\ CurrentTravel}$ identifies the impact of a person's reported travel time to their principal grocery store with their willingness to choose a grocery store with a higher grocery bill. This interaction term controls for potential nonlinearity between people's marginal WTP for groceries and the time they travel to get to the grocery store. The interaction term $\beta_8 X_{iBill} * \gamma_{i\ Distance}$ controls for nonlinearities between a respondents' marginal WTP for groceries and their proximity to a grocery store. The $\beta_9 X_{iHQProduce} * \gamma_{i\ Income}$ and $\beta_{10} X_{iAQProduce} * \gamma_{i\ Income}$ terms identify the impact of income level on WTP for produce quality. Positive and significant coefficients imply respondents with higher income gain higher utility from high quality or average quality produce over low quality produce.

By differentiating the utility equation U_{ij}, the β terms represent the marginal utilities of each attribute. The ratio between the non-monetary parameters and the cost parameter $-\beta_1$ provide the marginal value of each attribute. For example, the marginal value of shopping at a specialty store instead of a supercenter based on equation (13) is as follows:

$$MWTP_{Specialty} = -\frac{\partial U/\partial X_{Specialty}}{\partial U/\partial X_{Bill}} = -\beta_3/(\beta_1 + \beta_7 * \gamma_{i\ Distance} + \beta_8 * \gamma_{i\ CurrentTravel}) \quad (14)$$

The econometric models were estimated using Stata's clogit and mixlogit commands. Results of the model are presented and discussed in Chapter Five.

3.3.5 Expected Results of Choice Experiment Analysis

Both endogenous sorting as well as environmental treatment effects can influence respondent preferences for grocery store attributes. If endogenous sorting is the driving influence of consumption behavior differentials across respondent groups, I expect their preferences, and thus their marginal WTP for grocery store attributes to be different across groups. For example, I expect that respondents who live near to a grocery store will have a low or negative marginal WTP for added travel time if their proximity to a grocery store is a result of their desire to shorten commutes to the grocery store. The "Over 1 mile" respondents, by contrast, I would expect to have a higher or less negative marginal WTP for added travel. Because greater travel distance to a grocery store increases the opportunity cost of purchasing perishable produce instead of non-perishable food items, I expect the "Over 1 mile" respondents to have a lower marginal WTP for produce quality compared to the "1 mile or less" group. I am agnostic on my prediction of respondent preferences for store type. However, I do expect the "Constrained" group to have a higher marginal WTP for a store type different from that which they typically shop at, and the converse

to be true for the “Ideal” group. For example, if I find that the “Constrained” group mainly shops at supercenters, I would expect their marginal WTP to shop at a grocery or specialty store to be high relative to other groups.

On the other hand, if environmental treatment effects, i.e. the exogenous food geography, are driving consumption behavior across respondents, then I do not expect to see significant differences in their preferences. Further, I would expect their consumption behavior to be inconsistent with their preferences. For example, if “Over 1 mile” respondents exhibit high substitution between fresh and non-fresh produce, and very high marginal WTP for produce quality, this would imply these respondents prefer quality produce, but must compromise fresh for non-fresh produce due to an exogenous constraint. Such a result would provide evidence that geography, independent of preferences, can alter produce consumption. I expect the “Constrained” and “Ideal” group comparison to exhibit greater behavioral differences than the “Over 1 mile” and “Less than 1 mile” respondent group comparison. Because the “Constrained” respondents are, by definition, dissatisfied with their food environment, endogenous sorting likely has little effect on their consumption behavior.

Comparing the “Opt to Travel” group to the “Constrained” group will inform whether travel time or produce quality serves as the “constraining” factor for the “Constrained” group. For example, I can infer that low access to quality produce “constrains” the “Constrained” group if their marginal WTP for travel time is similarly low to that of the “Opt to Travel” group, but their marginal WTP for produce quality is significantly higher than that of the “Opt to Travel” group.

Alternatively, if the “Opt to Travel” group has similar preferences and behavior to that of the “Constrained” group, this will suggest that the exogenous “food desert” definition overestimates the constraining effects of food geography because the one-mile cut off fails to account for individuals impervious to their food environment.

While the methods presented in this thesis do not provide a comprehensive means for disentangling the endogenous from exogenous forces impacting food consumption, they do help to inform which of these two forces may have greater influence on food consumption.

Chapter Four: The Data

4.1 Data Collection

The following section summarizes key statistical findings from the survey discussed in the previous chapter. In addition to the choice experiment, respondents provided copious information on both their grocery shopping between December 2012 and January 2013, as well as they typical grocery shopping routines. Survey data includes information on where respondents purchased groceries, how often they shopped, the travel time to these stores, their mode of transportation, their average grocery expenditures per household, as well as qualitative questions on produce quality preferences and grocery store preferences. Respondents were asked to provide detailed information on which food types they purchased at each food retailer, and the quantity of fruits & vegetables they purchased. The responses include demographic information, as well as identification of the respondent's local neighborhood in Champaign & Piatt counties. I collected price data contemporaneously from 14 grocery and food retailers in Champaign County so as to estimate individual budget shares spent on staples as well as fruits & vegetables. I also geocoded all food retailer locations in Champaign & Piatt counties and calculated the distance from each respondent's neighborhood to their surrounding food stores. Food retailer locations and industry classifications were sourced from LexisNexis database.

4.2 Summary Statistics: Demographics, Food Retail Environment, and Budget Shares

Table 1 compares demographic characteristics of the respondent sample and Champaign County. The sample is biased toward a higher income and higher education level than is representative of Champaign County. Because the marginal effects of grocery store proximity and produce consumption are of greatest interest for this thesis, this bias does not diminish the relevance of my results. Rather, the affluence of the sample ensures any effect I find of food geography on produce consumption will be conservative. Applying the methodology of this thesis to a lower-income area would likely result in stronger effects, especially in areas of lower car ownership rates, or public transportation access.

Table 1: Demographic comparison of Champaign County population and survey sample

Demographic	**Sample** Average	**Champaign County*** Average
Income (thousand)	$87.5	$59.2
Age (adults over 18)	53	39
Education (% of adults with Bachelor's degree)	99.6%	42.1%
Household size	3	2
Car Access (% of households with)	99.0%	98.8%

* Source: American Fact Finder 2010 US Census

While car access is ubiquitous across our sample, a minority of respondents reported they use alternative means to travel to the grocery store. Figure 8 illustrates the types of transportation used by our sample. Eighty-five percent of respondents drive to their grocery store, on average. Ten percent ride a public bus, and 4.1 percent walk or bike.

Out of 991 total respondents to my survey, 746 provided some information on their geographic location. Of the 746 respondents, 685 indicated they live in either Champaign or Piatt counties. Finally, of these 685 individuals, 575 provided detailed information on their neighborhood location. Subsequent analyses based on neighborhood location and proximity to food retailers will rely on this subset of 575 individuals.

Figure 9 provides a visual for respondent neighborhood locations in Champaign & Piatt counties relative to the location of food stores. Respondents to the survey live among 36 distinct neighborhoods in Champaign & Piatt counties.[9] As indicated by the dot density continuum in Figure 9 as well as Table 2, the majority of respondents live in the West Urbana neighborhood of Urbana or the Downtown neighborhood of Champaign.

Distance from respondent's neighborhoods to surrounding food retailers were estimated by generating geographic centroids for each neighborhood and calculating the distance in miles from the centroid to surrounding food retailers. All geocoding, centroid calculation, and near distance calculations were performed using Esri ArcGIS mapping software. Nearest distance to grocery stores, supercenters, and convenience stores by neighborhood can be found in Table 2. LexisNexis designated retailers as grocery stores and convenience stores following the Standard Industrial Classification (SIC) system. Because the SIC system does not differentiate supercenters from

[9] Geographic definitions for neighborhoods of Champaign & Piatt counties were sourced from Homeplace Advisors, LLC (http://yourchampaignhome.com/about-homeplace-advisors/)

grocery stores, I designated Wal-Mart, Meijer, Target, and Sam's Club as supercenters. The "convenience store" designation includes drug stores, pharmacies, dollar stores, and gas stations.

The Downtown neighborhood of the city of Champaign claims the largest portion of survey respondents, 13.9 percent of the sample, followed by West Urbana, East Urbana and Clark Park of Champaign. Residents of Savoy face the shortest distance to their nearest grocery store, supercenter, and convenience store. The geographic centroid of Savoy is 0.1 miles from the nearest grocery store and supercenter, and 0.2 miles from the nearest convenience store. Residents of Longview, by contrast, must travel the furthest to get to a grocery store. The geographic centroid of Longview is 15.3 miles from the nearest grocery or supercenter store.

Table 2: Neighborhood sample population and average distance to nearest food retailer

No.	Neighborhood	City	Sample Pop.		Census Pop.	Miles to Nearest Store		
			Number	Share		Grocery	Super center	Convenience
1	Downtown	Champaign	114	13.9%	857	0.6	1.7	0.1
2	West Urbana	Urbana	104	12.7%	4,354	0.6	1.8	0.3
3	Clark Park	Champaign	80	9.8%	675	0.9	2.2	0.3
4	East Urbana	Urbana	80	9.8%	1,705	0.4	1.7	0.1
5	Savoy		63	7.7%	7,312	0.1	0.1	0.2
6	Cherry Hills	Champaign	60	7.3%	2,451	1.6	2.3	0.5
7	Mahomet		51	6.2%	10,042	0.3	8.1	0.2
8	Ironwood	Champaign	38	4.6%	3,489	1.8	4.6	1.6
9	Sawgrass	Champaign	26	3.2%	1,118	0.2	3.1	0.8
10	Ponds of Windsor	Champaign	22	2.7%	1,048	0.7	1.6	0.8
11	Boulder Ridge	Champaign	20	2.4%	1,060	0.7	2.6	1.6
12	Campus Town	Champaign	19	2.3%	17,026	0.5	2.6	0.2
13	Monticello		19	2.3%	4,384	0.2	17.3	0.1
14	St. Joseph		17	2.1%	4,407	6.9	6.9	0.3
15	Myra Ridge & South Ridge	Urbana	15	1.8%	872	0.5	0.5	0.9
16	Stone Creek	Urbana	12	1.5%	497	1.0	1.1	1.1
17	Trails at Abbey Fields	Champaign	10	1.2%	38	1.4	4.0	0.7
18	Ashland Park	Champaign	8	1.0%	1,011	0.7	0.7	0.9
19	Savannah Green	Urbana	7	0.9%	3,492	0.5	1.1	0.4
20	Beringer Commons	Urbana	6	0.7%	1,282	0.6	0.6	0.5
21	Tolono		6	0.7%	3,460	4.7	4.7	0.3
22	Rantoul		6	0.7%	13,263	1.2	1.2	0.5
23	Sidney		5	0.6%	1,248	7.0	7.0	3.7
24	Philo		4	0.5%	1,474	5.8	5.8	0.2
25	Pesotum		3	0.4%	557	9.8	9.8	5.2
26	Ogden		3	0.4%	802	10.8	10.8	3.5
27	Longview		3	0.4%	153	15.3	15.3	0.1
28	Ivesdale		3	0.4%	267	8.3	13.2	6.5
29	Fisher		3	0.4%	1,881	8.3	8.9	0.4
30	Thomasboro		2	0.2%	1,127	4.1	4.1	0.4
31	Seymour		2	0.2%	360	5.9	9.2	6.1
32	Rolling Acres	Champaign	2	0.2%	282	2.1	2.2	0.4
33	Hensley		2	0.2%	801	3.3	3.3	3.7
34	Bondville		1	0.1%	414	3.2	6.2	3.1
35	Somer		1	0.1%	1,178	4.6	4.9	4.3
36	Sadorus		1	0.1%	412	8.0	8.0	4.7

Table 3 states the sample's average geographic distances to food retailers by type for the full sample as well as the five subsamples under review. Of all food retailers, convenience stores are highest in number totaling 140 in Champaign & Piatt counties, while supercenters are lowest in number, totaling eight in Champaign & Piatt counties. As expected, the full sample's average distance to supercenters is an extra 2.2 miles further than distance to a grocery store. Distance to a convenience store is half a mile, on average. Respondents in the "Ideal" category live closest to all store types while respondents in the "Constrained" category live furthest. The "Status Quo" column in Table 3 refers to the choice experiment survey in which respondents answered six questions asking them to choose between two hypothetical new grocery stores, or their current grocery store, e.g., their "status quo" situation. See Figure 7 as an example of the choice experiment questions. The more often a respondent chose their status quo, the more likely they were happy with their current situation, and preferred their status quo grocery shopping scenario to any other possible scenario. "Ideal" respondents chose the status quo option five times out of six, on average, whereas the "Constrained" respondents chose the status quo option approximately one time out of six, on average.

Table 3: Distance (miles) to nearest food retailers by type and 'Status Quo' selection frequency

	Retailer (Population)										
	Grocery (40)			Supercenter (8)			Convenience (140)			Status Quo*	Observations
	Mean	Min	Max	**Mean**	Min	Max	**Mean**	Min	Max		
Full Sample	**1.2**	0.1	15	**3.4**	0.1	17.3	**0.5**	0.1	6.5	3.3	886
Less than 1 mile	**0.5**	0.1	1	**2.9**	0.1	17.3	**0.3**	0.1	1.6	3.4	455
Over 1 mile	**4.3**	1.2	15	**5.2**	1.2	15.3	**1.3**	0.1	6.5	3.2	116
Ideal	**0.5**	0.1	0.9	**2.5**	0.1	17.3	**0.2**	0.1	1.6	5.0	221
Constrained	**4.1**	1.0	11	**5**	1.1	13.2	**1.4**	0.2	6.5	1.1	56
Opt to Travel	**0.5**	0.1	1	**3.5**	0.1	17.3	**0.3**	0.1	1.6	3.0	260

*Number of times out of 6 respondent chose their current grocery shopping situation over a new situation

Table 4 compares grocery shopping habits across respondent groups. Recall, all respondents shopped at a grocery store at least once in the month prior to the survey. Table 4 highlights how many respondents shopped at non-grocery store retailers in addition to a grocery store. "Specialty" food retailers include grocery stores specializing in ethnic and regional cuisine, health or natural foods, specialty produce or meat and cheese products, as well as bakeries. Respondents in the "Ideal" group had the highest probability of shopping at a specialty store. Over 65 percent of respondents in the "ideal" group indicated that they had shopped at a specialty store in the month prior to the survey. These respondents also had the least incidence of both supercenter patronage of all four groups, 48.4 percent, and convenience store patronage, 37.1 percent. In contrast, respondents in the "Constrained" group were least likely to shop at a specialty store and most likely to shop at both a supercenter and convenience store compared to all other groups. Over 89 percent of the "Constrained" respondents shopped at a supercenter in the month prior to the survey, and 48.2 percent shopped at a convenience store prior to the survey. Surprisingly, the "Constrained" group does not appear to be income constrained, as they had the highest mean and median income

of all other groups. A comparison of "Less than 1 mile" and "Over 1 mile" respondents mimics that of the "Ideal" and "Constrained" group, but to a lesser degree. Proximity to a grocery store appears to be correlated with a higher likelihood of shopping at convenience stores, and a lower likelihood of shopping at supercenters or convenience stores. Proximity is also correlated with more frequent grocery shopping. The "Ideal" group grocery shopped most often in the month prior to the survey, nearly six times whereas the "Over 1 mile" group grocery shopped the least often. Results of Table 4suggest that the respondents who are most satisfied with their food geography frequently shop at supercenters and infrequently shop at supercenter or convenience stores, whereas the opposite is true of respondents who are least satisfied with their food geography.

Table 4: Patronage of alternative food retailers, shopping frequency, and income bracket by respondent group

	Incidence of shopping at alternative retailers in past month			Grocery trips per Month*		Income Bracket^		
	Specialty	Supercenter	Convenience	Mean	Median	Mean	Median	Obs.
Full Sample	56.3%	60.3%	40.7%	5.2	4	5.4	5	886
Less than 1 mile	60.7%	57.6%	40.4%	5.4	5	5.4	5	455
Over 1 mile	52.7%	64.3%	41.8%	5.1	4	5.5	5	423
Ideal	65.6%	48.4%	37.1%	5.8	5	5.6	5	221
Constrained	50.0%	89.3%	48.2%	5.3	4	5.9	6	56
Opt to Travel	53.8%	71.2%	46.9%	4.4	4	5.5	5	260

*Number of times, on average, that the respondent shops at their main grocery store, ie the grocery store where they spent the largest portion of their grocery bill in the month prior to the survey
^Income bracket of 4 denotes annual income between $50-75K; 5 denotes $75-100K; 6 denotes $100-125K

All respondents of the survey patronized at least one grocery store in the month prior to the survey. No respondent was forced to, nor willingly shopped only at convenience stores. Approximately 40.7 percent of the sample purchased food at a convenience store at least once in the month prior to the survey. Comparing the distance to the nearest grocery between this subset of convenience store-shoppers and the entire sample, convenience store-shoppers on average live further from a grocery store, though the difference is minimal (see Figure 10). Convenience store-shoppers do not live any closer, on average, to convenience stores than does the entire sample. In fact, convenience store-shoppers live 0.6 miles on average from the nearest convenience store, compared to 0.5 miles for the entire sample. Extra distance to a grocery store may induce individuals to patronize a closer convenience store, but proximity to a convenience store itself does not seem to induce people to shop there.

Table 5 shows the prices, budget shares, zero-response frequency, and total grocery expenditure the sample incurred, on average, for a week's worth of groceries. Appendix B provides a complete list of each fruit & vegetable type for which survey respondents provided purchase information. Respondents could also provide a free

response for any fresh, canned, or frozen produce item that I did not ask about in the survey. Grocery purchases were categorized into seven broad categories for purposes of estimating the demand system: fresh vegetables, fresh fruit, canned vegetables, canned fruit, frozen vegetables, frozen fruit, and other foods. The price indices are based on the sample's average quantity purchased per category in one week, and the observed price of these items averaged across 14 grocery retailers. The non-fruit and vegetable "Other Foods" category is both the most expensive category and makes up the largest share of respondents' grocery bill, followed by fresh vegetables, fresh fruit, canned vegetables, frozen vegetables, and frozen and canned fruit. Respondents purchased "Other Foods" more often than any other category, since the number of zero entries ("Zero's") is the lowest. Frozen fruit products were purchased least often of all food categories. Average total expenditure on groceries per respondent per week was $117.6.

Table 5: Food share and expenditure summary statistics

			Budget Share			
	Price Index	**Zero's**	Mean	SD	Min	Max
(1) Fresh Vegetables	$12.9	97	15.6%	14.4%	0.0%	91.9%
(2) Fresh Fruit	$10.3	125	11.9%	12.7%	0.0%	75.8%
(3) Canned Vegetables	$1.4	554	1.6%	3.3%	0.0%	32.1%
(4) Canned Fruit	$0.4	704	0.5%	1.5%	0.0%	15.5%
(5) Frozen Vegetables	$1.2	467	1.6%	3.6%	0.0%	46.7%
(6) Frozen Fruit	$0.4	749	0.5%	1.8%	0.0%	21.5%
(7) Other Foods	$65.1	23	68.1%	24.5%	0.0%	100.0%
Weekly grocery expenditure			$117.6	$67.0	$10.0	$500.0
Obs	*878*					

Table 6 provides correlations between respondents' stated travel time to their main grocery store, the distance to their nearest grocery store, and their grocery expenditure share spent on fresh fruits & vegetables, and frozen and canned non-fresh fruits & vegetables. Travel time and distance have a positive relationship, indicating people that live further from a grocery store travel further. Consumption of fresh fruits & vegetables has a negative relationship with both travel time and distance, while consumption of non-fresh fruits & vegetables has a small, but positive relationship. These correlations provide some evidence that individuals living nearby a grocery store are more likely to consume fresh fruit & vegetables, instead of non-fresh, however I cannot infer any causal relationship based on this information alone.

Table 6: Correlation Matrix

$\rho_{x,y}$	Travel time to store	Distance to store	Budget share of fresh F&V	Budget Share of all non-fresh F&V
Travel time to store	1.000			
Distance to store	0.293	1.000		
Budget share of fresh F&V	-0.046	-0.065	1.000	
Budget share of all non-fresh F&V	0.013	0.055	0.130	1.000

4.3 Comparison of Preferences and Budget Shares

My sample provides evidence that proximity to a grocery store has bearing on food purchasing behavior and preferences. Table 7 presents budget shares of respondents who live within one mile, respondents who live more than one mile from a grocery store, as well as budget shares of respondents who patronized a convenience store in the past month. One-fifth of the sample, 22.0 percent, must travel more than one mile to get to a grocery store.

Respondents in the "1 mile or less" group spent 3.5 percent more of their grocery bill on fresh fruits & vegetables than those in the "Over 1 mile" group. These individuals spent 0.3 percent less of their grocery bill on canned produce, and 3.0 percent less on non-fruit & vegetable food items compared to the "Over 1 mile" group. This suggests that proximity to a grocery store is positively related to consumption of fresh produce relative to non-fresh produce and non- fruit & vegetable foods.

The 22.0 percent of respondents that must travel over one mile to reach a grocery store account for 75.1 percent of the convenience store-shoppers,[10] which suggests the added travel constraint on grocery store access increases the probability of shopping at a convenience store. The full sample of convenience store-shoppers have similar expenditure shares to "Over 1 mile" respondents, spending a higher portion of their grocery bill on non-fruit & vegetables and frozen or canned fruits & vegetables compared to the "1 mile or less" respondent group.

[10] Based on the sample of 575 individuals for which I have neighborhood locations, 75.1% = 181 convenience store - shoppers living more than one mile from a grocery store divided by 241 total individuals that shopped at a convenience store in the past month.

Table 7: Average grocery expenditure share: Exogenous classification

	Distance to nearest grocery		Convenience
	1 mile or less	Over 1 mile	**Shoppers**
Fresh F&V	30.1%	26.6%	27.8%
Canned F&V	2.0%	2.3%	2.6%
Frozen F&V	2.1%	2.2%	2.3%
Non F&V	65.8%	68.8%	67.3%
Obs.	459	427	364

These three groups have similar preferences for produce. Table 8 displays how respondents ranked five characteristics of produce – Freshness, Seasonality, Appearance, Growing Method used to grow the produce, and Shelf Life – based on which have the greatest bearing on their produce purchases. Freshness and Appearance are the first and second most important determinants of produce purchases across all three groups. Respondents with a grocery store one mile or less from their neighborhood hold Growing Method in higher esteem relative to Shelf Life compared to convenience store-shoppers, and those living over one mile from a grocery store. People living one mile or less from a grocery store are far more likely to attend a farmers market or partake in a CSA. Proximity to a grocery store is positively related to farmers market and CSA patronage, which itself signals a penchant for fresh, high quality produce.

Table 8: Produce preferences: Exogenous classification

	Distance to nearest grocery		Convenience
Rank of Importance	1 mile or less	Over 1 mile	**Shoppers**
First	Freshness	Freshness	Freshness
Second	Appearance	Appearance	Appearance
Third	Growing Method	Shelf Life	Shelf Life
Fourth	Shelf Life	Growing Method	Growing Method
Fifth	Seasonality	Seasonality	Seasonality
Farmers Market/CSA attendance?*	79.5%	22.2%	10.2%

**Convenience* Farmers Market/CSA attendence based on percentage of total convenience store-shoppers with neighborhood location

Respondents living within one mile of a grocery store shop more frequently (5.4 times per month) than do respondents living over one mile from a grocery store (5.1 times per month – see Table 4). Respondents within one mile of a grocery store also spend more per household member per week on groceries ($51.6) than do respondents living over one mile from a grocery store ($48.4). Household size and composition were similar for these two groups. Respondents that lived over one mile from a grocery store had households of 2.7 people with 0.9 children, on average, while respondents living within on mile from a grocery had households of 2.5 people with

0.8 children, on average. Also, both groups had similar average income levels. Thus, income constraints are not necessarily driving the residential locations between samples.

The sample suggests close proximity to a grocery store does not guarantee patronage of that grocery store. Sixty percent of respondents reported that they travel more than five minutes to get to their main grocery store, yet 80 percent live less than one mile from a grocery store. See Figure 11 and Figure 12 for a graphical illustration. This means that at least 20 percent of the sample traveled past a grocery store near their house and instead chose to shop at a store further away. The 20 percent opting to travel further accounts even for those who traveled to a grocery store to or from work, because the average distance to the nearest grocery store from the center of University of Illinois campus is 0.6 miles - less than one mile.

In the following discussion, I expand upon the results discussed above through refining the sample of respondents living within one mile and over one mile of a grocery store to those that have revealed satisfaction or dissatisfaction with their current situation. Table 9 presents budget shares of individuals in the "Ideal" and "Constrained" categories. Individuals who live within one mile of a grocery store, but reported a travel distance to their main grocery store greater than five minutes are grouped as individuals who "Opt to Travel" past their nearest grocery store. The five-minute cut-off travel time is conservatively based on a 12 mph average driving time.[11]

As expected, respondents in the "Ideal" category spent 6.0 percent more of their grocery bill on fresh produce and 4.8 percent less on non-fruit & vegetable foods, compared to "Constrained" respondents. Respondents in the "Ideal" group also spent a smaller share of their grocery bill on canned and frozen produce, compared to individuals in the "Constrained" group. "Opt to Travel" respondents reported similar budget share allocation as those in the "Ideal" group, but spent 2.0 percent less on fresh produce than the "Ideal" group.

Table 9: Average grocery expenditure share: Endogenous classification

	Distance to nearest grocery		
	1 mile or less (Ideal)	Over 1 mile (Constrained)	1 mile or less (Opt to Travel)
Fresh F&V	30.4%	24.4%	29.4%
Canned F&V	1.8%	2.6%	2.1%
Frozen F&V	1.8%	2.1%	1.9%
Non F&V	66.0%	70.8%	66.6%
Obs.	221	56	264

[11] The data verifies the validity of this assumption because of those respondents who provided the name of their neighborhood and who reported travel time less than five minutes, 89.5 percent indeed live within one mile of the nearest grocery store.

The "Opt to Travel" respondents tended to be cost-conscious, value-driven shoppers. Of the people who opt to travel, 71.2 percent indicated they had shopped at a supercenter in the month prior to the survey, whereas 48.4 percent of the "Ideal" group had shopped a supercenter in the past month (see Table 4). This difference was not due to locational convenience: the respondents who opt to travel lived approximately one mile further from a supercenter, on average, than do individuals in the "Ideal" group (see Table 3).

The people who opt to travel had similar income, household size, vehicle access, and family structure as those in the "Ideal" group.[12] Yet, those that opt to travel spend less per person per week on groceries ($49 per person per week versus $54 per person per week), and grocery shop less frequently (4.4 times per month versus 5.8 times per month) than individuals in the "Ideal" group. Further, the far travelers held Shelf Life in higher esteem relative to Seasonality compared to those in the "Ideal" group. Table 10 shows how respondents ranked five characteristics of produce based on which have the greatest bearing on their produce purchases. Freshness and Appearance are, again, the first and second most important determinants of produce purchases across all groups, however those opting to travel preferred to purchase produce with a greater shelf life rather than in-season produce. Respondents in the "Constrained" group clearly must prioritize extending the shelf life of their purchases, compared to the respondents who live within one mile of a grocery store. "Constrained" respondents have a surprisingly high farmer's market attendance rate compared with the aggregate of respondents who live over one mile from a grocery store in Table 7. Since farmer's market produce often has significantly longer shelf life than does produce from the grocery store, the "Constrained" respondents may place a high premium on obtaining farmer's market produce to meet their need for long-lasting fresh produce.

Table 10: Produce preferences: endogenous classification

	Distance to nearest grocery		
Rank of Importance	1 mile or less ('Ideal')	Over 1 mile ('Constrained')	1 mile or less (Opt to Travel)
First	Freshness	Freshness	Freshness
Second	Appearance	Appearance	Appearance
Third	Growing Method	Shelf Life	Growing Method
Fourth	Seasonality	Growing Method^	Shelf Life
Fifth	Shelf Life	Seasonality^	Seasonality
Farmers Market/CSA attendance?*	81.9%	76.8%	75.4%

**Convenience* Farmers Market/CSA attendence based on percentage of respondents with neighborhood location
^Growing Method and Seaonality ranked equally

[12] Average income bracket, car access, household size, and number of children for the "opt to travel" versus "Ideal" groups are as follows: $75,000 to 100,000 income bracket per household for both; 89.8% vs. 94.0% of the sample can access a car whenever they need; 2.6 people per household for both; 0.9 vs. 0.8 children per household.

Respondents with ”Ideal” food geography consider growing method and seasonality before shelf life when purchasing produce. In contrast, respondents in the “Constrained” group consider shelf life before growing method or seasonality when purchasing produce. This difference may be due to differing preferences between these two groups, but may be a product of their differing food environments. Respondents in the “Opt to Travel” group have greater preference for value-driven shopping: they buy less fresh produce, hold shelf-life in high regard, shop less often, and stretch their food dollars further within their households compared to those in the “Ideal” group.

Chapter Five: Empirical Analysis and Results

5.1 Introduction

In the following chapter, I discuss results from the demand system estimation and MMNL regressions, conduct simulations combining significant results of both methods, and analyze the outcomes.

5.2 Demand Estimation Results

I estimated the AIDS model separately for the exogenous and endogenous constraint analyses. Each model was estimated for six fruit & vegetable categories (fresh vegetables, fresh fruit, canned vegetables, canned fruit, frozen vegetables, and frozen fruit) and all other foods, resulting in a seven-equation system. All elasticities were calculated at the sample medians. Table 11 through Table 14 report income elasticities, expenditure elasticities, and Marshallian own-price elasticities for respondents living within versus over one mile of a grocery store, and respondents with "constrained" versus "ideal" food geographies.

5.2.1 Exogenous Food Geography Demand Estimation Results

The results of Table 11 show all positive expenditure elasticities, thus each food category comprises a normal good within this demand system. Consistent with economic theory, nearly all own-price elasticities are negative. Negative own-price elasticity implies that an increase in the price of a good diminishes demand for that good. Canned fruit for the "Over 1 mile" subsample provides the only exception. This anomaly is likely due to the significant infrequency of purchases within this food group. Out of 116 respondents in the "Over 1 mile" sample, 74.1 percent did not purchase any canned fruit in the week or month prior to the survey.[13]

Table 11: Estimated expenditure and own-price elasticities for individuals living within 1 mile and over 1 mile of a grocery store

	1 mile or less		Over 1 mile	
Food Group	Food Expenditure	Marshallian own-price	Food Expenditure	Marshallian own-price
(1) Fresh Veg	0.31	-0.41	1.10	-1.02
(2) Fresh Fruit	0.48	-0.76	1.08	-0.98
(3) Canned Veg	1.16	-0.79	1.26	-1.23
(4) Canned Fruit	1.18	-2.12	1.11	4.16
(5) Frozen Veg	0.99	-1.24	1.17	-0.99
(6) Frozen Fruit	1.20	-0.67	1.21	-1.53
(7) Other Foods	1.01	-0.99	0.98	-1.01
Observations	455		116	

[13] Respondents could provide grocery purchase information from either the week or month prior to the survey depending on which time frame was easier for the respondent.

In general, the results presented in Table 11 suggest proximity to a grocery store is related to greater consistency of fresh produce consumption. The food expenditure elasticity estimates show "1 mile or less" respondents had less elastic food expenditure elasticities of 0.31 and 0.48 for fresh vegetables & fruit relative to "Over 1 mile" respondents, where expenditure elasticities were both greater than unity: 1.10 and 1.08, respectively. This result implies that "1 mile or less" individuals treat fresh vegetables & fruit as a staple item more so than "Over 1 mile" respondents; demand for fresh vegetables & fruit of the "1 mile or less" group responds minimally to changes in total food expenditure. Conversely, "Over 1 mile" respondents treat "Other Foods" as a staple item more so than "1 mile or less" respondents. "Over 1 mile" respondents had expenditure elasticity less than unity of 0.98 for "Other Foods" whereas "1 mile or less" respondents had expenditure elasticity above unity of 1.01 for "Other Foods."

Results also suggest proximity to a grocery store is related to less price sensitivity for all food groups, and fresh fruit & vegetables in particular. The magnitudes of own-price elasticities are greater among "Over 1 mile" residents for all food groups, excepting frozen vegetables. Thus, "Over 1 mile" residents change consumption of groceries items in response to own-price changes to a greater degree than their "1 mile or less" counterparts. Household income is not a likely driver of this variation in price sensitivity, as both groups had similar average income levels. Among "1 mile or less" respondents, fresh vegetable purchases are the least sensitive to price changes, which implies proximity to a grocery store is correlated with steady fresh vegetable consumption regardless of fresh vegetable price volatility.

Marshallian cross-price elasticities are reported in Table 12. They represent uncompensated demand responses to price and thus denote substitution effects taking into account the consumer's purchasing power change that results from a price change. Elasticity estimates should be read as the change in quantity demanded of good *x* by row for the change in price of good *y* by column. A positive cross-price elasticity indicates net substitutes, while a negative cross-price elasticity indicates net complements.

Table 12: Estimated Marshallian cross-price elasticities for individuals living within 1 mile and over 1 mile of a grocery store

$e_{x, p(y)}$	1 mile or less							Over 1 mile						
Food Group	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods
(1) Fresh Veg	**-0.50**	0.35	0.31	0.30	0.22	0.42	-0.26	**-1.02**	-0.05	-0.09	-0.04	-0.02	-0.07	0.05
(2) Fresh Fruit	0.02	**-0.85**	0.12	0.13	-0.01	0.18	-0.23	-0.03	**-0.98**	-0.07	-0.02	-0.01	-0.05	0.06
(3) Canned Veg	0.06	0.03	**-0.94**	-0.09	0.08	-0.21	0.08	0.19	0.14	**-1.23**	0.22	-0.02	-0.17	0.23
(4) Canned Fruit	0.06	0.06	-0.12	**-1.23**	0.07	0.09	0.10	-0.11	-0.10	-1.07	**4.16**	-1.07	-2.99	0.17
(5) Frozen Veg	-0.04	-0.10	0.27	0.15	**-0.90**	-0.35	-0.02	0.00	-0.03	0.04	-0.20	**-0.99**	-0.12	0.10
(6) Frozen Fruit	0.11	0.05	-0.13	0.05	-0.04	**-1.08**	0.11	-0.07	-0.05	-0.17	-2.99	-0.12	**-1.53**	0.20
(7) Other Foods	-0.03	-0.02	0.00	0.00	-0.01	0.01	**-0.99**	0.05	0.06	0.23	0.17	0.10	0.20	**-1.01**
Observations	455							116						

The results of Table 12 provide evidence that individuals who live over one mile from a grocery store are more likely to substitute frozen or canned vegetables for fresh vegetables compared to individuals who live within one mile of a grocery store. “Over 1 mile” residents increase consumption of canned vegetables by 0.19 percent for a 1.0 percent increase in the price of fresh vegetables. “1 mile or less” residents substitute non-fresh for fresh vegetables to a lesser degree of 0.06 percent. Also, “1 mile or less” residents treat fresh vegetables and frozen vegetables as complementary goods as indicated by the negative cross-price term of -0.04, while “Over 1 mile” residents exhibit behavior slightly closer to substitution with a positive cross-price elasticity of 0.004 (taken to the thousandth decimal place) for fresh to frozen vegetables. This behavior switches for the fruit categories. “Over 1 mile” residents treat fresh fruit as a complementary good to canned or frozen fruit with cross price elasticities of -0.10 and -0.05, respectively, whereas “1 mile or less” respondents are more likely to substitute fresh for frozen or canned fruits because the cross price terms are positive. For example, for a 1.0 percent price increase of fresh fruit canned fruit demand among “1 mile or less” residents increases by 0.06 percent and frozen fruit increases by 0.05 percent while canned fruit and frozen fruit demand both decrease with a price increase in fresh fruit for the “Over 1 mile” residents. For both groups the relationships between fresh and non-fresh fruits are inelastic, which is unsurprising if we consider that fresh fruit and non-fresh fruits are generally less substitutable for consumption and cooking than are fresh and non-fresh vegetables.

Using this exogenous classification of food access, the above results suggest residents who live within a mile of a grocery store have stable fresh produce consumption and are less price sensitive. People who live over a mile from a grocery store are more likely to substitute frozen or canned vegetables for fresh vegetables, and their consumption of other foods exhibits lower variation than the “1 mile or less” residents. Next, I use an endogenous classification of food access to further test the effect of geography on consumption behavior. Table 13 shows expenditure and own-price elasticities for individuals living with an “Ideal” food environment compared to respondents in the “Constrained” group. The last two columns also show expenditure and own-price elasticities for respondents who “Opt to Travel.”

5.2.2 Endogenous Food Geography Demand Estimation Results

Table 13 shows expenditure and own-price elasticities for individuals living with an "Ideal" food environment compared to respondents in the "Constrained" group. The last two columns also show expenditure and own-price elasticities for respondents who "Opt to Travel." The results of Table 13 show all positive expenditure elasticities, thus each food category comprises a normal good within this system. As with the previous analysis, nearly all own-price elasticities comply with economic theory and are negative. Frozen fruit for the "Constrained" and "Opt to Travel" subsamples provide the only exception, which is likely due to the significant infrequency of purchases within this food group. Of these groups, 84.6 of those who "Opt to Travel" and 83.9 percent of those in the "Constrained" group did not purchase any frozen fruit in the week or month prior to the survey. Similarly, the "Constrained" group did not purchase canned fruit 83.9 percent of the time, which may explain the extremely large canned fruit own-price elasticity of -28.07.

Table 13: Estimated expenditure and own-price elasticities for Ideal, Constrained, and Opt to Travel respondent groups

	Ideal		Constrained		Opt to Travel	
Food Group	Food Expenditure	Marshallian own-price	Food Expenditure	Marshallian own-price	Food Expenditure	Marshallian own-price
(1) Fresh Veg	0.33	-0.35	2.40	-2.68	0.73	-0.90
(2) Fresh Fruit	0.70	-0.88	0.94	-0.85	0.69	-1.01
(3) Canned Veg	1.15	-0.90	1.09	-1.44	1.05	-0.83
(4) Canned Fruit	1.16	-1.04	1.43	-28.07	1.01	-2.66
(5) Frozen Veg	1.15	-1.01	1.84	-2.85	0.83	-1.05
(6) Frozen Fruit	1.20	-1.01	1.15	0.77	0.98	0.68
(7) Other Foods	1.02	-0.99	1.03	-1.00	1.03	-0.99
Observations	221		56		260	

Based on these results, it again appears proximity to a grocery store is related to greater consistency of fresh produce consumption. Both the "Ideal" and "Opt to Travel" respondents have inelastic expenditure elasticity of demand for fresh produce, compared to those in the "Constrained" group. The "Ideal" and "Constrained" groups exhibit demand for canned and frozen fruits & vegetables above unity elastic whereas the "Opt to Travel" group treats frozen fruit & vegetables as relatively more of a staple item. The "Opt to Travel" group may be more likely to frequent big box stores in an effort to garner deals on non-perishable frozen produce. It is unclear whether the endogenous food access classification renders significantly different consumption behavior of non-fresh food items. "Constrained" respondents have lower income elasticity of demand for canned vegetables and frozen fruit, but higher income elasticity of demand for canned fruit and frozen vegetables compared to the "Ideal" group. Expenditure elasticity for "Other Foods" is similar for all groups. Respondents who live within one mile of a grocery store and prefer their status quo have stable demand for fresh produce, but demand for all other food groups is elastic compared to the "Opt to Travel" group.

Results are also inconclusive with respect to own-price sensitivity. “Ideal” scenario respondents are significantly less price sensitive to fresh vegetables than “Constrained” respondents, but slightly more price sensitive to fresh fruit. Both sets of respondents who live within one mile of a grocery store are less price-sensitive to all non-fresh foods compared to the “Constrained” group, though demand for “Other Foods” is robustly unit elastic for all groups. While the positive sign on own-price elasticity of frozen fruit may be an error caused by non-response bias, it is also plausible that consumers in the “Constrained” and “Opt to Travel” groups purchase frozen fruit as a complement to fresh fruit and thus, do not budget frozen fruit separately from fresh or other food categories. Respondents who “Opt to Travel” are less price-sensitive to fresh vegetables, since the own-price elasticity is -0.90 compared to the more elastic -2.68 of the “Constrained” group who must travel over one mile to reach their grocery store. The opposite is true, however, for fresh fruit. The “Constrained” group is less price-sensitive to fresh fruit with an own-price elasticity of -0.85 whereas the “Opt to Travel” group exhibits fresh fruit own-price elasticity of -1.01. Thus, respondents who elect to travel are less price- sensitive to all food groups except fresh fruit, compared to those who are forced to travel over one mile.

Marshallian cross-price elasticities for the endogenous “food desert” analysis are reported in Table 14. Table 15 provides cross-price elasticities for respondents who “Opt to Travel.” Recall, a positive cross-price elasticity indicates net substitutes, while a negative cross-price elasticity indicates net complements.

Table 14: Estimated Marshallian cross-price elasticities for Ideal and Constrained respondent groups

$e_{x,\ p(y)}$	Ideal							Constrained						
Food Group	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods
(1) Fresh Veg	**-0.35**	0.31	0.36	0.33	0.28	0.49	-0.29	**-2.68**	-1.82	-2.49	-2.04	-1.69	-2.24	-3.08
(2) Fresh Fruit	-0.12	**-0.88**	0.05	0.01	0.05	0.11	-0.14	-0.02	**-0.85**	0.07	0.00	-0.08	0.03	0.04
(3) Canned Veg	0.09	0.03	**-0.90**	0.13	-0.10	-0.32	0.08	0.14	0.11	**-1.44**	-0.47	0.64	0.22	-0.06
(4) Canned Fruit	0.07	-0.05	0.37	**-1.04**	-0.02	-0.42	0.12	0.71	0.80	6.93	**-28.07**	4.15	15.06	-0.77
(5) Frozen Veg	0.00	0.03	-0.10	-0.01	**-1.01**	0.00	0.10	1.76	1.76	-5.58	2.37	**-2.85**	1.78	-1.38
(6) Frozen Fruit	0.13	0.04	-0.15	-0.06	0.01	**-1.01**	0.12	0.13	0.06	0.34	-1.66	-0.33	**0.77**	-0.11
(7) Other Foods	-0.04	-0.02	-0.01	-0.01	-0.01	0.00	**-0.99**	0.05	0.03	0.02	0.03	0.03	0.02	**-1.00**
Observations	221							56						

Table 15: Estimated Marshallian cross-price elasticities for Opt to Travel group

$e_{x, p(y)}$	Opt to Travel						
Food Group	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods
(1) Fresh Veg	**-0.90**	0.30	0.37	0.33	0.29	0.37	0.58
(2) Fresh Fruit	0.09	**-1.01**	0.27	0.23	0.16	0.24	0.58
(3) Canned Veg	0.15	0.13	**-0.83**	-0.03	0.12	-0.42	-0.11
(4) Canned Fruit	0.00	0.04	-0.11	**-2.66**	1.10	0.67	-0.04
(5) Frozen Veg	-0.27	-0.24	0.32	0.82	**-1.05**	-0.55	0.13
(6) Frozen Fruit	0.71	0.28	-2.30	1.01	-1.20	**0.68**	-0.19
(7) Other Foods	0.06	0.05	0.04	0.04	0.04	0.04	**-0.99**
Observations	260						

The results of Table 14 provide evidence that constrained food geography may intensify the propensity to substitute non-fresh for fresh produce. Respondents in the "Constrained" category exhibit a higher elasticity of substitution for both fresh-to-canned vegetables of 0.14 and fresh-to-frozen vegetables of 1.76 compared to "Ideal" category respondents. Similarly, the "Constrained" respondents have greater elasticity of substitution for fresh-to-canned fruit of 0.8 and fresh-to-frozen fruit of 0.6 compared to respondents in the "Ideal" category. Both sets of respondents treat "Other Foods" as complementary to fresh vegetables, as the cross-price elasticities are both negative. However, the "Constrained" group is more likely to substitute fresh fruit for "Other Foods" with a positive cross-price elasticity of 0.04 whereas the "Ideal" group treats "Other Foods" and fresh fruit as complementary goods with a negative cross-price elasticity of -0.14. This result is consistent with the notion hat the "Constrained" group may be more sensitive to perishability.

Comparing behavior between those who are forced to travel (the "Constrained" group) and those who elect to travel (the "Opt to Travel" group) in Table 15 provides inconclusive results. The "Constrained" group more readily substitutes frozen for fresh vegetables with a cross price elasticity of 1.76 compared to -0.27 for the "Opt to Travel" group. Yet, both groups substitute fresh for canned vegetables to the same degree, with an elasticity of 0.14 and 0.15, respectively. Similarly, the "Constrained" group substitutes fresh for canned fruit to a greater degree than those who "Opt to Travel" but substitutes fresh for frozen fruit to a lesser degree. It should be noted that the "Constrained" cross price elasticity of frozen-to-fresh vegetables of 1.76 indicates highly elastic substitution above unity. Thus, while the differences between fresh and non-fresh produce substitution are minimal when comparing those forced to travel and those who willingly travel, fresh to frozen vegetable substitution appears to be the most susceptible to change due to geographic constraint.

The highly elastic cross-price elasticity terms on fresh vegetables for the "Constrained" group indicates these respondents are likely to compromise their purchase of fresh vegetables more than any other food category if they experience a budget squeeze. They essentially treat fresh vegetables as a residual category. If the price increases on any other food category, they compensate by purchasing fewer fresh vegetables. Some of this result may be

driven by aversion to vegetables. The highly elastic income elasticity term of frozen vegetables in Table 13 would support this plausible hypothesis. However, the group has signaled that they are dissatisfied with some aspect of their grocery shopping experience, which may be relate to their produce quality perishability. A typical basket of fresh vegetables available in central Illinois grocery stores tends to perish faster than a typical basket of fresh fruit. Lettuce, peppers, and cucumbers, for example, have significantly shorter shelf life than do apples, bananas, and oranges. Recall the "Constrained" group regarded shelf life as an important aspect of produce quality (see Table 10). The "Constrained" group's highly elastic vegetable consumption may be a product of poor quality or highly perishable vegetables.

Respondents who "Opt to Travel" substitute fresh for canned vegetables to a greater degree than the "Ideal" respondent group with a cross-price elasticity of 0.15 compared to 0.09 for the "Ideal group", and treat fresh and frozen vegetables as complements with a negative cross-price elasticity of -0.27, unlike the "Ideal" group, who treat the two food categories independently. The "Opt to Travel" respondents substitute canned and frozen fruit for fresh fruit to a greater degree than does the "Ideal" group. For example, "Opt to Travel" respondents increase consumption of canned fruit by 0.04 percent for a one percent increase in fresh fruit, and increase consumption of frozen fruit by 0.28 percent for a one percent increase in fresh fruit. Whereas respondents in the "Ideal" group actually decrease canned fruit consumption by 0.05 percent and increase frozen fruit to a lesser degree of 0.04 percent for a one percent increase in fresh fruit. If we consider the effect of a price decrease in "Other Foods," people who "Opt to Travel" will decrease their demand of fresh fruits & vegetables by 0.58 percent, whereas respondents in the "Ideal" group will respond by increasing demand for fresh vegetables by 0.29 percent and fresh fruit by 0.14 percent.

Respondents who live near to a grocery store exhibit greater stability of fresh produce consumption, compared to those who live over one mile. Proximate respondents also are less sensitive to price volatility of fresh vegetables, as well as non-fresh fruit & vegetable categories. Constrained food geography may intensify the propensity to substitute non-fresh for fresh produce. Such is the case when comparing the significantly different respondent groups of "Constrained" versus "Ideal" respondents. Proximity to a grocery store may have a positive effect on consumption behavior, even if an individual does not patronize the store nearby their residence. The "Constrained" respondents substitute frozen for fresh vegetables to a significantly greater degree than do respondents who also travel far to get to their grocery store, but do so willingly, e.g. the "Opt to Travel" respondents.

The exogenous and endogenous definitions reveal proximity to a grocery store is positively related to stable consumption of fresh produce, and lower variation of fresh produce consumption in response to price changes. Further, these results provide some evidence that proximity to a grocery store decreases the likelihood of

substituting fresh for non-fresh vegetables, even if people do not shop at the store nearest their house. The following discussion explores whether the observed consumption behavior may be caused by food geography or consumer preferences.

5.3 Choice Experiment Results

The choice experiment analyses quantify consumer preferences for grocery store attributes. I compare these results with the previous demand system results to test whether "Over 1 mile" or "Constrained" respondents are constrained by their food geography. Following the previous methodology, I use both an endogenous and an exogenous classification of a "food desert" to test for constraints.

I first estimate the MMNL model with the full sample prior to splitting the sample into groups. Next, I estimate three sets of MMNL models to compare with the demand systems. The first set includes the entire sample of respondents and compares preferences between "Over 1 mile" and "1 mile or less" respondents following the exogenous "food desert" classification. The second set includes the subset of "Over 1 mile" and "1 mile or less" respondents who were less likely than average to choose their status quo shopping scenario – the "Constrained" respondents, and more likely than average to choose their status quo shopping scenario – the "Ideal" respondents, respectively. The third set includes the "Constrained" and "Opt to Travel" respondents. For all three cases, I report a main effects specification as well as interaction effects specifications.

5.3.1 Exogenous Food Geography Choice Experiment Results

The "main effects" refer to the four grocery store attributes in my choice experiment (cost of a week's worth of groceries, travel time required to get to the store, store type, and the produce selection). The coefficient on each of these main effects provides an estimate of the change in utility given a change in the associated attribute. *Weekly Grocery Bill* and *Travel Time* are numeric attributes, thus their coefficients can be interpreted on a continuum as the change in utility given a $1 change or 1-minute change in the attribute, respectively. On the other hand, store type and produce selection are ordinal, discrete attributes. Consequently, I disaggregated the store type attribute into *Grocery Store* and *Specialty Store* variables. Coefficients on *Grocery Store* and *Specialty Store* provide the change in utility from shopping at a grocery store or specialty store instead of the base case of a supercenter. Similarly, I disaggregated the produce selection attribute into *HQ Produce* and *AQ Produce* variables. The coefficient on these variables indicates the probability of choosing to shop at a store with high quality or average quality produce instead of the base case low quality produce. Recall, utility cannot be observed directly, but we infer the choice that garners the highest utility based on which choice the respondent selects. To compare grocery store preferences between "Over 1 mile" and "1 mile or less" residents, I multiply each main effect by a dummy variable, denoted by *FD*, equal to 1 if the respondent lives over one mile from a grocery store, and equal to 0 if the respondent lives less than one mile from a grocery store.

Results of the main effects MMNL regressions are presented in Table 16. The first two columns show regression results of the main effects MMNL for the full sample of respondents. The second two columns show regression results for the main effects MMNL incorporating the *FD* dummy variable. The columns marked "SD Pval" provide

the p-value associated with the standard deviation of each parameter. Each of the main effect parameters has significant standard deviation around their mean, which implies individual heterogeneity is significant for all attributes. This result justifies employing the MMNL model as opposed to the CL model because the MMNL model yields more conservative estimates of marginal WTP compared to the CL when individual heterogeneity is significant. All main effect variables are significant at one percent, excepting the *Specialty Store* variable. The *ASC* term captures attributes affiliated with the status quo choice that I did not explicitly account for in the experimental design. *ASC* is negative and significant meaning respondents generally preferred to choose their current store rather than the new hypothetical store options A or B. This result follows Tang et al. (2001) who found knowledge of store layout was a significant source of utility of grocery shopping.

The negative sign on *Weekly Grocery Bill* and *Travel Time* meet my a priori expectations that people will have disutility from spending more money on groceries or taking extra time to travel to a grocery store, ceteris paribus. *Specialty Store* and *Grocery Store* are both negative, which means respondents prefer to shop at a supercenter to either of these store types. This result can be explained by the fact that one of the most popular grocery stores in Champaign-Urbana is indeed a supercenter model, a regional retailer called Meijer.[14] Also, it should be noted that my choice experiment described "supercenter" as a retailer characterized by their size, the number of non-food items available, and the number of extra services such as pharmacy or florist that are available. I did not provide any inferences on product quality or ownership structure in my definition of store type. The respondent's preference for supercenters mimics the findings of Tang et al. (2001) who found that a larger product assortment in a store including breadth and depth resulted in a greater likelihood that a consumer would choose that store because large assortment facilitates "one-stop shopping" and obviates the need to make separate trips to other stores. In other words, it lessens the fixed costs associated with grocery shopping. *Specialty Store* is insignificant across the models.

HQ Produce and *AQ Produce* are positive and significant, which indicates as I expect, that people garner higher utility from high quality or average quality produce compared to low quality produce. Respondents perceive the difference between low and high quality to be significantly greater than the difference between low and average quality, so they are willing to pay significantly more to consumer high quality as opposed to average quality produce.

The main effect results are robust between the MMNL model without the *FD* dummy and the MMNL with the *FD* dummy. If the main effect variables interacted with the *FD* dummy are significant, "Over 1 mile" respondents have significantly different utility for that main effect. The change in utility for a change in *Weekly Grocery Bill*, *Grocery*

[14] Of all grocery stores in Champaign County, the highest share of respondents chose Meijer as their primary grocery store (30.9%), followed by Schnucks (30.6%) and County market (13.9%).

Store, *HQ Produce* and *AQ Produce* are significantly different for "Over 1 mile" respondents compared to respondents who live within one mile of a grocery store. "Over 1 mile" respondents experience more disutility from an increase in the grocery bill. For example, a $1 increase in the grocery bill decreases the probability of choosing that store by a factor of -0.021 – 0.026 = -0.047. The greater disutility from a higher grocery bill is consistent with my findings from the demand estimation in Section 5.2 where I found "Over 1 mile" residents to be more price-sensitive. Price-sensitivity for the "Over 1 mile" group is not likely attributable to income constraints, since on average, they have higher income than do the "Less than 1 mile" group (see Table 4). The addition of the "Food Desert" dummies improved the model fit since the AIC and BIC values dropped significantly in the "MMNL w/ Dummy" model.[15]

Table 16: MMNL results with Main Effect and "Food Desert" dummies

	MMNL w/o dummy		MMNL w/ dummy	
Variable	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]
Main Effects				
ASC	-1.413 ***		-1.315 ***	
	0.054		*0.073*	
Weekly Grocery Bill	-0.024 ***		-0.021 ***	
	0.003		*0.003*	
Travel Time	-0.109 ***	0.000	-0.110 ***	0.000
	0.007		*0.010*	
Specialty Store	-0.114	0.000	-0.171	0.000
	0.095		*0.128*	
Grocery Store	-0.615 ***	0.000	-0.813 ***	0.000
	0.081		*0.117*	
HQ Produce	2.718 ***	0.000	2.651 ***	0.000
	0.112		*0.153*	
AQ Produce	0.649 ***	0.004	0.558 ***	0.047
	0.076		*0.107*	
Main Effect 'Food Desert' Dummies				
ASC_FD			-0.470 *	0.000
			0.278	
Weekly Grocery Bill_FD			-0.026 ***	0.105
			0.009	
Travel Time_FD			-0.014	0.016
			0.022	
Specialty Store_FD			0.423	0.004
			0.329	
Grocery Store_FD			0.601 **	0.349
			0.266	
HQ Produce_FD			1.520 ***	0.000
			0.395	
AQ Produce_FD			0.469 *	0.426
			0.273	
Observations	15,534		10,098	
Log Liklihood	-3878.05		-2490.24	
Prob > Chi2	0.000		0.000	
AIC	7770.10		5008.48	
BIC	7823.65		5109.56	

*** significant at 1%; ** significant at 5%, * significant at 10%

[a]Significance of standard deviations of individual heterogeneity

Coefficient estimates for remaining attributes are better interpreted through a marginal WTP measure. Marginal WTP values for the main effects and "Food Desert" dummy MMNL models are shown in Table 17**.** Unlike the

[15] AIC (Akaike's Information Criterion) and BIC (Bayesian Information Criterion) are common statistical measures used in probability-based regressions to assess the fit of a model based on measures of information. When comparing two models, the smaller the value of the AIC or BIC statistic, the better the fit of the model. BIC identifies the model that is more likely to have generated the observed data. The AIC offers a relative estimate of improved fit versus deviance. (Williams, Richard. Class notes. Categorical Data Analysis. Fall 2011. University of Notre Dame. < http://www3.nd.edu/~rwilliam/stats3/>

coefficient estimates presented in Table 16, the marginal WTP measures depend on respondent's disutility from the weekly grocery bill. Thus, the marginal WTP measure is related to, but should be interpreted differently from a utility measure.

Table 17: Marginal Willingness to Pay - Main Effects

	Full Sample	**1 mile or less**	**Over 1 mile**	Significant Difference?*
Travel Time ($/min. travel)	-$4.45	-$5.22	-$2.67	No
Specialty Store ($/specialty vs supercenter)	-$4.68	-$8.08	$5.41	No
Grocery Store ($/grocery vs supercenter)	-$25.21	-$38.48	-$4.53	Yes
HQ Produce ($/HQ vs LQ produce)	$111.36	$125.54	$89.34	Yes
AQ Produce ($/AQ vs LQ produce)	$26.59	$26.41	$21.98	Yes

*Significant difference between sample groups of utility for the attribute, not marginal WTP for the attribute

The marginal WTP for added travel time is smaller for the "Over 1 mile" respondents than for the "1 mile or less" respondents. Individuals living over one mile from a grocery need to be compensated $2.67 to travel an extra minute as opposed to "1 mile or less" individuals who need to be compensated $5.22. This result supports an endogenous sorting explanation whereby people who choose to live closer to a grocery store may do so because they value short trips to the grocery store, whereas individuals who live further away find other locational amenities more desirable than proximity to a grocery store.

Shopping at a specialty store instead of a supercenter does not significantly change utility for any of the respondent groups. "Over 1 mile" respondents, however, garner higher utility from shopping at a grocery-type store than do the "1 mile or less" respondents. "1 mile or less" respondents require more compensation, $38.48, to shop at a grocery store instead of a supercenter compared to "Over 1 mile" residents who require a smaller compensation of $4.53. Though their marginal WTP for a grocery store is negative, it is significantly less negative than that of the "1 mile or less" group, meaning they are significantly less averse to shopping at a grocery store over a supercenter than the "1 mile or less" group. This result is evident despite the fact that the "Over 1 mile" respondents were more likely than the "1 mile or less" respondents to have shopped at a supercenter in the month prior to the survey. Likely, the "Over 1 mile" group shopped at supercenters more frequently than the "1 mile or less" group because they are more price-sensitive. Note the *Weekly Grocery Bill_FD* term in Table 16 is negative and significant, meaning the "Over 1 mile" group has more disutility from paying a larger grocery bill than does the "1 mile or less" group, all else equal. This result is consistent with their highly elastic expenditure- and price- elasticity terms for all food categories, excepting "Other Foods" noted in the Table 11**.** Some of this price

sensitivity may be inherent. For example, the "Over 1 mile" group has a slightly lower average annual income than does the "1 mile or less" group (see Table 3). But some of the price sensitivity may be a product of possible quality or availability constraints from their food environment and the supercenters at which they often shop, in particular. For example, they may be dissatisfied with the produce quality at their current stores since high quality is valuable to both "Over 1 mile" and "1 mile or less" respondents, however the "Over 1 mile" group are not considerably less satisfied with their current situation than the "1 mile or less group" based on their average status quo selection of 3.2 versus 3.4 times out of six (see Table 3).

The most striking result from Table 17 is the significant difference in WTP for high quality produce over average quality produce across all sample groups. Marginal WTP for high quality produce is extremely high. This result is driven by the large, positive coefficient on *HQ Produce* in each MMNL model. Respondents to my survey are evidently insensitive to significant price differentials that were associated with the produce attribute, and consistently chose the grocery store option with the best produce availability. "1 mile or less" respondents exhibited higher marginal WTP for both average and high quality produce compared to the "Over 1 mile" respondents. Yet, "Over 1 mile" respondents garner significantly greater utility from higher produce quality than do the "1 mile or less" respondents`, since the *HQ Produce_FD* and *AQ Produce_FD* parameter estimates are both positive and significant. Preferences of the "Over 1 mile" group, therefore, appear at odds with the results of the demand system in Section 5.2 in which the "Over 1 mile" respondents displayed less consistent fruit & vegetable consumption. However, some of the preferential differences may be overstated unless I control for respondent income levels, average travel time to their main grocery store, or their proximity to the nearest grocery store. The main effects regressions are not able to control for respondent characteristics explicitly. To control for such characteristics, I introduce interaction terms into the MMNL model.

Table 18 provides the results of five regressions using interaction effects for respondents' proximity to a grocery store (*Distance*), travel time to their principal grocery store (*CurrentTravel*), and income level (*Income*). I interact *Distance* with *Weekly Grocery Bill* as well as *CurrentTravel* with *Weekly Grocery Bill* to control for the effect of proximity and travel time on respondents' willingness to pay for other attributes. I interact *Income* with both *HQ Produce* and *AQ Produce* to control for any effect of income on produce quality. Thus, each model in Table 18 has at least four interaction terms in addition to main effects terms. The first two columns show regression results from the MMNL without a "food desert" *FD* dummy variable. The four specifications titled "MMNL w/ dummy" from 1 to 4 each include the *FD* dummy on one of the four interaction terms. When interacted with the dummy these four variables are triple interaction terms. The models also include the main effects "food desert" dummy interactions from the previous table. Due to computing limitations, maximum likelihood estimation was not able to converge for a model with all main effect, interaction effect, and dummy triple interaction variables. Consequently, I test each triple interaction term for significance one at a time.

The MMNL models with interactions have similar coefficient results as that of the main effects model in sign and magnitude with the exception of the *AQ Produce* and *Weekly Grocery Bill* variables, which are insignificant. Across all models, the significance level, magnitude, and sign of all variables are relatively robust.

Considering the first MMNL model without the "food desert" dummy, *Distance x Grocery Bill* and *CurrentTravel x Grocery Bill* are negative and significant. This implies that respondents who live further from a grocery store, or have a longer travel time will have a lower probability of choosing a store with a high grocery bill. The added distance or travel time, itself, adds a cost premium to the grocery bill, so it is not surprising that respondents faced with this additional "cost" to grocery shop would be averse to expensive groceries. Even though the *Weekly Grocery Bill_FD* interaction terms are insignificant, the additional demographic interaction terms continue to signify that the "Over 1 mile" respondent group is more price-sensitive than are those who live close to a grocery store.

The interactions of *Income x HQ Produce* as well as *Income x AQ Produce* are both positive and significant, which indicates higher income is correlated with greater likelihood of choosing a store with better quality produce. This finding reflects that of Wilde, McNamara, and Ranney (1999 and 2000) who found low-income households were more likely to consume non-fruit & vegetable products, even if fruits & vegetables were subsidized.

Among, the four models with triple interaction terms, the double interaction effects without the dummy are robustly significant across all models. However, each triple interaction term is insignificant, suggesting the aforementioned interaction effects of distance and travel time on marginal WTP, and income on produce quality are not significantly different for "Over 1 mile" compared to "1 mile or less" respondents. Consequently, I drop the triple interaction terms for simplicity to calculate the marginal WTP for grocery store attributes. Model fit improved without the triple interaction terms, since the AIC and BIC measures are consistently lowest in the final "MMNL w/ dummy w/o triple effects" model. The results of the interaction effects MMNL without triple interaction terms are presented in the last columns of Table 18**.**

Table 18: MMNL Results with Interaction Effects and "Food Desert" dummies

	MMNL w/o dummy		MMNL w/ dummy 1		MMNL w/ dummy 2		MMNL w/ dummy 3		MMNL w/ dummy 4		MMNL w/ dummy w/o triple effects	
Variable	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]
Main Effects												
ASC	-1.383 ***		-1.304 ***		-1.371 ***		-1.331 ***		-1.361 ***		-1.357 ***	
	0.073		*0.078*		*0.084*		*0.079*		*0.082*		*0.082*	
Weekly Grocery Bill	-0.008		-0.012		-0.003		-0.006		-0.001		-0.004	
	0.006		*0.010*		*0.007*		*0.006*		*0.007*		*0.007*	
Travel Time	-0.115 ***	0.000	-0.116 ***	0.000	-0.124 ***	0.000	-0.117 ***	0.000	-0.125 ***	0.000	-0.124 ***	0.000
	0.009		*0.011*		*0.012*		*0.011*		*0.012*		*0.012*	
Specialty Store	-0.203	0.000	0.248 *	0.000	-0.264 *	0.000	-0.261 **	0.000	-0.237	0.000	-0.251 *	0.000
	0.129		*0.141*		*0.152*		*0.148*		*0.150*		*0.152*	
Grocery Store	-0.749 ***	0.000	-0.764 ***	0.000	-0.902 ***	0.000	-0.869 ***	0.000	-0.917 ***	0.000	-0.911 ***	0.000
	0.105		*0.116*		*0.131*		*0.115*		*0.128*		*0.128*	
HQ Produce	1.422 ***	0.000	1.625 ***	0.000	1.491 ***	0.000	1.937 ***	0.000	1.691 ***	0.000	1.707 ***	0.000
	0.423		*0.422*		*0.424*		*0.452*		*0.430*		*0.427*	
AQ Produce	-0.249	0.015	-0.075	0.029	-0.164	0.515	-0.112	0.083	-0.166	0.984	-0.168	0.508
	0.332		*0.331*		*0.356*		*0.330*		*-0.166*		*0.344*	
Main Effect 'Food Desert' Dummies												
ASC_FD			-0.831 ***	0.000	-0.952 ***	0.000	-0.737 ***	0.000	-0.599 *	0.000	-0.764 **	0.000
			0.308		*0.317*		*0.281*		*0.323*		*0.324*	
Weekly Grocery Bill_FD			0.000	0.008	-0.032	0.246	-0.006	0.233	-0.015	0.007	0.002	0.593
			0.018		*0.023*		*0.014*		*0.018*		*0.014*	
Travel Time_FD			-0.035	0.000	-0.029	0.007	-0.023	0.185	-0.028	0.001	-0.011	0.166
			0.023		*0.029*		*0.027*		*0.027*		*0.025*	
Specialty Store_FD			0.440	0.010	0.783 **	0.016	0.664 *	0.008	0.510	0.000	0.449	0.001
			0.342		*0.394*		*0.351*		*0.430*		*0.387*	
Grocery Store_FD			0.677 **	0.381	0.600 **	0.899	0.571 **	0.004	0.665 **	0.366	0.585 *	0.164
			0.296		*0.303*		*0.270*		*0.328*		*0.300*	
HQ Produce_FD			1.713 ***	0.000	1.550 ***	0.338	0.921	0.736	2.450 ***	0.000	1.492 ***	0.003
			0.444		*0.451*		*1.250*		*0.560*		*0.400*	
AQ Produce_FD			0.379	0.567	0.428	0.054	0.447	0.560	0.693	0.144	0.279	0.368
			0.287		*0.317*		*0.300*		*1.254*		*0.298*	
Interaction Effects												
Distance x Grocery Bill	-0.004 **	0.186	0.012	0.518	-0.010 **	0.365	-0.006 **	0.620	-0.009 **	0.805	-0.010 ***	0.813
	0.002		*0.016*		*0.004*		*0.003*		*0.004*		*0.003*	
CurrentTravel x Grocery Bill	-0.001 **	0.000	-0.002 **	0.002	-0.002 **	0.000	-0.002 ***	0.000	-0.002 ***	0.000	-0.002 **	0.000
	0.001		*0.001*		*0.001*		*0.001*		*0.001*		*0.001*	
Income x HQ Produce	0.300 ***	0.259	0.224 ***	0.810	0.307 ***	0.000	0.189 **	0.020	0.252 ***	0.000	0.248 ***	0.000
	0.078		*0.078*		*0.084*		*0.082*		*0.084*		*0.083*	
Income x AQ Produce	0.177 ***	0.031	0.125 **	0.606	0.157 **	0.058	0.137 **	0.811	0.156 **	0.336	0.157 **	0.297
	0.062		*0.062*		*0.068*		*0.062*		*0.068*		*0.065*	
Interaction effects 'Food Deser't Dummies												
Distance x Grocery Bill_FD			-0.021	0.176								
			0.017									
CurrentTravel x Grocery Bill_FD					0.002	0.278						
					0.002							
Income x HQ Produce_FD							0.083	0.941				
							0.205					
Income x AQ Produce_FD									-0.048	0.003		
									0.213			
Observations	9,609		9,609		9,609		9,609		9,609		9,609	
Log Liklihood	-2382.71		-2357.83		-2344.64		-2345.84		-2342.51		-2344.29	
Prob > Chi2	0.000		0.000		0.000		0.000		0.000		0.000	
LR Chi 2 (8)	474.48		511.38		534.28		537.21		542.52		540.69	
AIC	4787.43		4753.66		4727.29		4729.67		4723.02		4724.59	
BIC	4866.30		4889.90		4863.52		4865.91		4859.26		4853.66	

*** significant at 1%; ** significant at 5%, * significant at 10%

[a]Significance of standard deviations of individual heterogeneity

Table 19 provides the marginal WTP for each grocery store attribute based on the MMNL interaction effect specification in the final columns of Table 18. Unit values for the respondent characteristic variables of travel time, distance, and income are based on the sample average because I am interested in the marginal WTP for the average person in each group: full sample, "Over 1 mile" respondents, and "1 mile or less" respondents. In other words, I calculate the marginal WTP per grocery shopping trip for people in these groups.

Table 19: Marginal Willingness to Pay - Interaction Effects

	Full Sample	**1 mile or less**	**Over 1 mile**	Significant Difference?*
Travel Time ($/min. travel)	-$4.41	-$5.29	-$2.07	No
Specialty Store ($/specialty vs supercenter)	-$7.82	-$10.71	$3.03	No
Grocery Store ($/grocery vs supercenter)	-$28.78	-$38.79	-$4.97	Yes
HQ Produce ($/HQ vs LQ produce)	$117.06	$129.33	$71.02	Yes
AQ Produce ($/AQ vs LQ produce)	$27.28	$28.70	$15.73	No
Unit Values				
Average Travel Time (min.)	9.4	8.6	12.3	
Average Distance (mi.)	1.2	0.5	4.3	
Average Income Bracket^	5.4	5.4	5.9	

*Significant difference between sample groups of utility for the attribute, not marginal WTP for the attribute

^Income bracket of 4 denotes annual income between $50-75K; 5 denotes $75-100K; 6 denotes $100-125K

Marginal WTP results are similar to that of the main effects model. "Over 1 mile" respondents require less compensation for an added minute of travel time compared to "1 mile or less" respondents. "1 mile or less" respondents require more compensation to shop at a grocery store instead of a supercenter compared to "Over 1 mile" respondents. Similar to the main effects results, all sample groups are willing to pay significantly more for high quality produce over average quality produce. One notable difference between these results and that of the main effects marginal WTP values is that of the "Over 1 mile" respondents' WTP for high and average quality produce: marginal WTP for high- and average- quality produce over low quality produce decreased. The main effects marginal WTP for high- over low- quality produce was $89.34 compared to the interaction effects model of $71.02. Similarly, "Over 1 mile" residents exhibited a marginal WTP of $21.98 for average- over low- quality produce in the main effects model. After including the interaction effects, marginal WTP falls to $15.73. Thus the added travel time and distance to a grocery store diminish peoples' perceived value of better quality produce. Because average income for the two groups is very similar, (5.4 for "1 mile or less" respondents, and 5.9 for "Over 1 mile" respondents), income is not readily causing this variation in perceived produce quality value.

After controlling for proximity, travel time, and income, preferences of the "Over 1 mile" residents appear, again, to exhibit preferences somewhat inconsistent with their consumption behavior, but the interaction effects have

muted some of these disparities evident in the main effects MMNL. First, the "Over 1 mile" respondents have a higher WTP for grocery stores over supercenters than their "1 mile or less" counterparts, despite the fact that they shop more frequently at supercenters than the "1 mile or less" respondents. Neither travel time nor income appear to constrain the "Over 1 mile" group, however, their relatively inconsistent fruit & vegetable consumption and propensity to substitute frozen or canned vegetables for fresh vegetables, are somewhat anomalous with their preferences for high quality produce. The "Over 1 mile" group garners higher utility from high quality produce than does the "1 mile or less group" since the "*HQ Produce_FD*" parameter estimate is positive and significant. Nonetheless, they are less willing to pay for quality in absolute terms compared to the "1 mile or less" group. Consequently, it would be difficult to conclude based on these results that produce quality would drive these respondents from shopping at a supercenter to shopping at a grocery store if grocery stores were more accessible. The following analysis attempts to parse such effects by accounting for respondent's perception of their food environment.

5.3.2 Endogenous Food Geography Choice Experiment Results

In the following analysis, I consider a subset of the previous analysis and compare respondents who live within one mile of a grocery store who indicated satisfaction with their current food geography by selecting status quo option above average, and respondents who live over one mile of a grocery store who indicated dissatisfaction with the current food geography by selecting the status quo option less often than average.

To compare grocery store preferences between "Constrained" and "Ideal" respondents, I multiply each main effect by a dummy variable, denoted by *CD*, equal to 1 if the respondent is in the "Constrained" group, and equal to 0 if the respondent is in the "Ideal" group.

Results of the main effects MMNL regressions are presented in Table 20. Regression results for a pooled sample of the "Ideal" and "Constrained" group is uninformative, thus I show only regression results for the main effects MMNL incorporating the *CD* dummy variable to highlight where preferences diverge for the two sample groups. Each main effect parameter except for *AQ Produce* has significant standard deviation around their mean, which implies individual preferences are heterogeneous for each attribute except the utility gained from average over low quality produce. This result justifies, again, employing the MMNL model as opposed to the CL model. The sign, significance, and magnitude of each main effect coefficient are similar to that of Table 17 thus results for the "1 mile or less" respondents, which here correspond to the "Ideal" respondent group, are robust and mimic the discussion in section 5.3.1

The change in utility for a change in all main effect attributes excepting *Travel Time* are significantly different for respondents in the "Constrained" group compared to respondents in the "Ideal" group.

Table 20: MMNL Results with Main Effects and "Constrained" dummies

	MMNL w dummy		
Variable	Coef *(SE)*		SD Pval[a]
Main Effects			
ASC	-2.946	***	
	0.157		
Weekly Grocery Bill	-0.020	***	
	0.006		
Travel Time	-0.126	***	0.002
	0.021		
Specialty Store	-0.370	*	0.010
	0.223		
Grocery Store	-1.069	***	0.000
	0.198		
HQ Produce	2.091	***	0.005
	0.228		
AQ Produce	0.077		0.675
	0.207		
Main Effect 'Constrained' Dummies			
ASC_CD	3.144	***	0.000
	0.363		
Weekly Grocery Bill_CD	-0.036	**	0.002
	0.015		
Travel Time_CD	0.000		0.209
	0.032		
Specialty Store_CD	0.891	*	0.015
	0.469		
Grocery Store_CD	0.616	*	0.132
	0.363		
HQ Produce_CD	2.517	***	0.001
	0.652		
AQ Produce_CD	0.717	*	0.223
	0.381		
Observations	4,866		
Log Liklihood	-786.85		
Prob > Chi2	0.000		
AIC	1611.70		
BIC	1735.01		

*** significant at 1%; ** significant at 5%, * significant at 10%

[a]Significance of standard deviations of individual heterogeneity

I continue discussion of preferential differences between respondents in the “Ideal” and “Constrained” groups using their marginal WTP for grocery store attributes. Marginal WTP values for the main effects and “Constrained” dummy MMNL model is shown in Table 21**.**

Table 21: Marginal Willingness to Pay - Main Effects

	Ideal	**Constrained**	Significant Difference?*
Travel Time ($/min. travel)	-$6.17	-$2.25	No
Specialty Store ($/specialty vs supercenter)	-$18.08	$9.24	Yes
Grocery Store ($/grocery vs supercenter)	-$52.25	-$8.04	Yes
HQ Produce ($/HQ vs LQ produce)	$102.21	$81.77	Yes
AQ Produce ($/AQ vs LQ produce)	$3.79	$14.10	Yes

*Significant difference between sample groups of utility for the attribute, not marginal WTP for the attribute

The respondents in the "Ideal" group appear to place substantially higher premium on their time than do respondents in the "Constrained" group, requiring a compensation of $6.17 compared to $2.25 for the "Constrained" group. The "Constrained" group is likely insensitive to added travel time compared to the "Ideal" group who have the luxury of short commutes to the grocery store. The coefficient on *Travel Time* is insignificant, thus the "Constrained" respondents do not experience significantly different utility from a change in travel time than do respondents in the "Ideal" group.

The "Constrained" group, again, appears to be more price-sensitive than the "Ideal" group, since the *Weekly Grocery Bill_CD* variable is significant and negative. The "Constrained" group experiences more disutility from a higher grocery bill than does the "Ideal" group, all else equal. Similarly, fresh produce demand for the "Constrained" group was highly expenditure- and price- elastic compared to the "Ideal" group (see Table 13).

"Ideal" respondents and "Constrained" respondents have significantly different preferences for specialty stores. "Ideal" category respondents require compensation of $18.08 to shop at a specialty store instead of a supercenter whereas "Constrained" respondents will *pay* roughly half as much, $9.24, to shop at a specialty store instead of a supercenter. Both types of respondents require compensation to shop at a grocery store relative to a supercenter, though "Ideal" respondents require over five times as much compensation, of $52.25 compared to that of the "Constrained" group who require $8.09. The "Ideal" group places high premium on variety and access to non-food items and services when grocery shopping.

Similar to the previous exogenous analysis, respondents perceive the quality differential between low and high quality to be significantly different from the quality differential between low and average quality. The "Ideal" group perceives high quality produce to be nearly 25 times more valuable than average quality produce based on their marginal WTP of $3.79 for average and $102.21 for high quality. The "Constrained" group has a lower

marginal WTP for high quality produce compared to the "Ideal" group, yet they have a higher marginal WTP for average quality produce. Further, the "Constrained" group perceives the quality premium between high quality and average quality produce to be approximately six times as much, a substantially lower differential than that of the "Ideal" group. Their marginal WTP for average over low quality is $14.10, and their marginal WTP for high over low quality is approximately six times higher at $81.77. The "Constrained" group not only values average quality produce more than the "Ideal" group, but values average quality produce relative to high quality produce to a higher degree than the "Ideal" group. Some of the difference in quality perception between the groups may be explained by preference endogeneity. The "Ideal" group may have better access to high quality products where the "Constrained" group may have limited or no access to products they consider to be high quality. Consequently, their limited consumption of very high quality produce diminishes the value they can hypothetically place on such produce because they may be less familiar with high quality produce relative to their "Ideal" food geography counterparts.

Based on these results, the "Constrained" group appears to be dissatisfied with the amenities of their current grocery retailers. Recall that 89.3 percent of the "Constrained" group shopped at a supercenter in the month prior to the survey, a proportion well above the sample average of 60.3 percent. Yet, this group would actually prefer to shop at a specialty store instead of a supercenter. Also, the "Constrained" respondents were least likely to have shopped at a specialty store in the month prior to the survey compared to other sample groups, thus is appears this group is dissatisfied with their lack of access to specialty stores. Travel time and produce quality, on the other hand, do not appear to be the constraining factor. The "Constrained" group is agnostic to added travel time, and has lower marginal WTP for high quality produce compared to the "Ideal" group. Thus their dissatisfaction may stem from some other grocery store attribute such as store ambience, layout, or product offerings. Income clearly is not driving the "Constrained" group to shop at supercenters, since they had the largest mean and medium income of all other groups. Based on their preferences for specialty and grocery store types, it appears the "Constrained" group indeed would prefer to shop at a store type different from a supercenter, but geographic constraints prevent them from doing so. In Section 5.3.3 I perform a robustness check on preferences of the "Constrained" group.

In order to control for the any possible effects of proximity, travel time, and income on the marginal WTP results, I estimated a second set of MMNL models with interaction terms. Table 22 presents results from the MMNL model with *Distance*, *CurrentTravel*, and *Income* interaction terms with the *CD* "Constrained" dummy variable. Again, a pooled interaction model provides little information, so I present only results of the MMNL regression incorporating the *CD* dummy to highlight differences between the "Constrained" and "Ideal" subsamples. The MMNL interaction models have similar results to that of the main effects model, however the *Weekly Grocery Bill_CD* variable is no longer significant. Of the interaction terms, *Distance x Grocery Bill* is solely significant for all

models. The negative value implies that people who live further away from their nearest grocery store will have a lower marginal willingness to pay for grocery store attributes than those who live closer. As with the previous analysis, each triple interaction term is insignificant. This indicates the interaction effects of distance and travel time on marginal WTP, and income on produce quality are not significantly different for respondents in the "Constrained" group compared to respondents in the "Ideal" group. Consequently, I again drop the triple interaction terms for simplicity to calculate the marginal WTP for grocery store attributes. The results of the interaction effects MMNL without triple interaction terms are presented in the last columns of Table 22. Again, the simpler model appears to have the best fit based on the low AIC and BIC values.

Table 22: MMNL results with Interaction Effects and "Constrained" dummies

	MMNL w/ dummy 1		MMNL w/ dummy 2		MMNL w/ dummy 3		MMNL w/ dummy 4		MMNL w/ dummy w/o triple effects	
Variable	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]
Main Effects										
ASC	-2.875 ***		-2.856 ***		-2.864 ***		-2.846 ***		-2.863 ***	
	0.156		*0.151*		*0.152*		*0.153*		*0.151*	
Weekly Grocery Bill	-0.041 **		-0.013		-0.017 *		-0.009		-0.013	
	0.018		*0.010*		*0.009*		*0.010*		*0.009*	
Travel Time	-0.114 ***	0.011	-0.118 ***	0.004	-0.119 ***	0.000	-0.114 ***	0.000	-0.117 ***	0.007
	0.020		*0.020*		*0.019*		*0.018*		*0.019*	
Specialty Store	-0.544 **	0.022	-0.674 **	0.000	-0.570 **	0.000	-0.586 **	0.005	-0.595 **	0.000
	0.247		*0.271*		*0.244*		*0.268*		*0.261*	
Grocery Store	-1.074 ***	0.015	-0.996 ***	0.005	-1.047 ***	0.004	-1.031 ***	0.189	-1.029 ***	0.006
	0.204		*0.192*		*0.197*		*0.195*		*0.193*	
HQ Produce	1.278 **	0.011	1.390 **	0.135	1.473 **	0.291	1.356 **	0.789	1.621 ***	0.034
	0.648		*0.649*		*0.666*		*0.643*		*0.630*	
AQ Produce	-0.495	0.534	-0.464	0.886	-0.472	0.948	-0.421	0.688	-0.220	0.380
	0.666		*0.651*		*0.620*		*0.679*		*0.631*	
Main Effect 'Constrained' Dummies										
ASC_CD	2.685 ***	0.000	2.646 ***	0.001	2.812 ***	0.000	3.064 ***	0.000	2.744 ***	0.000
	0.399		*0.492*		*0.361*		*0.352*		*0.343*	
Weekly Grocery Bill_CD	0.010	0.248	-0.019	0.573	-0.016	0.036	0.007	0.025	-0.002	0.005
	0.024		*0.028*		*0.020*		*0.019*		*0.018*	
Travel Time_CD	0.002	0.238	-0.027	0.043	-0.004	0.140	-0.009	0.737	-0.006	0.397
	0.032		*0.041*		*0.032*		*0.031*		*0.030* **	
Specialty Store_CD	1.133 **	0.001	1.184 **	0.299	0.885 *	0.018	1.029 **	0.166	1.156	0.027
	0.486		*0.548*		*0.500*		*0.509*		*0.485* **	
Grocery Store_CD	1.009 ***	0.803	0.889 **	0.488	0.961 **	0.700	1.001 ***	0.537	0.942	0.366
	0.394		*0.415*		*0.376*		*0.387*		*0.388* ***	
HQ Produce_CD	2.029 ***	0.886	3.060 ***	0.000	2.044	0.000	2.200 ***	0.033	2.375	0.008
	0.524		*0.936*		*1.417*		*0.567*		*0.603*	
AQ Produce_CD	0.421	0.006	0.873 *	0.008	0.743 *	0.313	-0.027	0.617	0.620	0.016
	0.462		*0.455*		*0.392*		*1.270*		*0.407*	
Interaction Effects										
Distance x Grocery Bill	0.038	0.722	-0.012 *	0.453	-0.010 **	0.172	-0.014 ***	0.578	-0.012 ***	0.810
	0.028		*0.006*		*0.004*		*0.005*		*0.004*	
CurrentTravel x Grocery Bill	0.000	0.658	0.000	0.794	0.000	0.891	0.000	0.443	0.000	0.950
	0.001		*0.001*		*0.001*		*0.001*		*0.001*	
Income x HQ Produce	0.142	0.869	0.124	0.901	0.107	0.023	0.124	0.004	0.077	0.825
	0.114		*0.114*		*0.118*		*0.113*		*0.109*	
Income x AQ Produce	0.101	0.438	0.098	0.879	0.099	0.593	0.089	0.824	0.050	0.349
	0.118		*0.116*		*0.110*		*0.122*		*0.111*	
Interaction effects 'Constrained' Dummies										
Distance x Grocery Bill_CD	-0.045	0.251								
	0.028									
CurrentTravel x Grocery Bill_CD			0.001	0.149						
			0.002							
Income x HQ Produce_CD					0.087	0.101				
					0.219					
Income x AQ Produce_CD							0.099	0.358		
							0.216			
Observations	4,614		4,614		4,614		4,614		4,614	
Log Liklihood	-752.36		-749.50		-750.47		-749.23		-748.62	
Prob > Chi2	0.000		0.000		0.000		0.000		0.000	
AIC	1542.72		1536.99		1538.95		1536.46		1533.23	
BIC	1665.02		1659.29		1661.25		1658.76		1649.09	

*** significant at 1%; ** significant at 5%, * significant at 10%

[a]Significance of standard deviations of individual heterogeneity

Table 23 shows the marginal WTP for each grocery store attribute based on the interaction effect specification in the final columns of Table 22**.** As with the previous analysis, unit values for the respondent characteristic variables of travel time, distance, and income are based on the average for each sample group.

Table 23: Marginal Willingness to Pay - Interaction Effects

	Ideal	**Constrained**	Significant Difference?*
Travel Time ($/min. travel)	-$5.78	-$1.83	No
Specialty Store ($/specialty vs supercenter)	-$29.48	$8.36	No
Grocery Store ($/grocery vs supercenter)	-$50.96	-$1.30	No
HQ Produce ($/HQ vs LQ produce)	$101.62	$66.40	Yes
AQ Produce ($/AQ vs LQ produce)	$3.04	$10.39	No
Unit Values			
Average Travel Time (min.)	7.4	12.3	
Average Distance (mi.)	0.5	4.1	
Average Income Bracket^	5.6	5.9	

*Significant difference between sample groups of utility for the attribute, not marginal WTP for the attribute

^Income bracket of 4 denotes annual income between $50-75K; 5 denotes $75-100K; 6 denotes $100-125K

After controlling for respondent proximity, current travel time, and income, "Constrained" respondents no longer garner significantly different disutility from a higher grocery bill than do the "Ideal" respondents, since the *Weekly Grocery Bill_CD* coefficient is insignificant. Their price-sensitivity from the preference perspective is not dissimilar from that of the "Ideal" group. The "Constrained" group, however, appeared price sensitive to fruit & vegetable purchases in the demand system. This group is impartial to total grocery bill changes, but highly sensitive to fresh produce prices, suggesting that their food geography may affect fruit & vegetable consumption.

Utility from grocery or specialty stores and average quality produce is not significantly different between the "Constrained" and "Ideal" groups, however the differentials have gotten wider. "Constrained" group willingness to pay to grocery shop at a specialty store instead of a supercenter was formerly over $27 higher than that of the "Ideal" group ($9.24 - $-18.08). Now, the "Constrained" group is willing to pay almost $38 more than the "Ideal" group to shop at a specialty store over a grocery store ($8.36 - $-29.48). Willingness to pay to shop at a grocery store experienced a similar change. It should be noted, however, that the "Constrained" group is willing to pay, in earnest, to patronize a specialty store instead of a supercenter, whereas both the "Constrained" and "Ideal" respondents need compensation to shop at a grocery store instead of a supercenter. After taking into consideration distance, current travel time, and income of the respondents, the "Constrained" group has noticeably lower willingness to pay for high quality produce, $66.40 compared to their previous marginal WTP of $81.78. Further, their willingness to pay for average quality produce dropped in absolute terms. Their marginal

WTP to shop at a grocery store with average quality instead of low quality produce declined from $14.10 to $10.39. Thus, the "Constrained" group perceives high quality produce to be less valuable, after controlling for their income levels. In absolute terms, however, the "Constrained" group has a lower willingness to pay for produce quality than does the "Ideal" group. Their tendency to treat vegetables as a residual category may indeed be driven by legitimate displeasure from vegetables as opposed to some exogenous constraint. Results of the second MMNL regressions do, however, support the hypothesis that the "Constrained" group would likely shop at a specialty-type store instead of a supercenter if such a store were accessible. How much this would change their fresh produce consumption is an open question. Yet, if we believe the findings of Blitstein, Snider, and Evans (2011), as well as the correlations suggested by my own data, the more satisfied a consumer is with their produce retailer, the more fresh fruits & vegetables they will consume.[16]

5.3.3 Comparison of Constrained versus Willing Travelers

Another way to investigate whether quality, travel time, or some other factor might drive volatile fresh produce consumption is to compare the preferences of respondents who willingly travel further than they have to, the "Opt to Travel" respondents, against the "Constrained" group. For example, I can infer that lack of access to quality produce "constrains" the "Constrained" group if their marginal WTP for travel time is similarly low to that of the "Opt to Travel" group, but their marginal WTP for produce quality is significantly higher than that of the "Opt to Travel" group. Recall from Table 3 that the "Opt to Travel" respondents chose their status quo option 3.0 times out of six, on average, and the "Constrained" group chose their status quo option 1.1 times out of six on average. Both groups are relatively dissatisfied with their current grocery shopping situation compared to the full sample of respondents who chose their status quo option 3.3 times out of six. I investigate how their preferences and behavior compare.

Table 24 presents MMNL regression results for a comparison of the "Constrained" and "Opt to Travel" respondent groups. The dummy variable here is equal to 1 if the respondent falls in the "Opt to Travel" group and equal to 0 of the respondent falls in the "Constrained" group. Pooled regression results for these two groups are uninformative, thus I present only results of the main effects model with the "Opt to Travel" dummy variable interacted with each main effect. The "Opt to Travel" dummy is denoted by *OCD.* Interaction effects MMNL specification results are presented on the right columns. Triple interaction terms with the *OCD* dummy were each insignificant, as with the previous models. Results from the triple interaction specifications are available upon request. The second MMNL model with interaction terms has improved fit compared to the main effects specification, with a lower AIC and BIC. I will focus the following discussion on the Interaction effects specification.

[16] Based on the results shown in Table 3, Table 7, Table 9, ranking of respondent satisfaction with their current grocery stores - as indicated by the number of times they selected the status quo option – follows "Ideal", "1 mile or less", "Over 1 mile", and "Constrained." This ordering is identical when considering the grocery budget shares devoted to fresh produce.

The *ASC* term is insignificant, implying the "Constrained" group has negligible probability of choosing their status quo option. The "Constrained" group has markedly different behavior than the other consumer groups because they do not garner greater utility from grocery shopping at the store they are most familiar with. The *ASC_OCD* "Opt to Travel" dummy is negative and significant, however. This implies the "Opt to Travel" group is more likely to choose their status quo option. Such results mimic my findings in in Table 3 whereby the "Constrained" group were least likely to choose their status quo option.

Not surprising is both groups have similar disutility for added travel time. This result confirms that the "Constrained" group respondents are not likely "constrained" by their longer trips to the grocery store. Even though the "Constrained" group must travel far to get to a grocery store, their preferences are no different than respondents who choose to travel far regardless of their proximity to a grocery store. All other preferences for attributes are statistically similar: both are more likely to choose a grocery store with high quality produce instead of low quality produce, and both are less likely to choose a grocery store over a supercenter. Since both groups have relatively high average income brackets, it is not surprising that they appear equally ambivalent about a higher grocery bill. Yet the expenditure elasticities for all food groups were significantly more elastic for the "Constrained" group compared to the "Opt to Travel" group, with the exception of "Other Foods" (see Table 13)**.** This lends evidence that the "Constrained" group have volatile fruit & vegetable consumption because of dissatisfaction with some attribute of their food geography unrelated to income constraints or traveling constraints. Results of Table 25 illustrate each group's marginal willingness to pay for each store attribute.

Table 24: MMNL results: Main and Interaction Effect specifications with "Opt to Travel" dummies

	MMNL Main Effects		MMNL Interaction Effects	
Variable	Coef *(SE)*	SD Pval[a]	Coef *(SE)*	SD Pval[a]
Main Effects				
ASC	0.136		0.048	
	0.201		*0.221*	
Weekly Grocery Bill	-0.035 ***		-0.009	
	0.008		*0.018*	
Travel Time	-0.082 ***	0.001	-0.096 ***	0.007
	0.018		*0.021*	
Specialty Store	0.051	0.025	0.005	0.000
	0.262		*0.319*	
Grocery Store	-0.524 **	0.002	-0.522 *	0.000
	0.247		*0.275*	
HQ Produce	2.976 ***	0.007	2.074 ***	0.056
	0.351		*0.694*	
AQ Produce	0.650 ***	0.424	-0.306	0.753
	0.223		*0.559*	
Main Effect 'OptConstrained' Dummies				
ASC_OCD	-1.807 ***	0.000	-1.839 ***	0.000
	0.272		*0.297*	
Weekly Grocery Bill_OCD	-0.008	0.000	-0.019	0.014
	0.010		*0.016*	
Travel Time_OCD	-0.032	0.260	-0.025	0.000
	0.022		*0.024*	
Specialty Store_OCD	-0.089	0.008	-0.194	0.083
	0.315		*0.372*	
Grocery Store_OCD	-0.161	0.330	-0.213	0.203
	0.283		*0.320*	
HQ Produce_OCD	0.403	0.000	0.391	0.000
	0.400		*0.437*	
AQ Produce_OCD	-0.190	0.831	0.002	0.569
	0.271		*0.311*	
Interaction Effects				
Distance x Grocery Bill			-0.007 *	0.303
			0.003	
CurrentTravel x Grocery Bill			-0.001	0.019
			0.001	
Income x HQ Produce			0.247 **	0.043
			0.105	
Income x AQ Produce			0.167 *	0.263
			0.088	
Observations	5,601		5,346	
Log Liklihood	-1360.24		-1283.54	
Prob > Chi2	0.000		0.000	
AIC	2748.47		2603.09	
BIC	2841.30		2721.60	

*** significant at 1%; ** significant at 5%, * significant at 10%

[a]Significance of standard deviations of individual heterogeneity

Table 25: Marginal Willingness to Pay - Main Effects and Interaction Effects

	Main Effects		Interaction Effects		
	Constrained	Opt to Travel	Constrained	Opt to Travel	Significant Difference?*
Travel Time ($/min. travel)	-$2.36	-$2.67	-$2.11	-$3.02	No
Specialty Store ($/specialty vs supercenter)	$1.47	-$0.88	$0.12	-$4.68	No
Grocery Store ($/grocery vs supercenter)	-$15.11	-$16.11	-$11.45	-$18.25	No
HQ Produce ($/HQ vs LQ produce)	$85.78	$79.41	$77.45	$93.07	No
AQ Produce ($/AQ vs LQ produce)	$18.74	$10.82	$14.91	$14.01	No
Unit Values					
Average Travel Time (min.)	N/A	N/A	12.3	12.1	
Average Distance (mi.)	N/A	N/A	4.0	0.4	
Average Income Bracket^	N/A	N/A	5.9	5.2	

*Significant difference between sample groups of utility for the attribute, not marginal WTP for the attribute
^Income bracket of 4 denotes annual income between $50-75K; 5 denotes $75-100K; 6 denotes $100-125K

The "Constrained" group has similar marginal WTP for high quality produce as does the "Opt to Travel" group. In the main effects specification, the "Constrained" group actually has higher marginal WTP for high quality produce of $85.78 compared to $79.41 for the "Opt to Travel" group. After accounting for respondent proximity to their nearest grocery store, current travel, and income, the "Constrained" group indeed has lower marginal WTP for high quality, but the difference is small compared to previous respondent group comparisons. Despite their relatively similar value for produce quality, the "Constrained" group consumed five percent less fresh produce and almost five percent more non-fruit & vegetable "Other Foods" than did the "Opt to Travel Group" (see Table 9.) One explanation is the "Opt to Travel" respondents may be shopping at stores with significantly better produce quality than the "Constrained" group, even though both are traveling similar distances to get to their main grocery store. Note the average stated travel times are 12.3 and 12.1, respectively. Both groups have a relatively high incidence of shopping at supercenters in the month prior to the survey, as shown by Table 4. Yet the marginal WTP for store types suggest the "Opt to Travel" group actually prefers to shop at supercenters whereas the "Constrained" group would prefer to shop at a specialty store.

The "Opt to Travel" group has distinct consumption behavior from the "Constrained" group, despite the fact that their preferences are similar. Proximity to a grocery store is not the most likely cause of the divergent consumption behavior because the "Opt to Travel" group are traveling similar distances and shopping at similar stores as the "Constrained" group. This provides evidence that the exogenous food environment is not solely responsible for fresh produce consumption.

Chapter Six: Summary of Findings and Conclusions

The objective of this thesis was to identify a relationship between food geography and consumption behavior. My identification strategy involved examining the discrete choice of where to grocery shop against the continuous choice of which food products to purchase at the grocery store. I used a non-market valuation method known as a choice experiment to quantify value perception of grocery store attributes, and I estimated a system of demand for grocery items using original survey data from 991 residents of Champaign & Piatt counties, IL.

This research was motivated by growing awareness of the limitations of the "food desert" construct both in the media, and in the academic community. Current literature on food access and food geography has demonstrated that a greater number of supermarkets in a finite geographic area are positively correlated with greater health for residents of that area, but causality remains inconclusive. Literature has also found that "food desert" residents have relatively elastic demand for fresh produce, however it remains undetermined whether such demand behavior is a direct result of food geography, or a result of inherent preferences among "food desert" residents.

The methodology described in this thesis contributes to the current body of literature on food demand and food access in several ways. First, I address the simultaneous choice problem inherent in grocery shopping in which the choice of where to grocery shop has direct implications on what one decides to purchase. Second, I provide an approach for separating endogenous sorting from exogenous treatment effects in order to isolate how the food environment can influence fruit & vegetable consumption.

I find proximity to a grocery store is positively related to stable consumption of fresh produce and decreases the propensity to substitute fresh for non-fresh vegetables. Lack of access to specialized, small-scale retailers, as well as higher quality produce appears to be the most significant drivers of price elastic produce consumption and a higher frequency of trading off non-fresh vegetables for fresh vegetables. Neither added travel time nor expense appears to affect consumption behavior among survey respondents. Previous literature and the results of summary statistics from this thesis' sample data imply satisfaction with one's grocery store and a higher perception of quality value are positively related to fruit & vegetable consumption. Thus travel time and geographic distance in isolation appear less likely to alter produce consumption than does access to quality products and quality retailers.

I further find that proximity to a grocery store is positively related to consumption of fresher foods even for respondents who travel further than necessary to grocery shop. Respondents who willingly traveled over 1 mile to grocery shop exhibited similar preferences as respondents who were forced to travel over 1 mile to grocery shop. However, the respondents who opt to travel consumed fresh produce to a greater extent. This finding suggests that the food environment may have some indirect spillover effects that cause nearby residents to choose fresh

foods more often than non-fresh foods. Alternatively, this finding suggests that an exogenous food desert definition based on geographic boundaries overstates the influence of the food environment, thus overstates the potential constraining effects of the food environment. A significant sample of respondents in this study appears to consume in a particular manner regardless of their food environment, and regardless of living in close proximity to a grocery store.

The exogenous, geographic-based analysis on food access suggested that faculty and staff of the University of Illinois who live over 1 mile from a grocery store exhibit greater stability of fresh produce consumption, compared to those who live over one mile. The endogenous, preference-based analysis on food access suggested that constraining food geography could intensify the propensity to substitute non-fresh for fresh produce. The endogenous analysis further demonstrated that lack of access to specialty food stores and dissatisfaction with supercenter stores were the most significant forces causing people who live over 1 mile from a grocery store to trade off non-fresh for fresh produce.

Both endogenous sorting and environmental treatment effects of the food environment influence food consumption. While respondents who live within one mile of a grocery store consumed greater proportions of, and had inelastic demand for, fresh produce compared to those who live over one mile from a grocery store, these proximate respondents also exhibited stronger preferences for proximity and high quality produce. On the other hand, a significant portion of respondents exhibited behavior and preferences that suggested they would consistently purchase fresh produce if better quality stores different from supercenters were accessible to them.

Current policy response to inadequate food access focuses first on a geographic, exogenous definition to define areas as “food deserts,” and second encourages grocery chains to enterprise in underserved areas. This approach may be more likely to encourage fresh produce consumption of nearby residents if policy makers prioritize the quality improvement and focus efforts on advancing store amenity and product quality offerings. However, alternative means of health intervention, such as nutrition education are likely to be necessary for populations of people for whom personal preferences are stronger determinants of consumption behavior than their exogenous food environment.

The methods suggested in this thesis are not without limitations. I do not have a statistically sound technique for concluding whether results of the demand system are consistent with preferential results of the choice experiment analyses. My conclusions are specific to a sample of people who are not representative of the US as a whole, thus I cannot generalize my findings beyond the sample of residents of Champaign & Piatt counties affiliated with the University of Illinois. Lastly, a clear limitation of this thesis is that it does not take into account the endogeneity of household location.

References

Abdulai, Awudu (2002). "Household Demand for Food in Switzerland: A Quadratic Almost Ideal Demand System. *Swiss Society of Economics and Statistics*, 138(1): 1-18.

Alpizar, Francisco, Fredrik Carlsson, and Peter Martinsson. 2001. "Using Choice Experiments for Non-Market Valuation." Working paper. Göteborg University Department of Economics.

Andreyeva, Tatiana, Michael Long, and Kelly Brownell. 2010. "The Impact of Food Prices on Consumption: A Systematic Review of Research on the Price Elasticity of Demand for Food." *American Journal of Public Health,* 100(2): 216-222.

Arnold, Stephen J, Tae H Oum & Douglas J. Tigert. 1983. "Determinant Attributes in Retail Patronage: Seasonal, Temporal, Regional, and International Comparisons." *Journal of Marketing Research* Vol. 20 (May 1983): 149-57.

Banks, James, Richard Blundell, and Arthur Lewbel. (1997). "Quadratic Engel Curves and Consumer Demand." *The Review of Economics and Statistics*, 79(4): 527-539.

Becker, Gary. 1965. "A Theory of the Allocation of Time." *Economic Journal*. 75:493-517.

Bentham, Jeremy. 1789. *Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press.

Berg, Nathan and James Murdoch. 2007. "Access to Grocery Stores and Food Security in Dallas." Research Brief. The J. McDonald Williams Institute Center for Urban Economics. University of Texas – Dallas.

Bitler, Marianne and Steven J. Halder. 2011. "An Economic View of Food Deserts in the United States." *Journal of Policy Analysis and Management.* 30(1): 153-176.

Blisard, Noel, Hayden Stewart, and Dean Jolliffe. 2004. *Low-Income Households' Expenditures on Fruits and Vegetables.* Washington DC: US Department of Agriculture, ERS Agricultural Economic Report Number 833, May.

Blitstein, Jonathan, Jeremy Snider, and W Douglas Evans. 2011. "Perceptions of the food shopping environment are associated with greater consumption of fruits and vegetables." *Public Health Nutrition* 15(6): 1124-1129.

Brown, Daniel. 1978. "An Examination of Consumer Grocery Store Choice: Considering the Attraction of Size and the Friction of Travel Time." *NA – Advances in Consumer Research.* Vol 5.

Brown, Stephen. 1989. "Retail Location Theory: The Legacy of Harold Hotelling," *Journal of Retailing* 65(4): 450-470.

California Center for Public Health Advocacy. 2007. *Searching for Healthy Food: The Food Landscape in California Cities and Counties*. Davis, CA.

Carlson, Andrea, Mark Lino, WenYen Juan, Kenneth Hanson, and P. Peter Basiotis. 2007. *Thrifty Food Plan, 2006*. Washington DC: US Department of Agriculture, ERS CNPP-19, April.

Carrasco, N. (2008) Deciding where to shop: disaggregate random utility destination choice modeling of grocery shopping in canton Zurich, *Master Thesis,* IVT, ETH Zurich, Zurich. < www.ivt.ethz.ch/docs/students/sa223.pdf>

Centers for Disease Control & Prevention. National Center for Chronic Disease Prevention and Health Promotion. Division of Nutrition, Physical Activity, and Obesity. "State Initiatives Supporting Healthier Food Retail: An Overview of the National Landscape." www.cdc.gov/obesity/downloads/Healthier_Food_Retail.pdf

Chen, S., Florax, R.J.G.M., and Snyder, S.D. (2010). Does where you live make you fat? Obesity and access to chain grocers. *Economic Geography*, 86, 431-452.

Currie, J., DellaVigna, S., Moretti, E., and Pathania, V. (2009). The effect of fast food restaurants on obesity and weight gain. *NBER Working Paper No. 14721*, Retrieved March 2012 from http://www.nber.org/papers/w14721.

Deaton, Angus (1988). "Quality, quantity, and spatial variation of price." *American Economic Review*, 78(3): 418-430.

Dissanayake, Sahan and Amy Ando (2011). "Valuing restoration: conservation tradeoffs and respondent experience." Working paper.

Deaton, Angus and John Muellbauer. 1980a. "An Almost Ideal Demand System." *American Economic Review*, Vol. 70, 899-36.

Deaton, Angus and John Muellbauer. 1980b. *Economics and Consumer Behavior*. Cambridge University Press.

"Detroit Grocery Stores, 'Food Desert' Myth Examined in Noah Stephens' Photography Survey." Huffington Post. 22 October 2012. http://www.huffingtonpost.com/2012/10/22/detroit-grocery-stores-food-desert-noah-stephens_n_2001896.html. Viewed 16 February 2013.

Drewnoski, A (2003). "Fat and sugar: An economic analysis" *Journal of Nutrition,* 133, 838S-840S.

Drewnowski, A., Darmon, N., & Briend, A. (2004). "Replacing fats and sweets with vegetables and fruit – a question of cost." *American Journal of Public Health*, 94, 1555-1559.

Durham, Catherine and James Eales. (2010) "Demand elasticities for fresh fruit at the retail level." *Applied Economics*, 42: 1345-1354.

Eales, James S. and Laurian J Unnevehr. 1994. "The inverse almost ideal demand system." *European Economic Review,* 38:101-115.

Erdil, Erkan. 2005. "Demand systems for agricultural products in OECD countries." *Applied Economics Letters*, 13:163-169.

Gibson, John and Scott Rozelle. (2005). "Prices and Unit Values in Poverty Measurement and Tax Reform Analysis." World Bank Economic Review 19(1): 69-98.

Gibson, D.M. (2011). The neighborhood food environment and adult weight status: estimates from longitudinal data. *American Journal of Public Health*, 101: 71.

Giskes, K., van Lenthe, F., Avendano-Pabon, M., and Brug, J. (2011). A systematic review of environmental factors and obesogenic dietary intakes among adults: Are we getting closer to understanding obesogenic environments? *Obesity Review*, 12: e95-e106.

Gómez, Miguel, Edward McLaughlin, and Dick Wittink. 2004. "Customer satisfaction and retail sales performance: an empirical investigation." *Journal of Retailing*. 80(2004): 265 – 278.

Gorman, W.M. (1981). "Some Engel Curves" , in A. Deaton (ed). *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone,* Cambridge.

Greene, William. (a) Lecture notes. "Discrete Choice Modeling." Discrete Choice Modeling. New York University. <https://files.nyu.edu/mrg217/public/unordered_multiresponse.pdf> Visited 24 February 2013.

Greene, William. (b) Lecture notes. "Discrete Choice Models I." Discrete Choice Modeling. New York University. <http://pages.stern.nyu.edu/~wgreene/DiscreteChoice.htm> Visited 3 March 2013.

Hanemann, W.M. (2006). The economic conception of water, In: *Water Crisis: myth or reality?* Eds. P.P. Rogers, M.R. Llamas, L. Martinez-Cortina, Taylor & Francis plc., London.

Harvard School of Public Health. "The Nutrition Source." <http://www.hsph.harvard.edu/nutritionsource/>

Hensher, Rose, Greene. *Applied Choice Analysis: A Primer* (2005). Cambridge University Press.

Heiss, Florian. 2002. "Structural choice analysis with nested logit models." *The Stata Journal.* 2(3): 227-252.

Herrmann, Roland and Claudia Roeder. 1998. "Some neglected issues in food demand analysis: retail-level demand, health information and product quality." *The Australian Journal of Agricultural and Resource Economics*. 42(4): 341 – 367.

Hoch, Stephen, Byung-Do Kim, Alan Montgomery, and Peter Rossi. 1995. "Determinants of Store-Level Price Elasticity." *Journal of Marketing Research* 32(1): 17 –29.

Hsieh, Ming-Feng and Kyle W. Stiegert. 2011. "Store Format Choice in Organic Food Consumption." *American Journal of Agriculture Economics*. 94(2):307-313.

Huang, Kuo S. and Giing-Hwan Lin (200). "Estimation of Food Demand and Nutrient Elasticities from Household Survey Data." Economic Research Service, U.S. Department of Agriculture. Technical Bulletin No. 1887.

Huang, Yan and Harmen Oppewal (2006), "Why consumers hesitate to shop online: An experimental choice analysis of grocery shopping and the role of delivery fees", *International Journal of Retail & Distribution Management,* Vol. 34.4:334-53.

Katchova, Ani L. and Wen S. Chern (2004). "Comparison of Quadratic Expenditure System and Almost Ideal Demand System Based on Empirical Data." *International Journal of Applied Economics*, 1(1): 55-64.

Kuhfeld, Warren. 2010. "Marketing Research Methods in SAS: Experimental Design, Choice, Conjoint, and Graphical Techniques," SAS 9.2 Edition MR-2010, 1 October 2010.

Lancaster, Kelvin. (1966) "A New Approach to Consumer Theory." The Journal of Political Economy. 74(2): 132-157.

Larson N, Story M, and Nelson M. (2009), "Neighborhood environments: disparities in access to healthy foods in the U.S.", *American Journal of Preventative Medicine* 2009, 36: 74-81.

Lee, Jonq-Ying, Mark G. Brown and James L. Seale, Jr. 1994. "Model Choice in Consumer Analysis: Taiwan, 1970-89." *American Journal of Agricultural Economics*, 76:504-512

Lewbel, Arthur. (1990). "Full Rank Demand Systems." *International Economic Review*, 31(2): 289-300.

Liese, Angela D. Kristina E. Weis, Delores Pluto, Emily Smith, et al. 2007. "Food Store Types, Availability, and Cost of Foods in a Rural Environment." Journal of the American Dietetic Association 107:1916-1923.

Louviere, J.J., D.A. Hensher and J.D. Swait, (2000). *Stated Choice Methods: Analysis and Applications.* Cambridge University Press.

Mari Gallagher Research & Consulting Group. 2006. *Examining the Impact of Food Deserts on Public Health in Chicago.* Chicago, IL.

McKelvey, Christopher. 2011. "Price, unit value, and quantity demanded." *Journal of Development Economics.* 95(2011): 157-169.

Monsivais, Pablo and Adam Drewnowski (2007). "The Rising Cost of Low-Energy-Density Foods," *Journal of the American Dietetic Association* 107(12): 2071-2076.

Morland K, Wing S, Diez-Roux A, Poole C. "Neighborhood characteristics associated with the location of food stores and food service places." *American Journal of Preventative Medicine.* 2002 (22): 23-29.

Moro, Daniele and Paolo Sckokai (2000). "Heterogeneous preferences in household food consumption in Italy." *European Review of Agricultural Economics*, 27(3): 305-323.

Myers, James and Mark Alpert. 1968. "Determinant Buying Attitudes: Meaning and Measurement." *Journal of Marketing*, 32 (October): 13-20.

Nelson, Charles. Class Lecture Notes. Demand, Supply, Firms and Households. University of Illinois, Urbana-Champaign, IL. Spring 2013.

Nicholson, Walter and Christopher Snyder. 2008. *Microeconomic Theory: Basic Principles and Extensions*, 10th ed. Mason, OH: Thomson South-Western.

Pearson, Tim, Jean Russell, Michael J Campbell, and Margo E Barker. 2005. "Do 'food deserts' influence fruit and vegetable consumption? – a cross-sectional study." *Appetite*, 45: 195-197.

Raper, Kellie Curry, Maria Namakhoye Wanzala, and Rodolfo M. Nayga, Jr. 2002. "Food expenditures and household demographic composition in the US: a demand systems approach." *Applied Economics*, 34: 981-992.

Rose, Donald and Rickelle Richards. 2003. "Food store access and household fruit and vegetable use among participants in the US Food Stamp Program." *Public Health Nutrition,* 7: 1081-1088.

Tang, Christopher S, David R Bell, Teck-Hua Ho (2001), "Store choice and shopping behavior: How price format works," *California Management Review* 43(2): 56-74.

Teklu, Tesfaye and S.R. Johnson. 1986. "A Review of Consumer Demand Theory and Food Demand Studies in Indonesia." Staff Report 86-SR7. Center for Agricultural and Rural Development, Iowa State University.

Tiebout, Charles. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64: 416-424.

Tiffin, Richard, Kelvin Balcombe, Matthew Salois, and Ariane Kehlbacher. 2011. "Estimating Food and Drink Elasticities." University of Reading.

Weatherspoon, Dave, James Oehmke, Assa Dembele, Marcus Colema, Thasanee Satmanon, and Lorraine Weatherspoon. 2013. "Price and Expenditure Elasticities for Fresh Fruits in an Urban Food Desert." *Urban Studies*. 50(1): 88-106.

Wilde, Park, Craig Gundersen, Katherine Baylis & Joseph Llobrera. “National Food Study Geography Component”, Updated Study Plan, November 2012.

Wilde, Park, Paul McNamara, and Christine Ranney. 1999. “The Effect of Income and Food Programs on Dietary Quality: A Seemingly Unrelated Regression Analysis with Error Components.” *American Journal of Agricultural Economics*, 81: 959-71.

Wilde, Park, Paul McNamara, and Christine Ranney. 2000. *The Effect on Dietary Quality of Participation in the Food Stamp and WIC Programs*. Washington DC: US Department of Agriculture, ERS Food Assistance and Nutrition Research Report No. 9, September.

Wood, Ben, Carl Nelson and Lia Nogueira. “Food Price Crisis: Welfare Impact on Mexican Households.” Paper presented at the *International Agricultural Trade Research Consortium* Analytical Symposium June 22-23, 2009. Seattle Washington.

Wrigley, Neil, Daniel Warm, Barrie Margetts, and Amanda Whelan. 2002. “Assessing the Impact of Improved Retail Access on Diet in a ‘Food Desert’: A Preliminary Report.” *Urban Studies*, 39: 2061-2082.

Ver Ploeg, Michele, Vince Breneman, Tracey Farrigan, Karen Hamrick, David Hopkins, Phillip Kaufman, Biing-Hwan Lin, Mark Nord, Travis A. Smith, Ryan Williams, Kelly Kinnison, Carol Olander, Anita Singh, and Elizabeth Tukermanty. 2009. *Access to Affordable and Nutritious Food - Measuring and Understanding Food Deserts and Their Consequences: Report to Congress.* Washington DC: US Department of Agriculture, ERS AP-036, June.

You, Zhikang, James Epperson, and Chung Huang (2011). “Consumer Demand for Fresh Fruits and Vegetables in the United States (1960 – 1993)”. The University of Georgia Cooperative Extension RB 431.

Figure 1: Graphical illustration of consumer demand:
The effect of budget or price ratio change on consumption

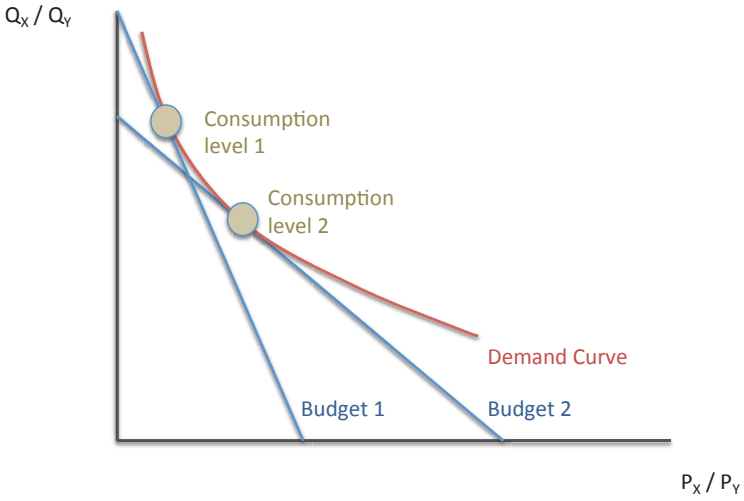
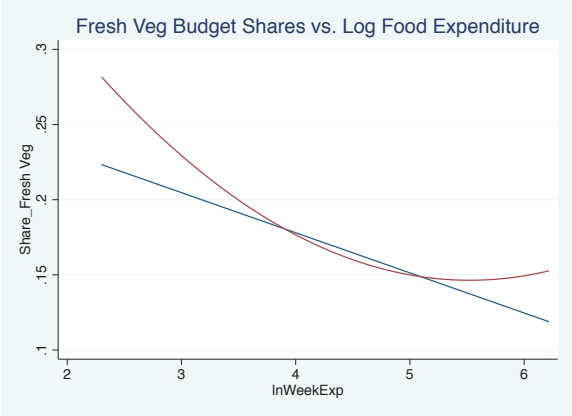
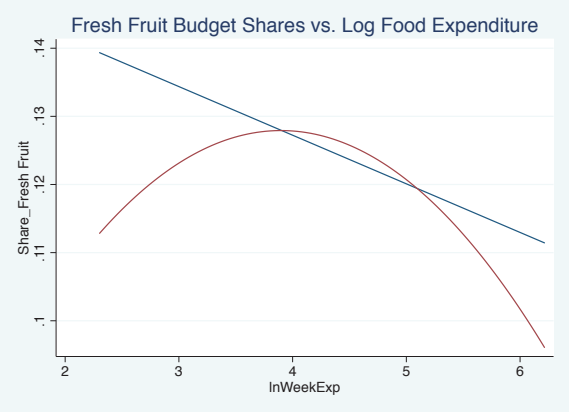
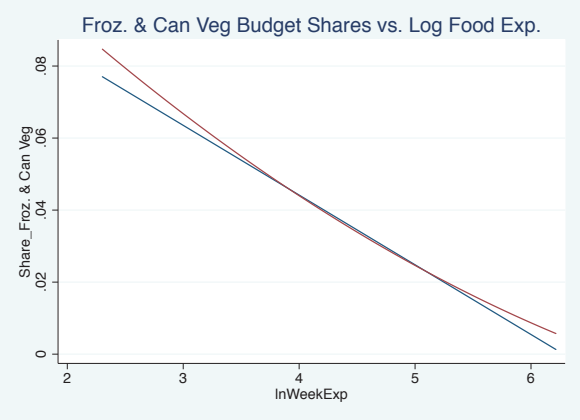
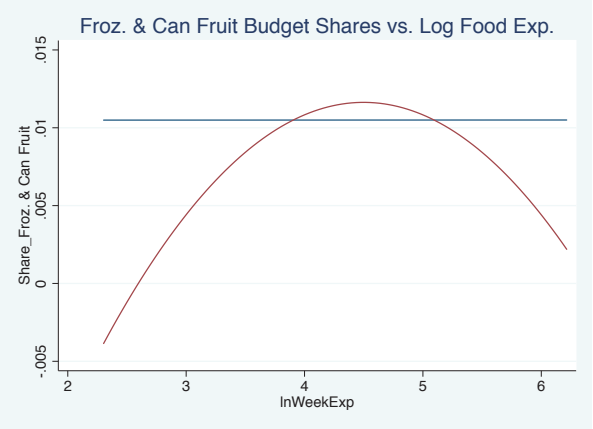


Figure 2: Linear and quadratic Engel Curves of food budget shares on log expenditure









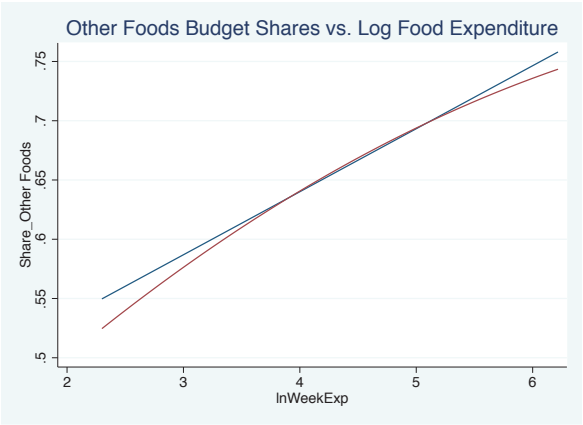


Figure 3: Tradeoff of timesavings vs. quantity

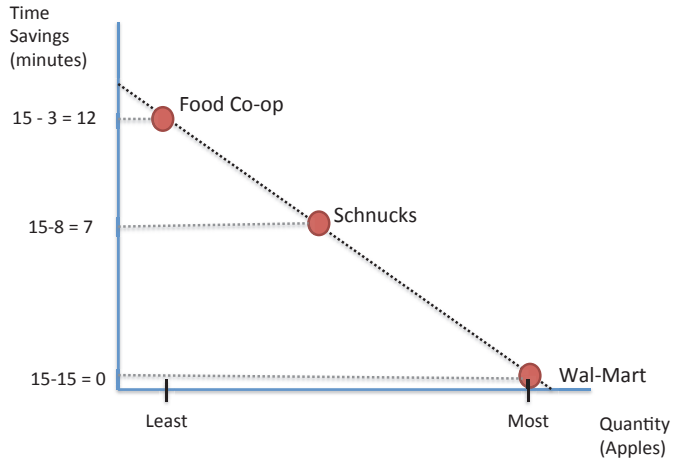


Figure 4: Tradeoff of quality vs. quantity

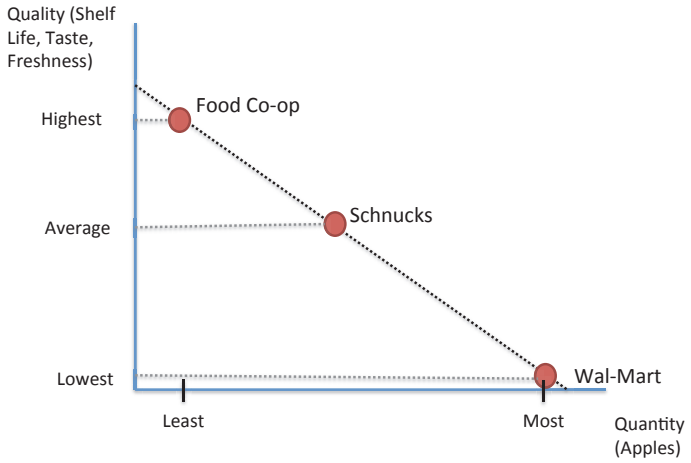


Figure 5: Tradeoff of quantity vs. store type

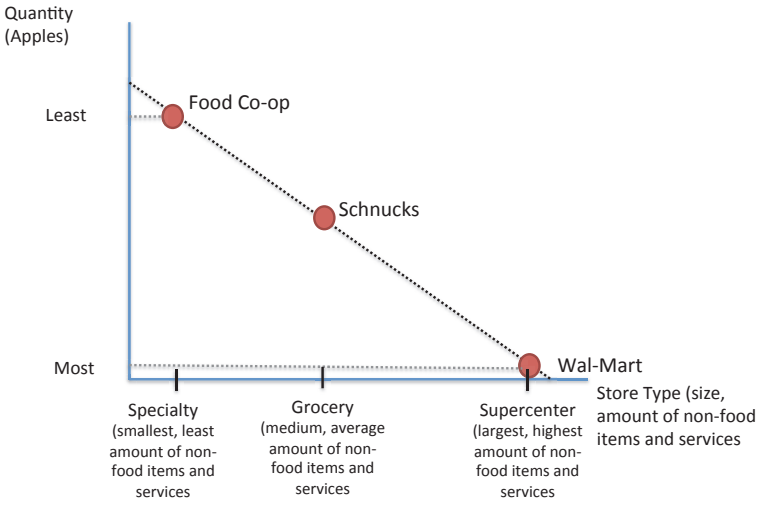
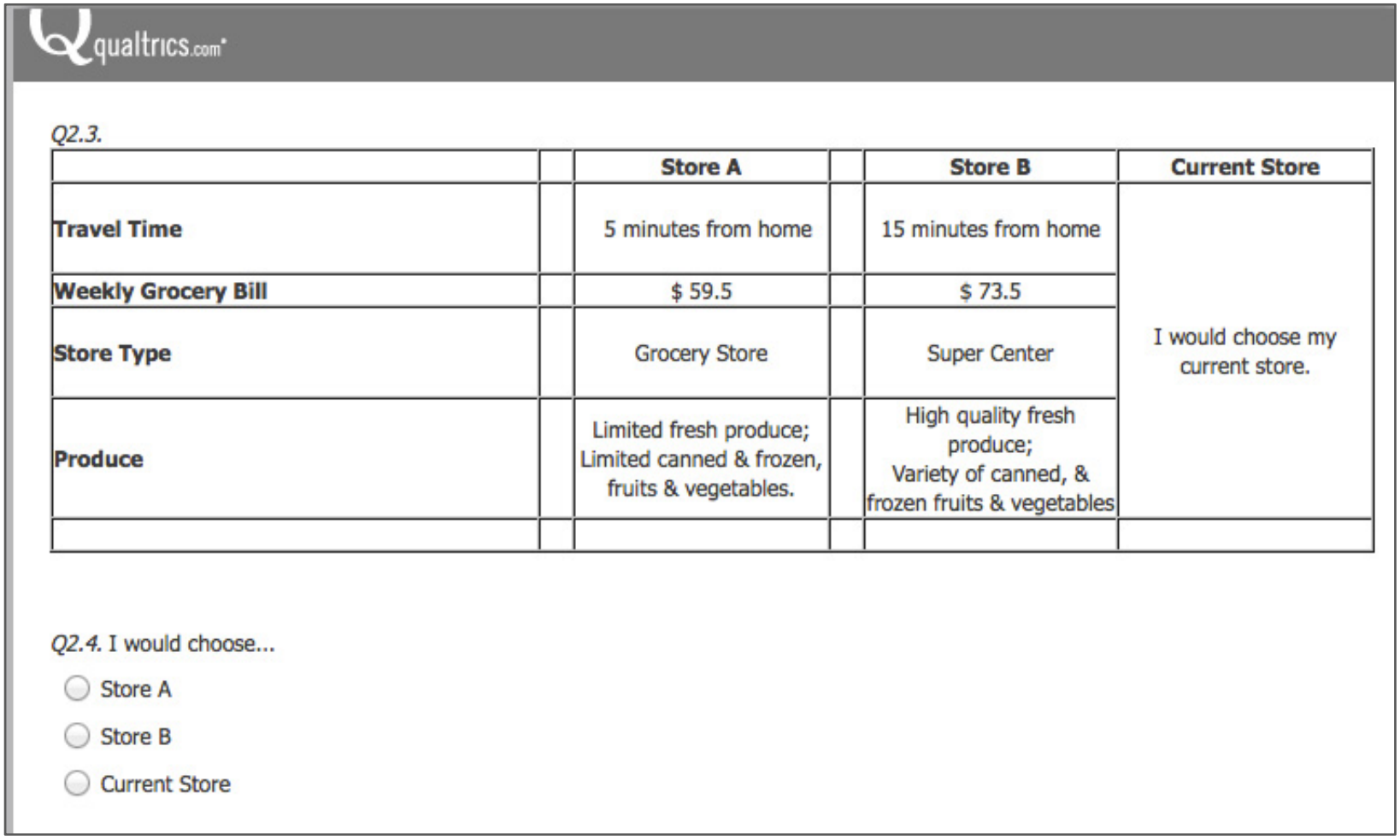


Figure 6: Choice Experiment attributes and levels

Attributes	Levels	Description	Examples
Store Type	Specialty Store	Small-sized store with few non-food items and no additional services (pharmacy, photo, florist, etc.)	• Strawberry Fields • World Harvest • corner store
	Grocery Store	Medium-sized store with an average amount of non-food items and few additional services	• Schnucks • County Market
	Supercenter	Large-sized store with many non-food items and many additional services	• Sam's Club • Wal-Mart • Meijer
Produce Quality & Variety	Limited	Limited Fresh Produce Limited canned or frozen fruits & vegetables	
	Average	Average quality fresh produce Some canned or frozen fruits & vegetables	
	High	High quality fresh produce Variety of canned or frozen fruits & vegetables	
Weekly Grocery Bill	85% 90% 95% 105% 110% 115%	Total grocery bill spent at the store for one week's worth of groceries presented in terms of percentage of respondent's current average weekly grocery bill	
Travel time	5 minutes 15 minutes	Total travel time required to get to the store using the mode of transportation most often used by the respondent	

Figure 7: Sample Choice Experiment question



qualtrics.com

Q2.3.

		Store A		Store B	Current Store
Travel Time		5 minutes from home		15 minutes from home	I would choose my current store.
Weekly Grocery Bill		$ 59.5		$ 73.5	
Store Type		Grocery Store		Super Center	
Produce		Limited fresh produce; Limited canned & frozen, fruits & vegetables.		High quality fresh produce; Variety of canned, & frozen fruits & vegetables	

Q2.4. I would choose...

- ○ Store A
- ○ Store B
- ○ Current Store

Figure 8: Modes of transport most often used by respondents to get to the grocery store

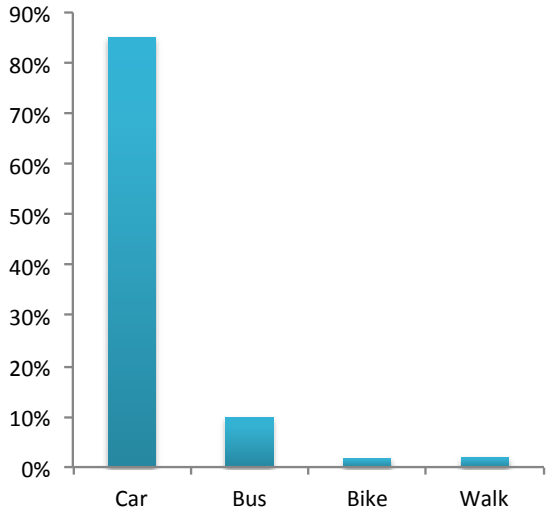


Figure 9: Sample residential distribution and food geography in Champaign & Piatt County, IL

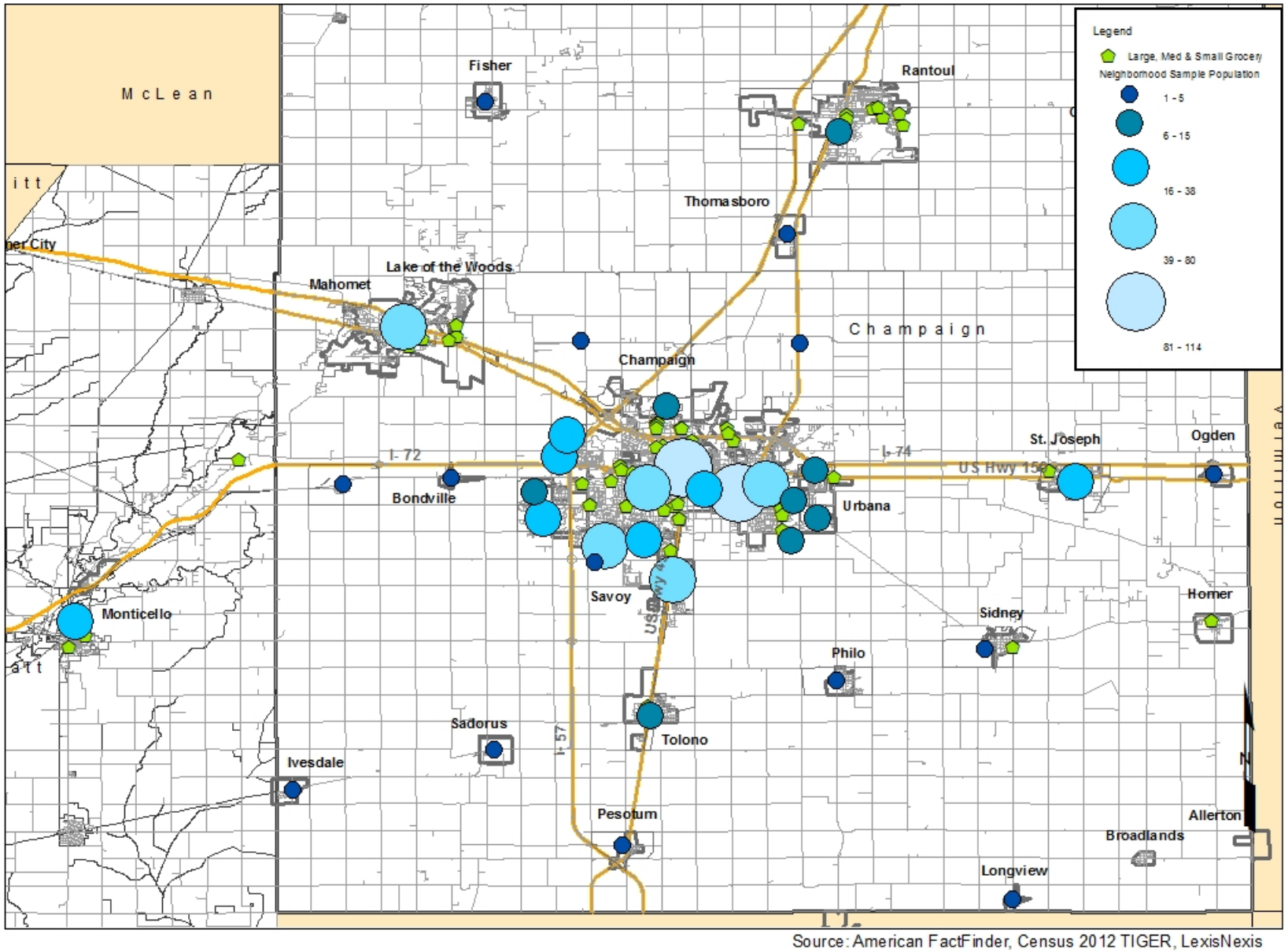


Figure 10: Distribution of geographic distance to nearest grocery store: comparison of convenience store shoppers

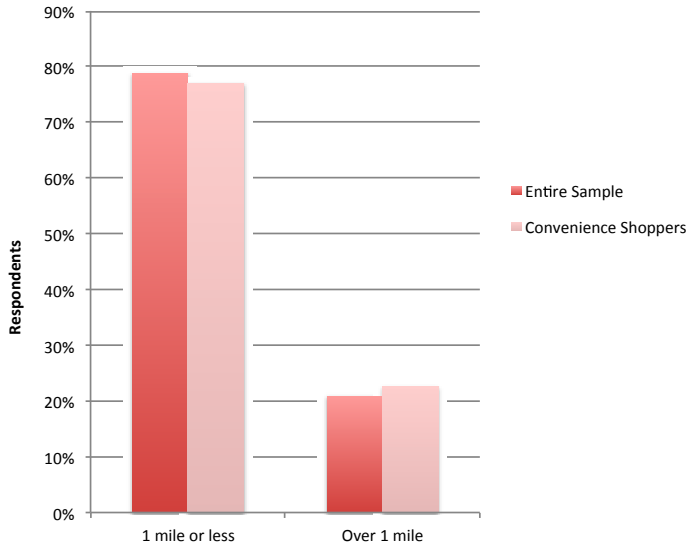


Figure 11: Distribution of stated travel time to principal grocery stores

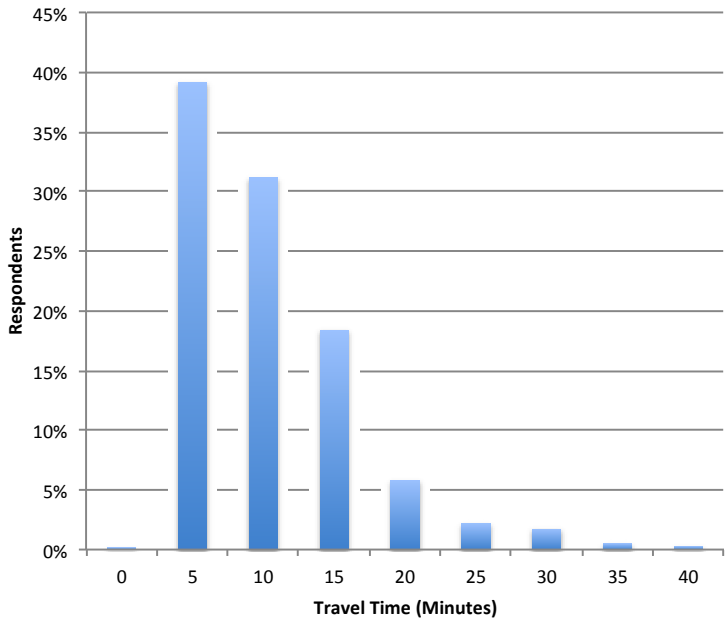
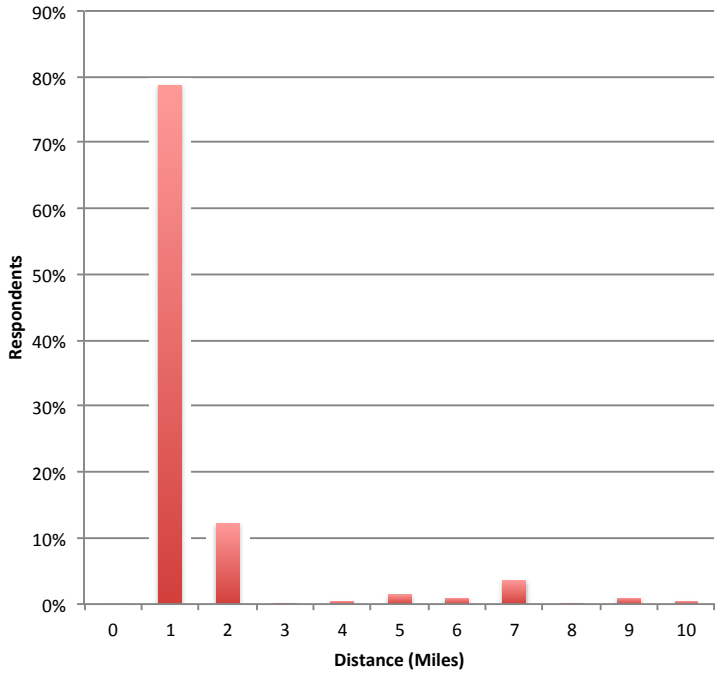


Figure 12: Distribution of geographic distance to nearest grocery store



Appendix A

Choice Experiment Design SAS macro

```
%mktruns (2 3 3 6)
%mktex (2 3 3 6, n=36)
%mktlab (data=design, int=f1-f2)
proc print; run;
%choiceff (data=final, model=class(x1-x4), nsets=18, maxiter=20,
flags=f1-f2, beta=zero);
proc print; by set; id set; run;
```

SAS Fractional Factorial Design Output

		Attribute Code				Attribute Description			
Profile	**Set**	**X1**	**X2**	**X3**	**X4**	**Time**	**Store**	**Produce**	**Price**
1	1	1	2	3	1	5 min	Grocery	Limited	- 15%
2	1	2	3	1	4	15 min	SuperCenter	HQ Fresh	+5%
3	2	1	3	2	1	5 min	SuperCenter	Avg Fresh	- 15%
4	2	2	2	1	6	15 min	Grocery	HQ Fresh	+15%
5	3	2	2	3	2	15 min	Grocery	Limited	-10%
6	3	1	3	1	3	5 min	SuperCenter	HQ Fresh	-5%
7	4	1	2	2	3	5 min	Grocery	Avg Fresh	-5%
8	4	2	3	3	5	15 min	SuperCenter	Limited	+10%
9	5	1	1	3	4	5 min	Specialty	Limited	+5%
10	5	2	3	2	1	15 min	SuperCenter	Avg Fresh	- 15%
11	6	2	3	1	3	15 min	SuperCenter	HQ Fresh	-5%
12	6	1	1	2	6	5 min	Specialty	Avg Fresh	+15%
13	7	1	2	3	2	5 min	Grocery	Limited	-10%
14	7	2	1	2	6	15 min	Specialty	Avg Fresh	+15%
15	8	1	2	2	4	5 min	Grocery	Avg Fresh	+5%
16	8	2	1	3	3	15 min	Specialty	Limited	-5%
17	9	2	3	2	2	15 min	SuperCenter	Avg Fresh	-10%
18	9	1	2	1	5	5 min	Grocery	HQ Fresh	+10%
19	10	1	3	3	6	5 min	SuperCenter	Limited	+15%
20	10	2	1	1	2	15 min	Specialty	HQ Fresh	-10%
21	11	1	3	1	4	5 min	SuperCenter	HQ Fresh	+5%
22	11	2	1	2	5	15 min	Specialty	Avg Fresh	+10%
23	12	2	2	1	5	15 min	Grocery	HQ Fresh	+10%
24	12	1	1	3	3	5 min	Specialty	Limited	-5%
25	13	1	1	1	1	5 min	Specialty	HQ Fresh	- 15%
26	13	2	2	2	3	15 min	Grocery	Avg Fresh	-5%
27	14	2	1	1	1	15 min	Specialty	HQ Fresh	- 15%
28	14	1	3	3	5	5 min	SuperCenter	Limited	+10%
29	15	1	1	2	5	5 min	Specialty	Avg Fresh	+10%
30	15	2	3	3	6	15 min	SuperCenter	Limited	+15%
31	16	2	1	3	4	15 min	Specialty	Limited	+5%
32	16	1	2	1	6	5 min	Grocery	HQ Fresh	+15%
33	17	1	3	2	2	5 min	SuperCenter	Avg Fresh	-10%
34	17	2	2	3	1	15 min	Grocery	Limited	- 15%
35	18	2	2	2	4	15 min	Grocery	Avg Fresh	+5%
36	18	1	1	1	2	5 min	Specialty	HQ Fresh	-10%

Appendix B: List of produce for which respondents provided consumption information

Vegetables	Fruits
Potato	Melon
Onion	Berries
Fresh Lettuce	Apples
Leafy Greens	Bananas
Pre-packaged Lettuce	Oranges
Tomato	Tangerines
Cucumber	Lemon
Bell Pepper	Lime
Broccolli	Grapes
Cauliflower	Stone Fruit
Celery	Clementine
Carrots	Pears
Green Beans	Kiwi
Squash	Pineapple
Zucchini	Grapefruit
Asparagus	Pomegranate
Beets	Frozen Fruit
Avocado	Canned Fruit
Snap Peas	Canned Pineapple
Mushrooms	
Sweet Potatoes	
Brussel Sprouts	
Leeks	
Green Onion	
Garlic	
Fennel	
Celeric	
Turnip	
Parsnip	
Herbs	
Ginger	
Olives	
Canned Vegetables	
Pasta Sauce	
Frozen Vegetables	