Essays on the Economics of Food Access and Food Security

by

Phillip M. Warsaw

A dissertation submitted in partial fulfillment of the requirements of

Doctor of Philosophy
(Economics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Date of final oral examination: 3/30/2017

The dissertation is approved by the following members of the Final Oral Committee:
Alan T. Sorenson, Professor, Economics
Kenneth Hendricks, Professor, Economics
Daniel J. Phaneuf, Professor, Agricultural and Applied Economics
Michael Bell, Professor, Community and Environmental Sociology
Alfonso Morales, Professor, Urban and Regional Planning

ACKNOWLEDEGMENTS

My time in the Department of Economics at the UW-Madison has been invaluable to my growth as an individual and a scholar, and I owe a deep debt of gratitude to many individuals for facilitating that development. I would like to thank my advisor, Alan Sorenson, for his constant encouragement and help in finding my academic voice. I would also like to thank Daniel Phaneuf, without whose extensive technical guidance and general support I could not have completed this project. In addition, I wish to thank the rest of my committee: Kenneth Hendricks, Michael Bell, and Alfonso Morales, each of whom provided valuable insights to my work at various stages of my writing process.

During my time at the University of Wisconsin-Madison, I have also had the opportunity to meet many amazing colleagues, many of whom played a vital role in shaping my work. In particular, many thanks to Laura Alex Frye-Levine and Amanda McMillen-Lequieu, who volunteered many hours to reading and providing feedback on my work.

Finally, I wish to thank my family. My wife, Alexandra Allweiss, provided both the intellectual stimulus for the creation and development of this project, and also the emotional support and love necessary to see it to its completion. My parents provided me with the critical foundations of hard work and determination, and always encouraged me to pursue my goals, for which I cannot express deeply enough my love and gratitude. Finally, to my siblings, I wish to thank them for their unconditional love and support. I owe much to the many people who have aided in the completion of this project, but any mistakes within are solely my own.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

This project aims to develop economic models and methodologies to investigate the presence of food insecurity in urban areas, and the impacts of policies designed to address it. The concept and impacts of food insecurity have become increasingly prominent within academia and applied policy work over the last several decades. The rise in interest around the concept followed robust work in public health linking environmental factors, including local food access, to health outcomes such as BMI and obesity rates in the 1990's. This same line of research has revealed consistent patterns of correlation between food insecurity and socioeconomic status, namely race and income. In particular, urban areas with a high concentration of poor households and households of color face disproportionately high levels of food insecurity, weaving this issue into a broader narrative of social and environmental injustice. This connection makes the topic of food access one of interest not just to academics and policymakers in public health, but also to a broader group of researchers and activists interested in social justice.

Despite, or perhaps because of the broad interdisciplinary engagement on this topic, many open questions remain within the literature. Of them, *defining* food insecurity and *determining its causes* are of particular importance. While an intuitive concept, determining a broadly appealing and operational definition of food access has proven difficult. Doing so requires both defining what counts as a food source (i.e. grocery stores vs. other outlets like farmer's markets), and what constitutes access (physical proximity vs. affordability of available options), both of which are open to many interpretations. Similarly, establishing the cause of food insecurity plays a foundational role in attempts to address it. Whether food insecurity is caused by the decisions of rational individual consumers in residential markets or by discriminatory structures and institutions remains a central debate for researchers, and one of significant importance.

It is here that this project enters the current discussion. Using fourteen years of publicly available land use and residential transaction data from the City of Milwaukee, WI, I develop and estimate a pair of econometric models to characterize demand for local food access within the city. In my first chapter, I use a hedonic pricing model to define food access using the marginal implicit prices paid by households for local grocery stores. I further define food insecurity as the presence of inequality in those prices paid across space and socioeconomic characteristics. I find evidence of potential food insecurity in the city, specifically that households in neighborhoods consisting only of African-American and/or Latino households pay premiums of over 2% of their home price for an additional grocery store within .75-1 mile of their home, compared to premiums near or even below zero percent for homes located in neighborhoods without any households of color.

In my second chapter I use a residential sorting model in order to estimate consumer preferences for local grocery stores in Milwaukee. The use of the residential sorting model allows for the estimation of heterogeneous preferences. Specifically, I allow preferences to vary across socioeconomic status, aiding to identify the cause of the existing distribution of households with respect to food access. I find that African-American households at the average observed income are willing to make a one-time payment of over $5000 for an additional grocery store within a mile of their home, compared to a willingness-to-pay of less than $2000 for white households at the same income level. I further find that African-American households at the poverty line have a willingness-to-pay for an additional grocery store that is roughly equivalent to a white household at the maximum observed income in Milwaukee. These results not only highlight the conflation of race and income in structural preference models, but also suggests that any existing food insecurity in the City of Milwaukee is not the result of consumer preferences, but instead other market dynamics or institutional barriers. Counterfactual policy analysis supports this notion by suggesting that policies aimed at increasing the number of food outlets in low-income neighborhoods and/or neighborhoods with a high proportion of households of color lead to significant gains in welfare, despite resulting increases in average home prices. This contrasts with policies aimed to do the same in areas with higher income and/or white households, where the overall welfare impact is greatly diminished in comparison.

Together, the two chapters provide a complementary characterization of the demand for food access in Milwaukee. The hedonic model of the first chapter estimates equilibrium prices arising from a supply and demand framework in residential markets. However, in the context of this study, it does not allow for a robust investigation of the relative roles of supply and demand in determining the existing equilibrium, as the hedonic model does not allow for the identification of a unique demand curve. However, the horizontal sorting model, which relies on the use of a locational Nash equilibrium, allows for the identification of consumer preferences. This allows for a more thorough discussion of the demand side dynamics potentially influencing the hedonic model. The higher marginal willingness to pay for an additional grocery store in their neighborhood by African-American households explains, in part, the higher marginal implicit prices paid for homes neighborhoods with a high proportion of African-American and Latino households which are also in close proximity to local grocery stores. At the same time, the inequalities discovered through the use of the hedonic model suggest that these locations might also be sites for potential policy intervention, the effects of which can be evaluated through the locational sorting model.

It is important to note that the work presented here does not preclude the potential role of supply side dynamics in the existing food landscape. However, taken together, this project provides a compelling argument for both the need and potential value of interventions aimed to provide greater levels of food access in Milwaukee neighborhoods with a high proportion of low income households and households of color. In addition, this research further highlights the synergy between hedonic and residential sorting models in using housing markets to identify preferences for local public goods, and evaluating policies which alter the landscape within a given market.

CHAPTER 1

The Economics of Food Security: Investigating Price Inequalities for Food Access in Milwaukee, Wisconsin

ABSTRACT

One of the primary debates ongoing in the literature on food security is how to define the term in such a way that it is both operational and generalizable. The lack of such a concrete definition has led some scholars to question whether or not food insecurity truly represents a broad and systemic challenge to public health. In this essay, I enter this debate by deriving a definition of food security based on the implicit prices paid by households in residential markets for food access. Using the theory of the hedonic pricing model, I define food insecurity as the presence of unequal marginal implicit prices paid across space and/or socioeconomic characteristics (SES) for access to local food sources. I argue this definition provides a more robust characterization of the "functional" food access households have, as opposed to the standard definitions of food access which consider only the proximity of local food sources. Using publicly available land use and residential transaction data from the City of Milwaukee, I employ this definition to investigate potential inequalities in access to large, full service grocery stores. Using interaction terms between neighborhood racial characteristics and local groceries, I find evidence of potential food insecurity in neighborhoods with a high proportion of African-American and Latino households.

1.1 Introduction

Over the last two decades, food security has become an increasingly prominent research topic, both within academic and policy circles. In particular, inequalities in access to food, especially fresh food, within low socioeconomic status (SES) communities have been of interest to scholars, especially in urban spaces. While an important topic of research, defining food access in a way to operationalize investigations into food security has proven challenging, inducing a large and interdisciplinary literature seeking to reconcile these challenges. Researchers have also devoted significant time to characterizing a wide variety of solutions, such as increasing the number of retail options available to households, and more recently direct production by consumers in the form of community and urban gardens.

A recurring concept in these discussions is the issue of *scarcity*, with food insecurity representing a relative lack of healthy food options in one neighborhood compared to another. As such, economists are well positioned to contribute to this ongoing discussion. Residential home markets in particular provide a useful context for studying food access. While households may value living in close proximity to quality food sources, they cannot directly purchase food access. However, if food access is an amenity to households, local home prices should reflect their value in equilibrium, as agents seek homes with more food options in their vicinity.

Research over the last decade in urban economics has further considered how the impact of local amenities on local home prices vary across space and household characteristics, allowing for direct estimation of potential inequalities in the premiums paid by households in different locations across the city. Overlaying socioeconomic characteristics on the geographical landscape then allows for a simultaneous consideration of both spatial and socioeconomic variance in residential markets and their roles in influencing local prices for particular amenities.

Such a consideration would allow for a deeper investigation of food access beyond most existing studies, which focus on the proximity of local food sources. Rather, this robust form of analysis instead derives the “functional” access households face given their own budget constraints. Despite that, very little work has been done within economics investigating food access as an amenity to households.

In this paper, I address this gap in the economics literature by developing and estimating a model to characterize food access in the city of Milwaukee, Wisconsin. Using a residential supply and demand framework, I argue that in the presence of food insecurity, the marginal implicit price paid for an additional grocery store will vary across space and neighborhood SES. I then operationalize this definition by estimating a hedonic pricing model to determine whether large, full-service grocery stores are capitalized into home prices, and if this capitalization varies across space and socioeconomic characteristics.

To implement this model, I generate a unique dataset from publicly available sources characterizing home sale transactions and land-use characteristics over a period of fourteen years in Milwaukee. Detailed master property records combined with GIS parcel maps allow me to pinpoint the location of grocery stores along with other retail locations in relation to each home sale. This allows for a robust characterization of food access for homes, while also controlling for other retail options that could potentially bias the results. Further, a characterization of neighborhoods in Milwaukee, based on demographic and geographical features, in addition to high school district boundaries provide a natural set of spatial controls, which I use as fixed effects to control for the possibility of spatial heterogeneity. The close mapping of the neighborhoods to census tracts within the city provides a natural geography for testing inequalities in marginal prices paid for additional food outlets. Specifically, I interact racial characteristics at a census tract level with the number of grocery stores within a given proximity to each home transaction, after determining the average effect of an additional grocery stores on home prices throughout the city.

My results show that grocery store access is capitalized into home prices, which, after controlling for commercial access, suggests that households consider local food access an amenity separate from other forms of retail access. After controlling for spatial heterogeneity, I find that an additional grocery store between .25 – 1 miles of a home increases its sale price by 0.55 – 1.82 percent on average. Additionally, my results indicate that households in neighborhoods with a higher population of African-Americans and Latinos pay a higher marginal implicit price for food access, up to 2.39 percent of home prices in neighborhoods only contain African-American and/or Latino households, suggesting potential inequalities in food access across the city. These are consistent with much of the existing literature on food security, with the additional benefit of providing a stronger characterization of the functional food access available to households than many existing studies. These results indicate the promise of using this framework to investigate food security in other contexts, particularly those with a richer characterization of food sources. My results also highlight the importance of controlling for access to heterogeneous retail properties in addition to other land use types in models characterizing residential markets.

The rest of this paper is structured as follows. Section two provides a brief literature review of work on food access and estimating consumer preferences for housing. Section three outlines the theoretical and empirical models used to characterize and estimate food access. Section four describes the data sources used and the process to create the dataset for empirical testing. Section five presents the results and section six concludes.

1.2 Literature Review

Academic interest in food access date back to the 1960's, though interest in the topic surged in the 1990's. One of the most common concepts used in the food access literature is the "food desert", used to describe an area of low food access. Typically, food deserts are defined in absolute terms[1], though no common definition is used across the literature (Hendrickson et al. 2006). In general, studies characterize food deserts by one or more of four factors: accessibility, variety, quality, and affordability. Scholars have most frequently focused on accessibility, typically leveraging GIS data to quantify the number and relative proximity of grocery stores to the average household in a given neighborhood, and the resulting travel time to the nearest food source (see, for instance, Block et al 2004 or Moorland et al 2002). However, the lack of a common metric across the literature, even among studies using this strict proximity definition of food access, and the resulting variance in results as to the presence, severity, and consequences of food deserts, has caused some to question whether food deserts actually exist (Hendrickson et al 2006). The term has also been subject to much critique and debate, with some scholars, particularly food justice advocates, arguing that the term places the blame of low food access on those harmed by low food access[2]. As such, while I draw from the literature on food deserts, I use the term "food (in)security" moving forward in defining inadequacies and inequalities in food access for this study, as it lacks these troubling connotations, and more easily accommodates the comparative analysis used here.

While fewer in number, several researchers have noted the importance variety, quality, and affordability as factors which fold into accessibility. The simultaneous consideration of these

---

[1]http://www.ers.usda.gov/dataFiles/Food_Access_Research_Atlas/Download_the_Data/Archived_Version/archived_documentation.pdf

[2]http://www.wellhousegr.org/wordpress/wp-content/uploads/2014/05/What-is-Food-Justice.pdf

factors generates the "functional" food access that households in a given area have, extending beyond simple physical proximity of food sources. Even if a neighborhood has many grocery stores, if they sell products that are too expensive (Hendrickson et al 2006), or are of poor quality (Zenk et al 2006), they provide limited benefit to local residents. Within this work, scholars have posed a wider variety of questions, and employed a broader set tools to investigate food access. These range from direct comparisons of food prices across neighborhoods (Morris et al 1990, 1992), to surveys and questionnaires assessing resident beliefs within neighborhoods regarding the quality and variety of the food provided in nearby outlets (Hendrickson et al 2006, Block and Kouba 2006).

Despite the variance in methodologies and definitions used to characterize food security, a point of common inquiry across the literature is the connection between socioeconomic status and food access. Specifically, regardless of the specific metric used, many scholars and advocates have investigated whether poor households and households of color face lower food access than their counterparts. Almost universally, researchers agree that to the extent food insecurity exists, it has a strong correlation with socioeconomic inequality. Studies such as those performed by Chung and Myers (1999) and Powell et al (2007) indicate that low SES households have fewer grocery stores in their vicinity, and thus typically have to travel further to reach a food outlet (Alwitt and Donley 1997, Zenk et al 2005). Once they arrive at a storefront, these households typically face fewer product options (Morris et al 1992) than their better-off counterparts, and ultimately pay more for food that is often of a lower quality (Zenk et al 2006), a result borne out of both statistical analysis and household assessment (Hendrickson et al 2006).

To date, almost no studies exist within economics to directly investigate the dynamics of food access outlined above. Rather, most studies focus on understanding consumer demand for nutrition as it relates to consumer access for food (see, for instance, Handbury et al 2016). The results of these studies are mixed on whether additional food access significantly impacts consumption of “healthy” foods. While the existing studies provide critical insight to the relationship between access and health outcomes, they also rely on a narrow construction of the benefits of additional food access, built on a specific definition of “healthy foods”. By focusing on access to food as an amenity itself, it is possible to capture benefits that go beyond this narrow scope, and as such provide a useful complement to the demand for nutrition analysis typically seen in economic analyses. These studies, however, appear very infrequently within the economics literature. One such study was performed in 2008, in an investigation of the impact of community gardens on residential values in New York City (Voicu and Been 2008). Using a difference-in-difference specification of the hedonic regression, they find that the arrival of community gardens leads to an increase in local residential property values, particularly those within a 1000-foot radius of a given garden. These effects were found to be larger in poorer neighborhoods, providing evidence of the demand and benefits for additional local food sources in these neighborhoods. A study by Sheutz et al (2012) also explicitly discusses and explores the existence of food insecurity. In it they investigate the relationship between household income and retail density of several types (including food outlets) in 58 large US metropolitan areas, using employment statistics for various retail segments in a pooled sample from 1992-2006 as their measure of access. They find that high poverty neighborhoods have a lower density of every retail option considered, including grocery stores.

While limited attention has been given to demand for food access as an amenity in economics, a wide class of models have used to examine residential markets and estimate the demand for local urban and environmental amenities, and how demand varies across space and socioeconomic characteristics. Within urban economics, studies have often focused on estimating demand for local public goods, such as school quality and the socioeconomic makeup of one's neighborhood. For example, Bayer and McMillan (2012) investigate these questions in the San Francisco housing market. Other scholars have focused on the disamentity value of public "bads" in residential markets, typically those associated with crime (Linden and Rockoff 2008, Pope 2008). In addition, many scholars both within urban and environmental economics have investigated the amenity value of local desirable and undesirable land use. Kuethe (2012), for instance, estimated the implicit prices for neighborhood-level spatial fragmentation, the rate at which land-use type changes, and land-use diversity in Milwaukee, finding a U-shaped price curve for the former. Other studies, such as by Taylor et al (2016), Gamper-Rabindran and Timmins (2008), and Noonan et al (2007) focus on specific urban land-uses, most often hazardous waste sites and their negative impacts on home prices.

Methodologically, many of these models all belong to a much wider class of hedonic pricing models which have been used to explain consumer preferences and choices in housing markets. The hedonic model allows for an empirically straightforward estimation of marginal implicit prices paid for features in the housing market. They also require fewer theoretical assumptions than other models used to study residential markets, such as residential sorting models, and have minimal data requirements. This makes it an ideal fit for this study, which focuses on inequalities in prices paid for food access at a neighborhood level. Recent advances in this literature have focused on addressing omitted variable bias in the hedonic regression, particularly in the unobservable spatial heterogeneity arising from features that vary across space. Quasi-experimental approaches which exploit natural breaks in the data (see, for instance, Linden and Rockoff 2008 or Pope and Pope 2012) have emerged as an ideal way to address omitted variable bias, but are not often possible given their larger data requirements, or desirable due to their often very localized contexts (Bajari et al 2012). In this case, the use of spatial fixed effects has become standard to address unobserved local factors which may vary across space. Recent studies have gone beyond the use of spatial fixed effects in isolations, interacting them with time fixed effects, or even other covariates (such as prices within a neighborhood from the previous year), in order to additionally control for variables which vary across time and space (Linn 2012, Kuminoff et al 2010 Anderson and West 2006). I follow this approach in this study, as I lack the observation of repeated sales necessary to a quasi-experimental approach.

This paper contributes to the literature in two ways. First, I contribute to a limited literature investigating the amenity value of heterogeneous commercial land-use parcels in a housing market. While a growing segment of the urban and environmental economic literature has leveraged high-resolution GIS land-use data to investigate the amenity value of heterogeneous land-use types, the focus has largely been on investigating environmental amenities, particularly open space (see, for instance, Cho et al 2010 or Klaiber and Phaneuf 2010). Less attention has been given to the value of urban commercial and retail locations, with almost no attention given to the amenities provided by access to heterogeneous retail options. Using the Milwaukee's publically available yearly master property record, I am able to identify a rich selection of retail land-use categories, including grocery stores, which I use to define food access. Joining these records to GIS maps characterizing each parcel throughout the city allows me to characterize the location of each store in relation to home sale transactions, and define food access at several distances.

Second, I contribute to the literature on food access and food insecurity. Both within and outside of economics, this represents the first attempt to use dynamics in the housing market to investigate local food access as an amenity to households and the presence of food insecurity. By considering how grocery stores at various proximities are capitalized into home prices, I am able to provide a unique perspective on the affordability of local food sources, providing one way of measuring "functional" food access. In addition, the flexibility of the hedonic model allows for the possibility of future consideration of the role of grocery store quality or variety in affecting home prices as well. In addition, the output of the hedonic model, a single price for food access, is easily communicated across disciplines and contexts, giving it an advantage over many other studies investigating food access which rely on more complex tools to characterize "functional" food access.

1.3 Model and Estimation

In the traditional hedonic model, households have preferences over homes which are a function of its attributes. The homes are differentiated, varying in their structural and locational characteristics. The focus of this study is on food access in the form of large, full-service grocery stores. As such, I characterize the household utility function for housing as:

$$U^k = U(X_i, N_j, Groc_i, c, \alpha^k) \qquad (1)$$

where households are indexed by k, and X_i represents structural characteristics of individual homes, such as the lot size or number of bedrooms, indexed by i. Neighborhood level attributes are represented by N_j and indexed by j. These include the socioeconomic make-up of the neighborhood a home is located in, which consumers have preferences over. Food access is defined at the individual home level, and represented by $Groc_i$. The numeraire is denoted by c, and I allow for heterogeneous preferences across households with α^k. Given the prices of the available homes, households choose a house that maximizes their utility subject to their budget constraint.

While the supply of homes in hedonic studies are typically assumed to be exogenous, both the neighborhood and food access characteristics of homes, which households have preferences over, are the result of decisions made by both households and firms. Retail firms, namely grocery stores in this study, make decisions of where to locate within the city. Typically, the hedonic model frames the firm decision as one of minimizing labor costs in the production of a single composite good. However, this paper explicitly focuses on household preferences and consumption decisions of local retail options as local amenities. As such, the decision process more closely follows Hotelling's (1929) model and its variants of firm location choice. I do not provide a comprehensive overview of the theoretical models here, but the approach and intuition closely follows that of empirical studies such as Scheutz et al (2012), Zenk et al (2005), and Alwitt and Donley (1997). In particular, Scheutz et al find that firms are less likely to locate in areas with lower median income levels and higher crime rates, indicating that firms consider potential demand in location decisions. Zenk et al and Alwitt and Donley show that firms may also consider the racial composition of potential markets, perhaps as a proxy for demand, indicating the potential for discrimination in firm location decisions. Together, these results suggest that firm locations are not exogenous to household location decisions. Rather, as firms observe household decisions, they choose a location based on the socioeconomic landscape to maximize expected profit. Similarly, households, having demand for retail access and observing firm locations, will choose locations to maximize their utility based on these endogenous firm decisions.

For households and firms, the available housing/land stock is exogenous, along with a set of structural and environmental features within the city, such as the proximity to the city center. As previously stated, when firms and households choose their locations, it shapes the demographic and retail landscape of neighborhoods across the city[3], forming the supply of locational amenities available at any given house. This landscape then influences the decisions of marginal households and firms as they enter or relocate within the market, who take the current composition of the city as exogenous. The preferences of these households generates the demand for various house-types throughout the city. Following Rosen (1974), the interaction between supply and demand in the housing market for these local amenities leads to an equilibrium price for each, along with an equilibrium socioeconomic composition of neighborhoods and firm locations. Home prices can then be expressed as a function of their characteristics:

$$P_{ij} = P(X_i, N_j, Groc_i) \qquad (2)$$

Empirically, equation **(2)** is estimated via regression analysis. In investigating the impact of food access on home prices, I first determine whether or not grocery stores provide an amenity value. If so, then the coefficient for food access in equation two should be positive. To test this hypothesis, I use a simple semi-log specification to estimate the hedonic pricing equation:

$$\ln(P_{ij}) = \beta' X_i + \gamma' N_j + \omega' Groc_i + \varepsilon_i \qquad (3)$$

The semi-log model has the advantage of easier interpretation, as estimated coefficients represent percentage changes to the conditional mean home price. Previous research has also indicated that it performs better than other specifications when the model is misspecified (Cropper et al. 1988). The use of a log transformation also reduces the issue of homoscedasticity caused by the highly skewed price variable (Ottensmann et al 2008).

A challenge in the estimation of equation **(3)** is the endogeneity of both socioeconomic characteristics and retail locations. Given that retail stores choose locations based on the

[3] Typically, economists assume that firms and households choose their location freely. However, in reality housing stocks are limited and firms face restrictions in where they operate, both from local zoning regulations. That said, this assumption is reasonable when considering the construction of the "initial" landscape of the city by firms and households.

perceived demand of local households, and households choose their locations based on the desirability of the neighborhood as a whole, the possibility exists that grocery stores could serve as a proxy for the desirability of neighborhoods based on unobserved factors. A similar argument can be made for the socioeconomic composition of neighborhoods as well. To address this, I use spatial fixed effects to control for unobserved features of each neighborhood that might influence mean home prices, and thus the likelihood of a grocery store or other commercial storefronts locating there. Given the possibility of race affecting firm location decisions, I also use the racial composition of each neighborhood as a control in the estimation of implicit prices for commercial access.

Estimating equation **(3)** allows for the identification of the average effect of amenities on home prices. However, these marginal implicit prices need not be constant across a given market. Unlike most product markets, the nature of the housing market prevents arbitrage, which would otherwise guarantee constant prices. Using grocery stores as an example, purchasing two homes which both have one grocery store within one half mile is not equivalent to purchasing one home with two grocery stores within a mile, all else equal. As such, equilibrium prices for amenities will often vary across space, both as a result of heterogeneity in the physical landscape and in the characteristics of the households sorting into a given area in equilibrium. This heterogeneity provides the basis for testing the second hypothesis of this paper. I posit that *food insecurity* is characterized, in part, by higher marginal implicit prices for food access in a given neighborhoods than in other areas within the same residential market. Drawing from the description of the housing market provided here, if we consider two otherwise identical neighborhoods, and observe that it is more costly to purchase a home with an additional local grocery store in one neighborhood, then functionally that neighborhood has a lower level of food access, even if the number of grocery stores is equal in both neighborhoods. If we further observe that higher prices appear in neighborhoods as a function of their socioeconomic characteristics, this suggests some level of social inequality, particularly if high prices are paid in low-SES neighborhoods. According to the existing literature on food insecurity, there is evidence to suggest that higher marginal prices for food access may be paid in neighborhoods with a higher share of households of color, especially if they are low-income. Thus, I rewrite equation **(3)** to allow for this possibility:

$$\ln\left(P_{ij}\right) = \beta' X_i + f(N_j, Groc_i) + \varepsilon_i \qquad (4)$$

Unlike in equation **(3)**, I no longer specify a functional form for the impact of neighborhood characteristics and grocery stores on home prices. As such, I can interact the socioeconomic characteristics of a given neighborhood with the food access of homes in that neighborhood, providing an empirical test of the existence of inequity in food access. If neighborhoods with a high proportion of minority and low-income households do in fact face lower access to food, this should be reflected in higher marginal implicit prices for food access, estimated as positive coefficients for interaction terms between food access and neighborhood-level demographics.

1.4 Study Area and Data

1.4.1 Study Area

My study area is the City of Milwaukee. At 96.8 square miles, Milwaukee is the largest city in the state of Wisconsin. The city is comprised of 190 unofficial neighborhoods, as defined by the Milwaukee Neighborhood Identification Project. In their assessment, they distinguish neighborhoods based on by physical features, housing subdivisions, community opinion, and major street locations, among other characteristics[4]. Figure 1.1 illustrates the neighborhood demarcated by this project. While few academic projects have studied Milwaukee in the context of food access, the city makes for an interesting case study, in part due to its similarities to other urban, industrialized cities in the Midwest (McMillen 2001). These include cities such as Chicago and Detroit, each of which have been prominent in the discussion of food access in urban contexts (for instance, see Mari Gallagher Consulting Group 2006 or Zenk et al 2005). Much like those cities, the Milwaukee exhibits fairly significant racial segregation[5]. The center and northwest portions of the city have a significant majority of African-American households, the near southern portion of the city has a slight majority of Latino households, while the eastern segment of the city, positioned right next to Lake Michigan, has a mainly young, professional white population. Outlying cities, such as Whitefish Bay, have white populations making up over 90% of the population, mostly consisting of older families. Figure 1.2 illustrates the distribution of African-American and Latino households in the city. This unique socioeconomic division suggests that the residential market in Milwaukee is distinct from those of the surrounding locales. Given this, I limit this study to only include sales within the city of Milwaukee proper, without concern of potential bias from estimating equilibrium prices while not observing sales in a significant portion of the market.

Milwaukee also makes an appealing city to study due to its abundance of publicly available data. Estimating a hedonic model requires data characterizing home sale transactions, including the home's structural features, neighborhood characteristics, including the location of

[4] A map of the neighborhoods, along with the criteria used for characterizing them, can be found here: http://www.milwaukee.gov/ImageLibrary/Public/map4.pdf

[5] Milwaukee, Detroit and Chicago have all been ranked in the top five cities for racial segregation for the last ten years, with Milwaukee ranked number one for a majority of that period https://fivethirtyeight.com/features/the-most-diverse-cities-are-often-the-most-segregated/

relevant amenities and it socioeconomic makeup. The city provides tabular and GIS data characterizing the landscape for free to the public. This makes the results found here fairly easy to replicate, or extend in future projects.

1.4.2 Residential Transactions

Milwaukee's Assessor office provides data detailing land sale transactions dating back to 2002. These transactions include single-family homes, as well as condominiums, vacant lots, and commercial properties. In addition to the sale year and price, the data provides information on basic structural characteristics of each lot purchased, including the number of bedrooms and bathrooms, lot and home size, and home style[6]. For this study, I consider sales in the years 2002-2006, and 2011-2015. The hedonic pricing model relies on the assumption of equilibrium in the market in question, an assumption I argue is inaccurate during the recession years of 2007-2010. Figure 1.3 illustrates the number of home sales from 2002-2015, and figure 1.4 shows the average real home price in 2002 dollars over the same period. Residential sales were fairly stable from 2002-2006, but fall sharply from 2007-2010, from 4798 home sales (after removing unreasonable records), to 815 in 2010. After 2010, home sales begin to rise and do so steadily through the observed end of the study period. Real home prices experience a similar fall during the recession years, from a high of $116,264.12 in 2006 to $100,411.11 in 2011. Unlike home sales, though, home prices continue to fall into the 2010's. They do stabilize in the final five years of the study period, averaging $95,253, which provides additional evidence of stabilization in the housing market during this time. Together, this provides compelling evidence that the housing market was out of equilibrium during the Great Recession years, violating a key assumption in estimating a hedonic model. However, as a robustness check, I also run the same

[6] http://assessments.milwaukee.gov/

models with the full fourteen year panel of sales data. After removing incomplete or erroneous records, I also trim 2.5% of the records based on price from both tails of the distribution, and end up with 26,912 records used in this analysis. Table 1.1 reports summary statistics for these sales[7].

1.4.3 Land-Use Data

To characterize the local amenities available at each home, including food access, I leverage Milwaukee's Master Property Record (MPROP) and parcel shapefiles, which provide spatial and tabular data describing every parcel in the City of Milwaukee dating back to 2001[8]. These records provide structural and ownership information for each of the parcels in the city. The city also provides a parcel shapefile, which illustrates each parcel in a city-wide map. Each parcel has a unique taxkey, contained in both the parcel shapefiles and MPROP. This allows me to match each parcel's structural characteristics with its location within the city. The same taxkey is also present with the home sales transaction data, enabling me to match these transactions to their parcels, isolate these transactions on the map, and then construct variables describing their surrounding features in ArcGIS.

The structural features found in the MPROP contain both physical characteristics of the buildings themselves, including additional variables not found in the land-sale transactions data such as the presence of an attic or air conditioning, and land-use characteristics of the parcel as a whole. The MPROP contains two distinct land-use codes, which provide a fairly robust categorization of the city's parcels. The first of these is a four-digit code based on the Standard Industrial Classification code (SIC), which identifies they type of activity done on a given parcel. Among these codes is a code for grocery stores, which I use as the basis for defining food access.

---

[7] Given the relatively large number of neighborhoods in Milwaukee, I do not include the summary statistics for those dummy variables in the table

[8] http://city.milwaukee.gov/mapmilwaukee

A second land use variable separates the parcels into eight larger catch-all categories, such as commercial, residential, and manufacturing, which I use to construct of broader land-use variables used in this analysis.

1.4.4 Defining Food Access and other Neighborhood Amenities

To define food access, I draw from the existing literature, which typically focus on large, full-service[9] grocery stores. While many stores, including convenience stores, offer food products, full-service grocery stores generally have lower prices (Kaufman et al 1997) and a wider variety (Glanz et al 2007), including selections of fresh meat and produce often not found in any single convenience store location. From this position, two issues present themselves: defining *large*, full-service grocery stores in the context of Milwaukee, and deriving a metric of food access around this food source. To begin, I sorted all of the locations coded as grocery stores present in the MPROP by their lot size for the year 2005. In examining the locations, a clear break in grocery store size occurred around 10000 square feet, with the largest grocery store across the ten years under that size reporting at 8916 square feet, and the next largest reporting at 10406 square feet[10]. I then cross-checked the addresses provided by the MPROP for each of these stores with Google results, listings in the yellow pages, and store websites when available. This confirmed that stores above this size consisted of only grocery stores which met the full-service criteria laid out by the District of Columbia (such as Whole Foods or Pick N' Save), with stores below this cutoff consisting solely of convenience stores (of a similar style to Seven-Eleven) and a small number of corner stores, none of which met the full-service criteria.

---

[9] I use the rubric provided by the District of Columbia in defining a food desert: http://abra.dc.gov/page/full-service-grocery-stores

[10] Two stores were listed at approximately 10000 square feet across the study time period, but both were Family Dollars, which were not considered for this study

As a result, I used 10000 square feet as the cutoff between "large" grocery stores included in this analysis, and smaller stores excluded from the analysis.

In addition to these stores, I also include a number of locations that were not coded as grocery stores in the MPROP in my definition of large, full-service grocery stores. Most notably, Super Walmarts and Targets, listed under "department stores" in their land-use codes, which meet the criteria of a full-service grocery store, were also considered grocery stores for this study. A very small number of grocery stores were also listed as "mixed-use commercial" lots in the MPROP. For these, I again used a combination of Google searches along with business listings to determine which addresses belonged to grocery stores. In both of these cases, I manually changed their land-use codes to match the one for grocery stores.

To define food access, I draw from the definition of a food desert provided by the USDA, which defines an urban food desert as a census tract where 33% of households to not have access to a quality food source within one mile of their residence.[11] This definition captures the importance of "walkability" for food access in an urban area, where many households, particularly low-income households, do not own a car. Modifying this definition slightly for the sake of a hedonic study, I consider food access as the number of grocery stores found within a mile of a home. With such a definition, the estimated coefficient of food access would provide the marginal implicit price of an additional grocery store within a mile of a home. However, this definition does provide some challenges. In particular, given the emphasis on "walkability", it seems unlikely that households would value an additional grocery store a mile away the same as one only one-tenth of a mile away. Recent papers by Pope (2012) and Walls et al (2015) have considered this heterogeneity in estimating the value of a Wal-Mart, and landscape views,

[11] http://americannutritionassociation.org/newsletter/usda-defines-food-deserts

respectively. Drawing from their work, I further refine my definition of food access to be the number of grocery store within quarter-mile buffer rings up to one-mile, allowing me to capture possible heterogeneity in the value of grocery stores at different distances. Figure 1.5 illustrates the location of grocery stores relative to single-family home transactions in 2005.

In addition to grocery stores, I use the parcel shapefiles and MPROP to identify several other features within the city to address the potential for omitted variable bias in the hedonic model. Omitted variable bias in the hedonic arises from structural or neighborhood characteristics not included in the regression, but correlated with observable features. As a result, the vast majority of hedonic studies control for the presences of local parks, and proximity to the city center, at a minimum, regardless of their amenity of focus. In this study focusing on food access, I also attempt to control for the influence of other commercial properties on food access as a way to address potential bias. As a retail location, grocery stores typically appear in close proximity to other retail storefronts. The stores could act as either an amenity to households (additional shopping options) or a disamenity (congestion), which raises the concern for bias if not properly controlled for. To do this, I include two explicit measures of commercial land-use, both at the household level: the percentage of land within a mile which is denoted as commercial use, and the distance to the nearest non-grocery commercial parcel.

To complement these commercial controls, I also include the number of bars within quarter mile rings up to one mile as a second proxy for commercial density. While the percentage of commercial lots represents a reasonable measure of commercial density, one concern arises from the vague and broad definition of “commercial” lots as coded in the MPROP. Upon closer observation, this definition appears to contain a number of wholesale and other lots which are unlikely to serve any use to local residents, while potentially biasing the commercial percentage control. As such, including some count of a more specific type commercial lots likely to appear in high numbers in dense commercial helps provide a balance to this potential bias. To that end bars make an ideal candidate for this role. Functionally, they have their own land-use code in the MPROP, making them easier to identify than establishments such as coffee shops, which might otherwise better fulfill this role. In addition, there is evidence to suggest that bars are likely to appear in high concentrations in urban areas, even in those with a lack of other commercial locations. Previous research has suggested that while food sources tend to diminish in communities with a high percentage of people of color, the number of bars does not (Moorland et al 2002). In addition, Wisconsin in particular has a high density of bars, with previous studies suggesting that bars outnumber grocery stores by a factor of nearly three, as illustrated in figure 1.6[12]. Finally, high concentrations of bars and liquor stores have been associated with negative amenities in previous research (see, for instance, Romley et al 2007). As such, controlling for their presence at the same geographic level as the amenity of choice in this study could also serve as a control for potential negative bias in the estimates of implicit marginal prices for grocery stores.

1.4.5 Generating Spatial Fixed Effects

Spatial heterogeneity and its impact on home prices presents another challenge to producing unbiased estimates in hedonic studies. In addition, the endogeneity of grocery store and other locations requires control of unobserved neighborhood characteristics. To account for this, I derive a set of spatial fixed effects from data characterizing neighborhoods and high school districts in Milwaukee. I obtained shapefiles characterizing high school districts from the University of Wisconsin-Milwaukee, which makes available a number of shapefiles

[12] http://flowingdata.com/2014/05/29/bars-versus-grocery-stores-around-the-world/

characterizing the geography of the city[13]. There are 11 high schools in the City of Milwaukee, and I consider models with these districts as standalone fixed effects, and others with separate time trends for each district. The Milwaukee Neighborhood Identification Project provides shapefiles illustrating the neighborhood boundaries generated by their efforts, which I use to construct fixed effects for the city's residential areas. Due to the larger number of neighborhoods defined by the city (159 out of 190 had home sales in the study period), I do not consider separate time trends of each of these but do consider a model including both neighborhood fixed effects and the high school time trends, allowing for the possibility that they offer non-mutually exclusive levels of variation. To create the dummy variables, I performed a spatial join of the shapefiles described here with the city-wide parcel shapefiles in ArcGIS. Doing so provided each parcel, including homes that sold, with a unique code signifying its high school district and neighborhood. I then constructed the dummy variables from these code.

Milwaukee neighborhoods and high school districts form a compelling set of spatial controls. In partitioning the city into neighborhoods, those working with the Milwaukee Identification Project listed six criteria for defining a neighborhood: subdivisions, major streets and other physical/natural barriers, community group participation, common housing attributes, historic areas, and resident opinions. These criteria capture many of the potential sources of unobserved local amenities that contribute to neighborhood quality, which in turn attract both homeowners and potential commercial storefronts. Similarly, within the literature, school districts are one of the most commonly researched public goods in relation to housing choice. As such, including these fixed effects, particularly when interacted with time dummies, also

[13] http://uwm.edu/libraries/agsl/milw_data_coll/

provides an important measure of neighborhood quality to deal with potential bias in the coefficients of interest.

1.4.6 Socioeconomic data

To test the hypothesis that grocery stores have a marginal implicit price in neighborhoods with a high proportion of minorities, I require demographic data on the ethnic composition of the city of Milwaukee. The US government makes shapefiles of the geographies they use for the Census freely available to the public[14]. I use data at the tract level in this study, based on its very close mapping to the residential neighborhoods in the city. As mentioned previously, the neighborhoods and tracts can be seen in figures 1.1 and 1.2, respectively. In matching the census data to each home sale transaction, I use the census data closest in data to the year the home was sold. After using a spatial join to attach the geographical ID of the census tract each home resides in, I then join tabular census data to each home sale.

1.5 Results

1.5.1 Estimating the Average Effect of Grocery Stores on Home Prices

To begin, I estimate the average effect of a grocery store within a particular radius of a home on home prices. As a baseline, I estimate a regression excluding any controls for commercial land use. Table 1.2 reports these results. The estimates suggest that grocery stores are capitalized in home prices, and positively so. In particular, for each additional grocery store within 0.25 – 0.50 miles of a home the mean price is estimated to increase by between 1.4 – 1.8 percent. I also estimate a positive influence on prices for grocery stores within 0.25 miles, though they are not significant for the models with a finer level of spatial controls, possibly reflecting limited level of variation at the neighborhood level. The estimate for an additional grocery store

[14] https://www.census.gov/geo/maps-data/data/tiger-line.html

within 0.50 – 0.75 miles in both regressions using only high school fixed effects are negative, though it is only significant at a 10 percent level in the regression allowing a separate time trend for each high school dummy. For grocery stores within 0.75 – 1 miles, the estimate is negative and significant at a 1 percent level in each model. However, the models including neighborhood fixed effects also indicate a positive impact of nearby grocery stores. The change in signs for these coefficients from the high school fixed effects model to the neighborhood fixed effects model, along with the significantly improved fit, indicates that the finer additional spatial fixed effects do control for some spatial heterogeneity not captured by the high school districts alone.

The final specification adds high school time trends to the neighborhood fixed effects model in place of the individual time dummies. Within this model, the variation in home prices arises from changes within neighborhoods by high school district and year. Given the fine level of resolution provided by this combination of fixed effects, the possibility exists that the remaining variation would not be enough to provide significant estimates of the coefficients of interest. However, the regression using these controls not only exhibit and improved fit, but also retains the signs and significance level of the regression only containing neighborhood fixed effects and time dummies. As such, I move forward with the model including neighborhood fixed effects and time dummies by high school district as my preferred specification in the rest of the analysis presented here.

As mentioned in the previous section, one cause for concern is the potential for the coefficients for grocery stores to be biased by the omission of other commercial use properties. Grocery stores rarely appear in isolation, and instead are usually located in the middle of several commercial properties. To account for this potential source of bias, I introduce several controls for commercial land-use to my preferred specification: distance to the nearest commercial lot and its square, the percentage of land within one mile of each home dedicated to commercial use, and the number of bars in quarter-mile buffers up to one ring. Given that the direction of potential bias is unclear, the possibility exists that such bias could change the signs of these estimated coefficients. Table 1.3 presents these regressions.

In comparison to the initial results, the estimated signs for the reported coefficients remain the same across all three specifications. However, in comparison to the model without commercial controls, the estimated impact on home prices increases for grocery stores in each buffer ring as more controls are added, suggesting potential negative bias from not including them previously. In particular, as column three illustrates, an additional grocery store between one-quarter to one mile increases home prices by an estimated 0.60 – 1.85 percent, in contrast to 0.69 – 1.49 percent in the initial model with no commercial controls. This supports the hypothesis that grocery stores provide specific amenities which are distinguishable from general access to retail options.

The estimates for the commercial controls also present an interesting narrative, though these results must be seen through an appropriate lens, as they were constructed primarily as a control, and not as actual amenity. In line with the vast majority of existing hedonic studies, the estimates for proximity to a non-grocery commercial lot suggest a U-shaped effect on home prices. On average, home prices flip from acting as a disamenity to an amenity at 1,602 feet. The estimated coefficients on the proportion of land dedicated to commercial land use within one mile report as negative in columns two and three, though insignificant at a 10 percent level. This is likely the result of low variance in the variable at the spatial scale of the preferred specification. Results from column three also indicate that bars within 0.5 – 1 miles of a home have a negative average effect on home prices, suggesting the potential existence of noise and crime-related disamenities of local bars. It is also possible that the inclusion of bars as a control is better capturing the effect of commercial density than the commercial percentage variable, due to the coarseness of the construction of the percentage variable.

Together, these two models suggest a fairly robust positive average marginal effect for grocery stores within a close proximity of a home. In the next section I investigate how grocery store access impacts home prices differently across socioeconomic characteristics.

1.5.2 Estimating Heterogeneous Implicit Prices for Grocery Stores

In order to estimate how grocery stores capitalize into home prices differently across neighborhoods, I interact the number of grocery stores within a given distance ring with the combined percentage of African-Americans and Latino-Americans within the home's census tract. Table 1.4 reports the results of these regressions, where I progressively add commercial controls as a robustness check.

Column one presents a regression without any of the commercial land-use checks. The estimated signs and relative levels of the coefficients for grocery stores within 0.25 – 0.75 miles remain the same as the parallel regression. The corresponding interactions are also statistically indistinguishable from zero at even a 10 percent level. However, the sign for the mean effect of grocery stores between 0.75 – 1 miles of a home flips from positive to negative, now indicating that an additional grocery store reduces home prices by 0.86 percent, though this too is insignificant at a 10 percent level. The estimated sign the interaction term for groceries with 0.75 – 1 miles of a home is positive, and together the two coefficients imply that the marginal implicit price for grocery stores is positive in neighborhoods where the combined proportion of African-Americans and Latinos exceed 0.29.

Columns two and three highlight the results from models using additional commercial land-use controls. I use the same controls as in the previous section, but I also include interaction effects for the number of bars within each buffer ring of a home transaction. The signs and levels of the grocery store coefficients and their interaction with census tract minority proportions remain fairly consistent across the three specifications. In the final specification, I estimate that the marginal implicit price for an additional grocery store within 0.75 – 1 mile of a home is positive in neighborhoods with a combined proportion of African-American and Latino households above .25. However, after controlling for the presence of bars, and their interactions with race, the grocery interaction at .25 – .50 miles also becomes significant at the 1 percent level. These results not only suggest that the estimated interaction at .75 – 1.0 miles is fairly robust to a number of controls, but that these price differentials likely exist at closer proximities as well.

Table 1.5 shows the premiums paid for an additional grocery store within .75 – 1.0 mile of their home by households based on the percentage of black and Latino households in their neighborhood, both as a percentage and as a dollar amount based on mean home prices in Milwaukee. Table 1.6 does this same for grocery stores between .25 – .50 miles. Households in neighborhoods with only African-Americans and Latinos pay a premium of 2.39 percent or $2,915.05 for each additional grocery store located within 0.75 – 1.0 miles of their home. Figure 1.7 illustrates how the dollar premium at the home level varies across neighborhoods in Milwaukee during a representative year[15]. These premiums rise to an average as high as $3,150 in some predominately African-American neighborhoods. This figure provides a particularly compelling illustration of existence of food insecurity within the city. In contrast with the

[15] I choose 2005 as the representative year as it had the highest number of home sales in the observed time period.

percentage premiums estimated via regression, these dollar estimates rely on the price paid for homes within each neighborhood, and allow for a more straightforward comparison of the prices paid for an additional grocery store across the city. While the estimates suggested differences in the percent premium paid by households, in theory it is possible that the dollar premiums could be equal across the city, or even higher in areas with a lower percentage of African-American and Latino households. While this would not suggest that food inequality does not exist within the city, this figure illustrates both higher percentage and dollar premiums paid for an additional grocery store based on the racial composition of a given neighborhood, providing strong evidence of food insecurity within the city.

Though not a place of focus, the estimated interactions for additional bars also provide some interesting results. While the mean effect was estimated as negative or indistinguishable from zero at all proximities in the non-interaction model, here the mean effect for bars within .75 – 1.0 miles flips to reporting as positive once interactions are included. Additionally, all of the interaction terms are estimated as negative, with the coefficients for bars within .50 miles significant at the 1 percent level. This also supports hypothesis presented within smaller segments of the food access literature, noting the overwhelming presence of bars and fast food locations relative to grocery stores. These results suggest the potential for further investigation of this topic in future research.

1.5.3 Robustness Check: Including Recession Year Sales

In the analysis provided above I do not include residential transactions from the Great Recession years, as there is evidence to suggest the Milwaukee housing market was not in equilibrium during those years. To test the robustness of those results though, and assess the length to which these "disequilbrium" years affect the inference of this analysis, I rerun two preferred specifications of the hedonic model. The first regression is the average effects model including commercial controls, and the second contains the interaction effects model, also with commercial controls. The results are presented in table 1.6. Qualitatively, the results are identical to the corresponding models excluding the recession years, and nearly identical quantitatively. On average, grocery stores within .25 – 1.0 miles of a home are positively capitalized into home prices, increasing home prices from 0.4%-1.25%. Once interaction terms are added, the average effects for grocery stores between .25 – .75 miles remains positive, but becomes negative for grocery stores between .75 – 1.0 miles, with home prices estimated to fall 0.81% for each grocery store in this range. However, the interaction term for neighborhood racial composition and grocery stores is positive, and in neighborhoods with a proportion of African-American and Latino households above .25, the overall impact of grocery stores is positive. In neighborhoods with only these households of color, an additional grocery store increases home prices by 2.39%. Finally, the interaction term for grocery stores between .25-.50 miles of a home remains positive, although it is not significant at a 10 percent level. Despite the near identical results, I choose to maintain the model excluding the recession year results as the primary model, due to the likelihood that the equilibrium assumption was violated during these years.

1.6 Conclusion

This study sought to estimate how food access capitalizes on home prices in Milwaukee, and how this effect varied across socioeconomic characteristics. Using publically available land use and housing data from the City of Milwaukee, along with Census data, I compiled a dataset describing Milwaukee’s housing market over a decade from 2002-2015. Defining food access as the number of grocery stores within quarter-mile buffers up to one mile, I estimated hedonic models to assess the impact of food access. I found consistent evidence of a positive average effect of additional grocery stores within one mile on home prices, particularly after using neighborhood fixed effects and high school time trends to control for potential spatial heterogeneity.

In order to test the hypothesis that households in neighborhoods with a high minority population pay a higher premium for access to grocery stores, I constructed interaction terms between the proportion of African and Latino-Americans each home's census tracts and the grocery stores within each buffer. The results indicate that households living in neighborhoods with a high proportion of African-American and Latino households pay higher premiums for an additional grocery store, particularly for grocery stores within 0.25 – 0.50 miles of a home, premiums which include both the percentage and dollar increases in home prices as the result of an additional grocery store.

While the use of the hedonic model allows for a characterization of how food access is capitalized into home prices, the model presented here does not distinguish between supply and demand side dynamics in shaping the existing price equilibrium. While the use of additional analytical steps does allow for the recovery of marginal willingness to pay, a unique demand curve cannot be identified. In addition, the hedonic model is generally unable to capture heterogeneous preference for local amenities. A top priority for future research, then, is to go beyond the analysis presented here to understanding consumer preferences for food access, and how those preferences also vary across socioeconomic characteristics. Residential sorting models, particularly the horizontal sorting model first introduced by Bayer and Timmins (2005), have become increasingly prominent over the last decade, and allows for the direct estimation of consumer preferences in residential markets through a random utility framework. Its reliance on a locational equilibrium makes the model capable of solving for a new price equilibrium after an exogenous change to the amenity space of a given market, even if the change only affects a portion of the market. As a result, the horizontal sorting model is capable of robust counterfactual policy analysis, estimating the welfare effects of policies such as an increase in the number of food outlets within particular neighborhoods in a city. Thus, this framework would serve as a useful complement to the results presented here, providing a more intricate characterization of consumer demand for food access, and assessing the impacts of policies aimed to alter the level of food access within a given housing market.

The results here could also be refined by considering alternate definitions of food access. While full-service grocery stores still serve as a primary locus of food purchase, the last two decades have seen significant growth in outlets such as farmer's markets and community gardens as options for consumers, particularly in areas with low access to fresh produce. Given the size and economic impact of these food source (see, for instance, Hughes 2008) pairing them alongside conventional grocery stores in a similar analysis would provide a more holistic vision of the relationship between residential markets and food access. In addition, integrating restaurants, particular fast food establishments, would provide useful context to these results. In areas of low food access, fast food establishments often serve as one the primary purveyors of nutrition, often exacerbating local obesity rates (see, for instance, Walker et al 2010). As such, understanding the relative importance of these establishments with "healthier" food outlets would be critical in capturing the full picture of food access in relation to consumer decisions in residential markets.

1.7 References

Alwitt, L.F., Donley, T.D., 1997. Retail Stores in Poor Urban Neighborhoods. The Journal of Consumer Affairs 31(1), 139–164

Anderson, S.T. and West, S.E., 2006. Open space, residential property values, and spatial context. Regional Science and Urban Economics 36(6), 773-789

Bajari, P, Freuhwirth, J.C., Kim, K.I., Timmins, C., 2012. A Rational Expectations Approach to Hedonic Price Regressions with Time-Varying Unobserved Product Attributes: The Price of Pollution. American Economics View 102(5), 1898-1926

Bayer, P., McMillan, R., 2012. Tiebout sorting and neighborhood stratification. Journal of Pubilc Economics 96, 1129-1143

Bayer, P. and Timmins, C., 2005. On the equilibrium properties of locational sorting models. Journal of Urban Economics, 57(3), 462-477

Block, D., Kouba, J., 2006. A comparison of the availability and affordability of a market basket in two communities in the Chicago Area. Public Health Nutrition 9(7), 837-845

Block, J.P., Scribner, R.A., DeSalvo, K.B., 2004. Fast food, race/ethnicity, and income. American Journal of Preventative Medicine 27(3), 211-217

Cho, S.H., Lambert, D.M., Roberts, R.K., Kim, S.G., 2010. Moderating urban sprawl: is there a balance between shared open space and housing parcel size? Journal of Economic Geography 10, 763-783

Chung, C., Myers, S.L., 1999. Do the poor pay more for food? An analysis of grocery store availability and food price disparities. The Journal of Consumer Affairs 33(2), 276-296

Cropper, M.L., Deck, L.B., McConnell, K.E., 1988. On the choice of functional form for hedonic price functions. The Review of Economics and Statistics 70(4), 668-675

Gamper-Rabindran, S., Timmins, C., 2013. Does cleanup of hazardous waste raise housing values? Evidence of spatially localized benefits. Journal of Environmental Economics and Management 65, 345-360

Glanz, K., Sallis, J.F., Saelens, B.E., Frank, L.D., 2007. Nutrition environment measures survey in stores (NEMS-S). American Journal of Preventive Medicine 32 (4), 282–289

Handbury, J., Rahkovsky, I., Schnell, M., 2016. Is the focus on food deserts fruitless? Retail access and food purchases across the socioeconomic spectrum. Working paper

Hendrickson, D., Smith, C., Eikenberry, N., 2006. Fruit and vegetable access in four low-income food deserts communities in Minnesota. Agriculture and Human Values 23, 371–383

Hughes, D.W., Brown, C., Miller, S. and McConnell, T., 2008. Evaluating the economic impact of farmers' markets using an opportunity cost framework. Journal of Agricultural and Applied Economics 40(01), 253-265

Hotelling, H., 1929. Stability in Competition. Economic Journal 39(153), 41-57

Kaufman, P.R., MacDonald, J., Lutz, S., Smallwood, D., 1997. Do the Poor Pay More for Food? Item Selection and Price Differences Affect Low-Income Household Food Costs. AER 759, USDA-ERS

Kuethe, T. H., 2012. Spatial fragmentation and the value of residential housing. Land Economics 88(1), 16-27

Kuminoff, N.V., Jarrah, A.S., 2010. A New Approach to Computing Hedonic Equilibria and Investigating the Properties of Locational Sorting Models. Journal of Urban Economics 67(3), 322-35

Linden, L., Rockoff, J.E., 2008. Estimates of the Impact of Crime Risk on Property Values from Megan's Laws. American Economic Review 98(3), 1103-27

Linn, J., 2013. The effect of voluntary brownfields programs on nearby property values: evidence from Illinois. Journal of Urban Economics 78(1), 1-18

Mari Gallagher Research and Consulting Group (July 18 2006). Examining the impact of Food Deserts on Public Health in Chicago. Chicago IL: Available at (http://www. marigallagher.com/projects)

McMillen, D.P., 2001. Polycentric Urban Structure: The Case of Milwaukee. Economic Perspectives, Federal Reserve Board of Chicago 25(2), 15-27

Moorland, K., Wing, S., Diez Roux, A.V., Poole, C., 2002. Neighborhood characteristics associated with the location of food stores and food service places. American Journal of Preventive Medicine 22 (1), 23–29

Morris, P.M., Neuhauser, L. Campbell, C. 1992. Food security in rural America: a study of the availability and costs of food. Journal of Nutrition Education 24(1), 525-585

Noonan, D.S., Krupka, D.J., Baden, B.M., 2007. Neighborhood dynamics and price effects of Superfund site clean-up. Journal Regional Economics 47, 665-692

Ottensmann, J.R., Payton, S., Man, J., 2008. Urban location and housing prices within a hedonic model. The Journal of Regional Analysis & Policy 38(1), 19-35

Pope, D.G., Pope, J.C., 2015. When Walmart comes to town: Always low housing prices? Always? Journal of Urban Economics 87, 1-13

Pope, D.G., Pope, J.C., 2012. Crime and property values: evidence from the 1990's crime drop. Regional Sciences & Urban Economics 42(1), 177-188

Pope, J., 2008. Buyer Information and the Hedonic: The Impact of a Seller Disclosure on the Implicit Price for Airport Noise. Journal of Urban Economics 63(2), 498-516

Powell, L.M., Slater, S., Mirtcheva, D., Bao, Y., Chaloupka, F.J., 2007. Food store availability and neighborhood characteristics in the United States. Preventive Medicine 44, 189–195

Romley, J. A., Cohen, D., Ringel, J., Sturm, R., 2007. Alcohol and environmental justice: the density of liquor stores and bars in urban neighborhoods in the United States. Journal of studies on alcohol and drugs 68(1), 48-55

Rosen, S., 1974. Hedonic Prices and Implicit Markets: Product Differentiations in Pure Competition. Journal of Political Economy 82(1): 34-55

Schuetz, J., Kolko, J., Meltzer, R., 2012. Are poor neighborhoods "retail deserts"? Regional Science and Urban Economics 42, 269-285

Taylor, L.O., Phaneuf, D.J., Liu, X., 2016. Disentangling property value impacts of environmental contamination from locally undesirable land uses: Implications for measuring post-cleanup stigma. Journal of Urban Economics 93, 85-98

Voicu, I. and Been, V., 2008. The effect of community gardens on neighboring property values. Real Estate Economics, 36(2), 241-283

Walker, R.E., Keane, C.R. and Burke, J.G., 2010. Disparities and access to healthy food in the United States: A review of food deserts literature. Health & place, 16(5), pp.876-884

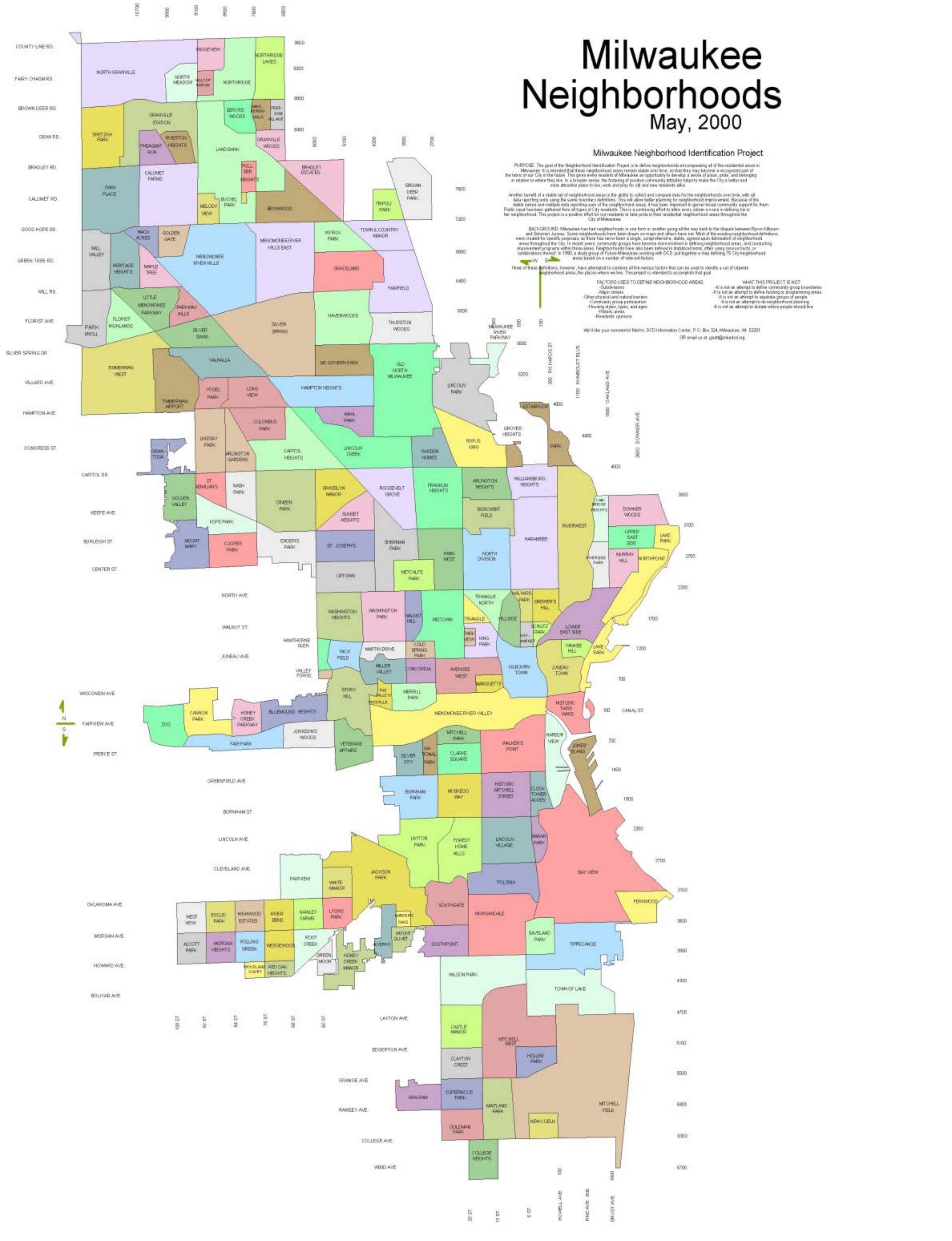
Walls, M., Kousky, C., Chu, Z. Is What You See What You Get? The Value of Natural Landscape Views. Land Economics 91(1), 1-20

Zenk, S.N., Schulz, A.J., Israel, B.A., James, S.A., Bao, S., Wilson, M.L., 2005. Neighborhood racial composition, neighborhood poverty, and the spatial accessibility of supermarkets in Metropolitan Detroit. American Journal of Public Health 95 (4), 660–667

Zenk, S.N., Schulz, A.J., Israel, B.A., James, S.A., Bao, S.M., Wilson, M.L., 2006. Fruit and vegetable access differs by community racial composition and socio-economic position in Detroit, Michigan. Ethnicity & Disease 16, 275-280

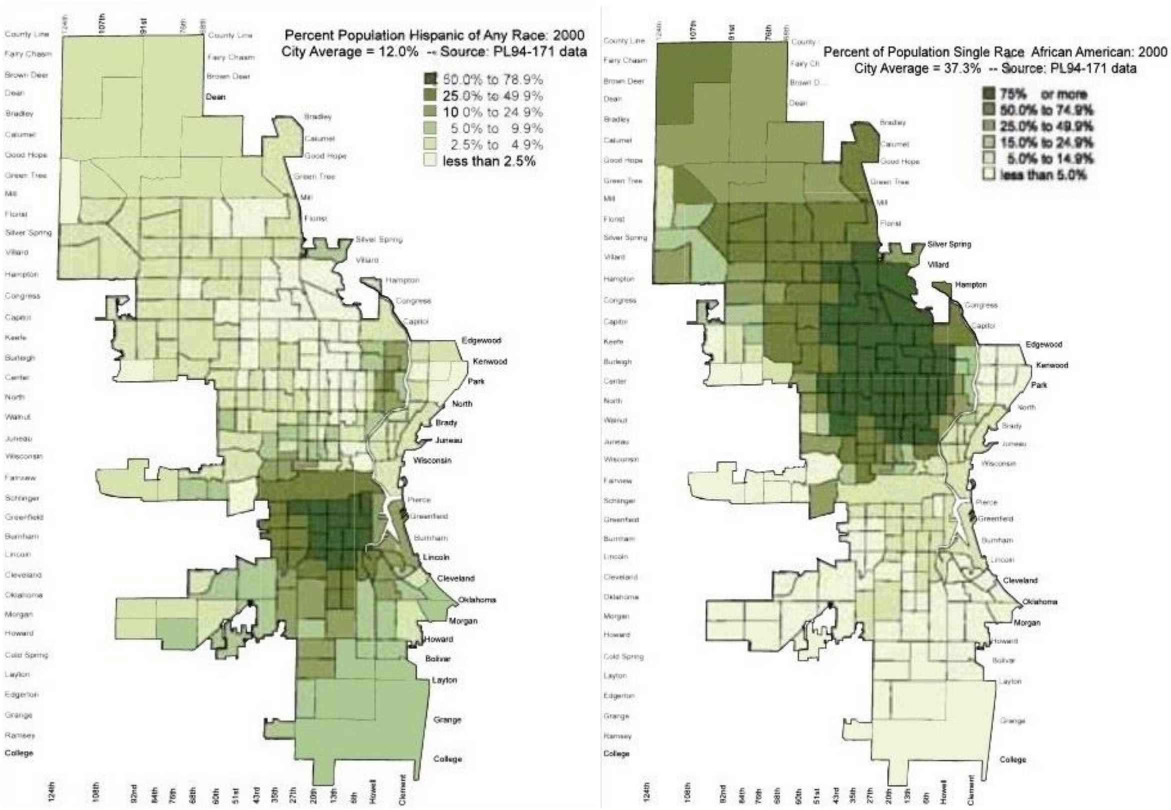
1.8 Tables and Figures

Figure 1.1: Milwaukee Neighborhoods[16]



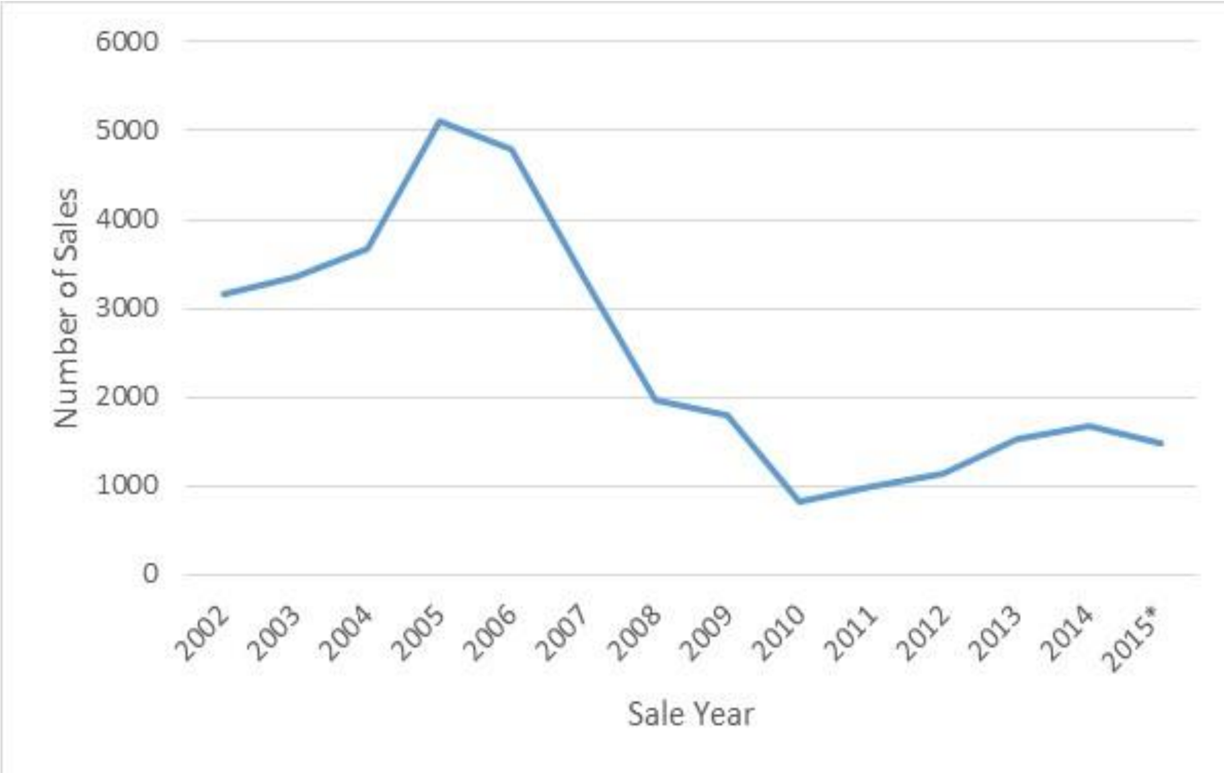
[16] http://www.milwaukee.gov/ImageLibrary/Public/map4.pdf

Figure 1.2: Race in Milwaukee by Census Tracts[17]



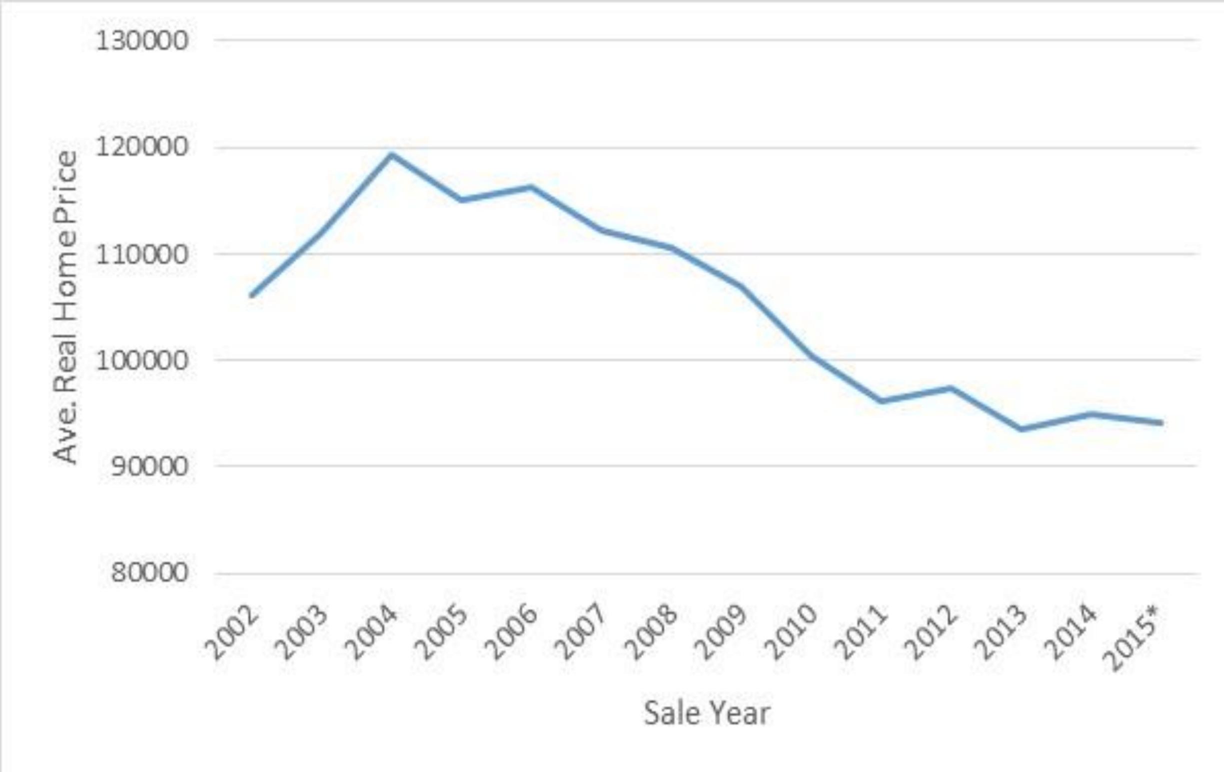
[17] http://city.milwaukee.gov/ImageLibrary/Groups/cityDCD/planning/plans/Citywide/plan/Data.pdf

Figure 1.3: Residential transactions by year



*Home Sales in 2015 are through the month of September

Figure 1.4: Average Real Home Sale Price by Year in Milwaukee in 2002 Dollars



* Home Sales in 2015 through the month of September

Table 1.1: Summary Statistics

Variable	Mean	Standard Deviation
Structural Characteristics		
Sales Price	109360.5	39587.5
Home Size (sq. ft.)	1238.113	364.692
Home Age	75.510	22.537
Bedrooms	3.025	0.762
Full Bathrooms	1.178	0.403
Half Bathrooms	0.306	0.482
Basement	0.973	0.161
Attic	0.439	0.496
Air Conditioning	0.551	0.497
Powder Rooms	0.302	0.478
Attached Garage	0.083	0.276
Detached Garage	0.795	0.403
Demographic and Neighborhood		
Proportion Black + Latino (Census Block)	0.334	0.301
Proportion Commercial Land Use within one mile	0.128	0.071
Distance to nearest Non-Grocery Commercial Lot (000's ft.)	0.550	0.456
Distance to the Central Business District (000's ft.)	24.552	9.959
Distance to the nearest Park or Playground (000's ft.)	1.333	0.977
Grocery Store Variables		
Number of Groceries within .25 miles	0.146	0.399
Number of Groceries within .25-.5 miles	0.397	0.626
Number of Groceries within .5-.75 miles	0.563	0.783
Number of Groceries within .75-1 mile	0.687	0.871
Bar Variables		
Number of Bars within .25 miles	0.213	0.520
Number of Bars within .25-.5 miles	0.543	0.900
Number of Bars within .5-.75 miles	0.860	1.174
Number of Bars within .75-1 miles	1.125	1.487

Figure 1.5: Grocery Store Locations (black/red triangle) relative to 2005 home sales

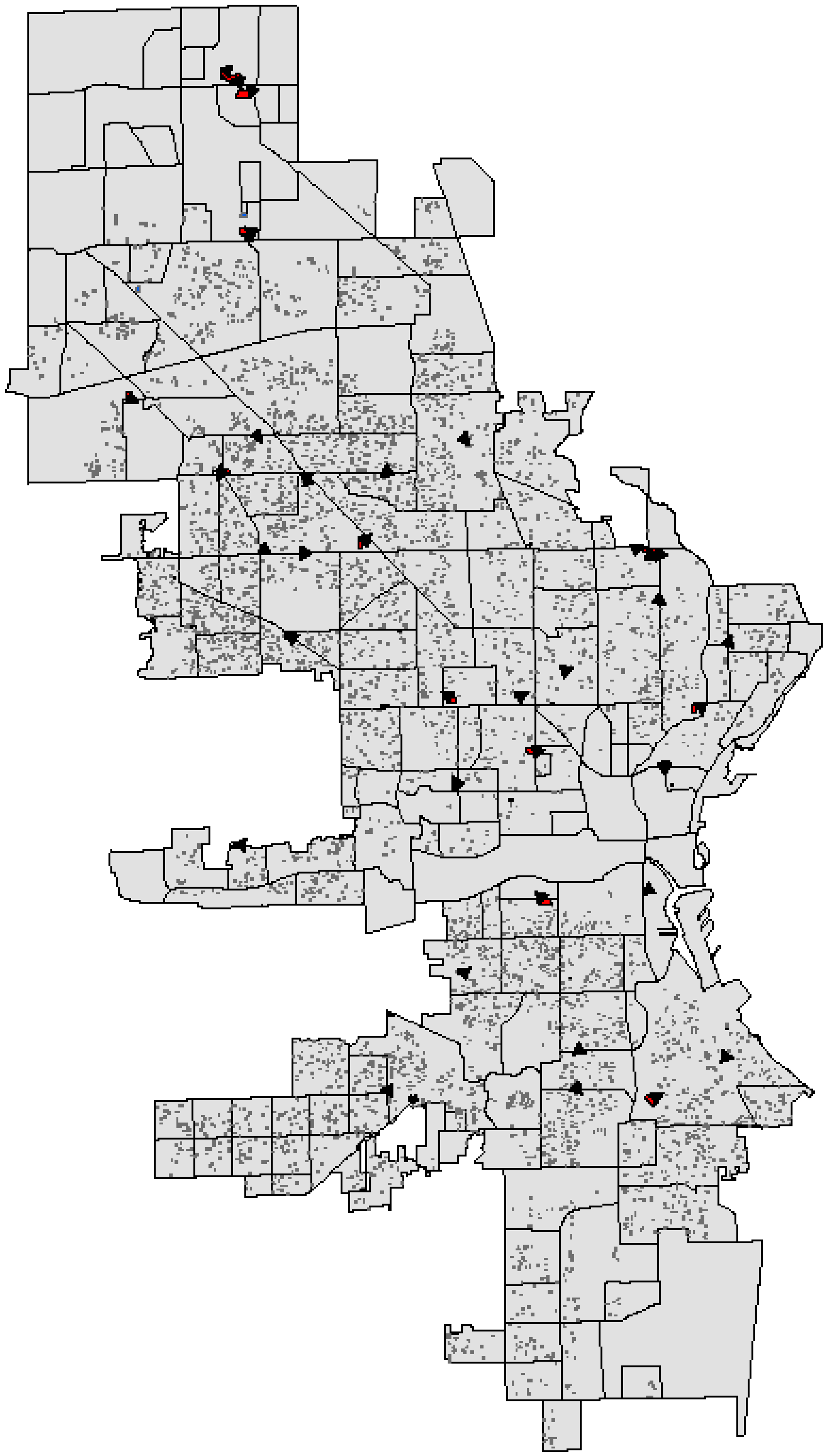
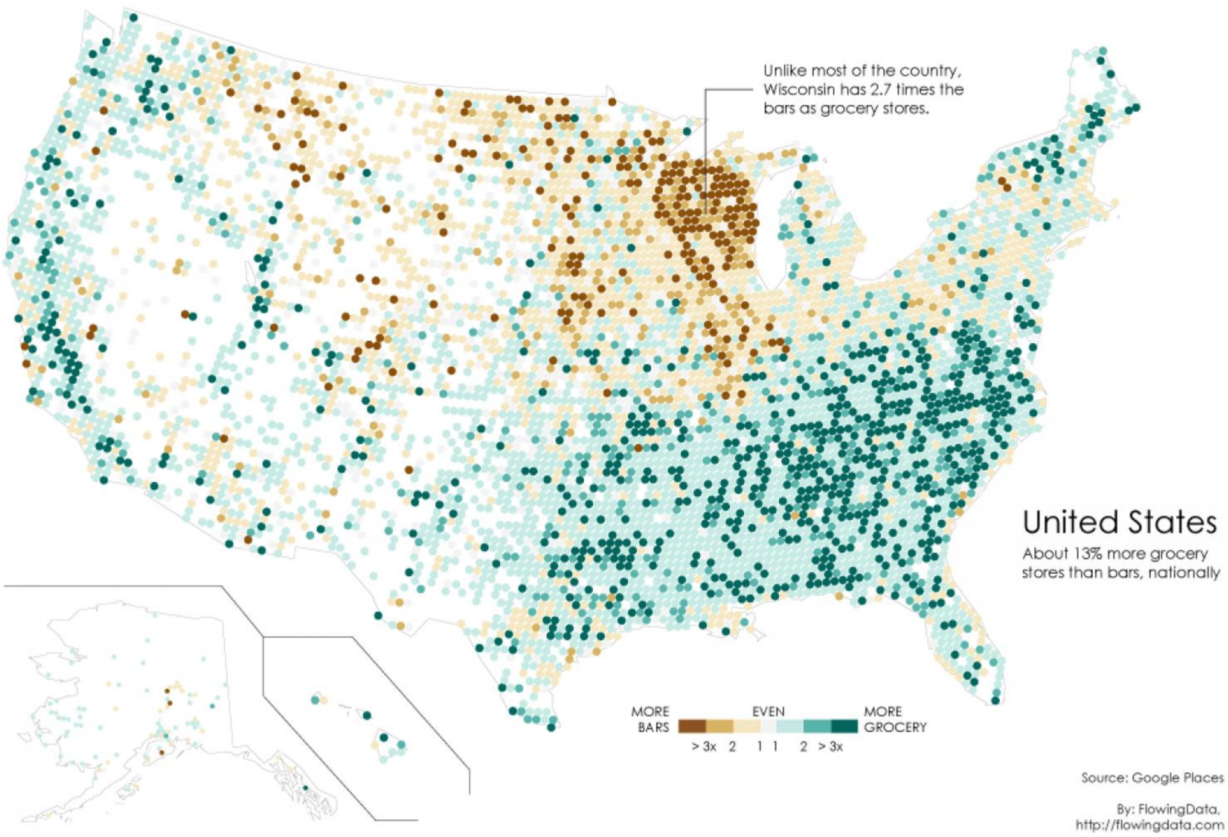


Figure 1.6: Density of bars vs. grocery stores in the United States in 2008[18]



[18] https://flowingdata.com/2014/05/29/bars-versus-grocery-stores-around-the-world/

Table 1.2: Regression results without interaction effects or commercial controls[1]

Variable (lnprice)	High School FE & Year of Sale Dummies	Separate Time Trends by High School District	Neighborhood FE & Year Dummies	NBHD FE & High School Time Trends
Grocery within .25 mi	0.0060 (0.0069)	0.0066 (0.0067)	0.0003 (0.0053)	$2.47 \cdot 10^{-6}$ (0.0048)
Grocery within .25-.5 mi	0.0141*** (0.0040)	0.0133*** (0.0038)	0.0182*** (0.0047)	0.0169*** (0.0025)
Grocery within .5-.75 mi	-0.0017 (0.0032)	-0.0034 (0.0031)	0.0129** (0.0045)	0.0123*** (0.0026)
Grocery within .75-1 mi	-0.0100*** (0.0015)	-0.0110*** (0.0020)	0.0062** (0.0024)	0.0049*** (0.0017)
% Black + % Latino census tract	-0.5916*** (0.0309)	-0.5806*** (0.0027)	-0.3301*** (0.0359)	-0.2628*** (0.0256)
Observations	26,912	26,912	26,912	26,912
Adjusted R-Squared	.6836	.6969	.7246	.7389

* significant at 10%, ** 5%, and *** 1% level

[1] Standard errors clustered at the year of sale level

Table 1.3: Preferred specification (Neighborhood FE with separate high school time trends) with commercial controls[1]

Variable Table (lnprice)	Model 1	Model 2	Model 3
Number of Groceries within .25 mi	0.0037 (.0047)	0.0039 (0.0046)	0.0040 (0.0050)
Number of Groceries within .25-.5 mi	0.0179*** (0.0024)	0.0181*** (0.0023)	0.0185*** (0.0023)
Number of Groceries within .5-.75 mi	0.0124*** (0.0026)	0.0125*** (0.0027)	0.0133*** (0.0024)
Number of Groceries within .75-1 mi	0.0056*** (0.0018)	0.0057*** (0.0018)	0.0060*** (0.0018)
Dist. to other comm.	0.0574*** (0.0077)	0.0577*** (0.0077)	0.0601*** (0.0077)
Comm. Distance^2	-0.0179*** (0.0036)	-0.0180*** (0.0035)	-0.0182*** (0.0035)
Comm. % one mile		-0.0005 (0.0006)	-0.0003 (0.0007)
Number of Bars within .25 mi.			-0.0004 (0.0053)
Number of Bars within .25-.5 mi.			-0.0031 (0.0034)
Number of Bars within .5-.75 mi.			-0.0069** (0.0022)
Number of Bars within .75-1 mi.			-0.0078** (0.0026)
% Black + % Latino census tract	-0.2632*** (0.0250)	-0.2647*** (0.0248)	-0.2556*** (0.0274)
Observations	26,912	26,912	26,912
Adjusted R-Squared	.7396	.7396	.7401

* significant at 10%, ** 5%, and *** 1% level
[1] Standard errors clustered at the year of sale level

Table 1.4: Preferred Specification with Interaction Effects[1]

Variable (lnprice)	Model 1	Model 2	Model 3
Number of Groceries within .25 mi	-0.0051 (0.0106)	0.0002 (0.0107)	-0.0055 (0.0103)
Number of Groceries within .25-.5 mi	0.0119*** (0.0037)	0.0145*** (0.0037)	0.0108** (0.0036)
Number of Groceries within .5-.75 mi	0.0163** (0.0057)	0.0165*** (0.0057)	0.0149** (0.0057)
Number of Groceries within .75-1 mi	-0.0086 (0.0048)	-0.0068 (0.0050)	-0.0081 (0.0051)
Grocery .25 x Black + Latino	0.0096 (0.0203)	0.0058 (0.0205)	0.0169 (0.0197)
Grocery .25 - .50 x Black + Latino	0.0125 (0.0097)	0.0090 (0.0055)	0.0163*** (0.0044)
Grocery .50 - .75 x Black + Latino	-0.0083 (0.0092)	-0.0085 (0.0091)	-0.0032 (0.0093)
Grocery .75 - 1 x Black + Latino	0.0301*** (0.0082)	0.0278*** (0.0082)	0.0320*** (0.0084)
% Black + % Latino Census Tract	-0.2908*** (0.0358)	-0.2878*** (0.0347)	-0.2535*** (0.0430)
Distance to Other Commercial		0.0561*** (0.0076)	0.0612*** (0.0075)
Comm. Distance^2		-0.0175*** (0.0036)	-0.0183*** (0.0036)
Commercial % 1 mile		-0.0005 (0.0006)	-0.0003 (0.0006)
Number of Bars within .25 mi			0.0252*** (0.0064)
Number of Bars within .25-.5 mi			-0.0097* (0.0053)
Number of Bars within .5-.75 mi			-0.0032 (0.0041)
Number of Bars within .75-1 mi			-0.0084** (0.0029)
Bar .25 x Black + Latino			-0.0475*** (0.0080)
Bar .25 - .50 x Black + Latino			-0.0287*** (0.0070)

Bar .50 - .75 x Black + Latino			-0.0032 (0.0085)
Bar .75 - 1 x Black + Latino			-0.0023 (0.0049)
Observations	26,912	26,912	26,912
Adjusted R-Squared	.7392	.7399	.7408

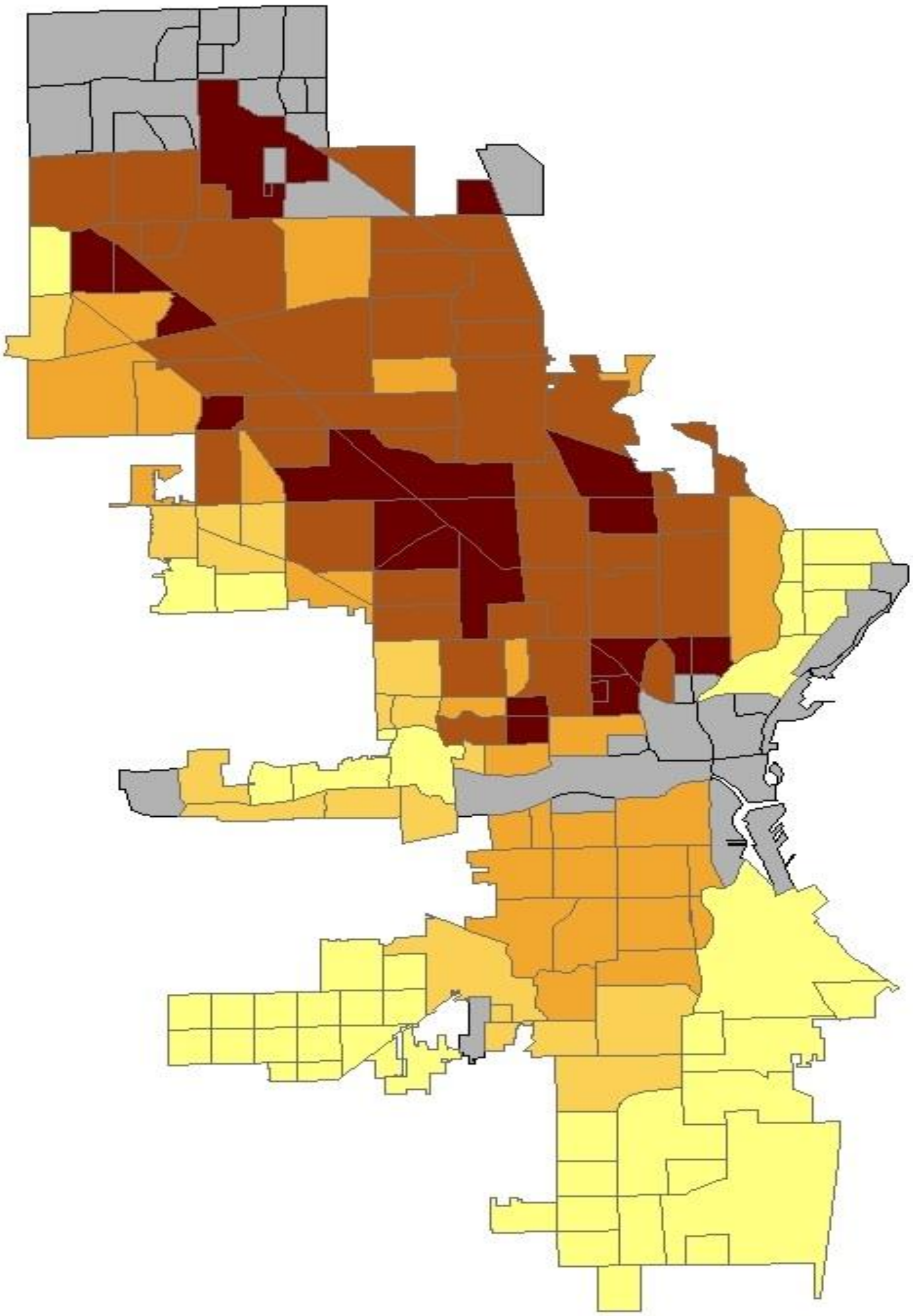
* significant at 10%, ** 5%, and *** 1% level
[1] Standard errors clustered at the year of sale level

Table 1.5: Premium paid for an additional grocery store within 0.75-1.0 miles of a home

Black + Latino	0%	20%	40%	60%	80%	100%
Percent Premium	-.808%	-.168%	.470%	1.11%	1.75%	2.39%
Dollar Premium*	-$985.51	-$204.91	$573.25	$1,353.85	$2,134.45	$2,915.05

*Based on mean home price

Figure 1.7: Dollar premiums for an additional grocery stores within .75-1 of a home in 2005



Color						
Premium ($)	-1800 - 750	750 - 0	0 - 950	950 - 1700	1700 -3150	No Sales

Table 1.6: Preferred Specifications including Recession Years[1]

Variable (lnprice)	Average Effects	Interaction Effects
Number of Groceries within .25 mi	-0.0008 (0.0048)	-0.0076 (0.0076)
Number of Groceries within .25-.5 mi	0.0125** (0.0052)	0.0108 (0.0060)
Number of Groceries within .5-.75 mi	0.0089** (0.0031)	0.0097* (0.0040)
Number of Groceries within .75-1 mi	0.0040** (.0018)	-0.0108** (0.0037)
Grocery .25 x Black + Latino		0.0156 (0.0145)
Grocery .25 - .50 x Black + Latino		0.0142 (0.0104)
Grocery .50 - .75 x Black + Latino		-0.0092 (0.0093)
Grocery .75 - 1 x Black + Latino		0.0338*** (0.0056)
% Black + % Latino Census Tract	-0.2650*** (0.0254)	-0.2628*** (0.0368)
Distance to Other Commercial	0.0640*** (0.0069)	0.0642*** (0.0066)
Comm. Distance^2	-0.0199*** (0.0029)	-0.0199*** (0.0029)
Commercial % 1 mile	-0.0001 (0.0006)	-0.0001 (0.0006)
Number of Bars within .25 mi	0.0045 (0.0043)	0.0247*** (0.0051)
Number of Bars within .25-.5 mi	-0.0027 (0.0027)	-0.0092*** (0.0044)
Number of Bars within .5-.75 mi	-0.0071*** (0.0021)	-0.0009 (0.0032)
Number of Bars within .75-1 mi	-0.0087*** (0.0021)	-0.0104*** (0.0026)
Bar .25 x Black + Latino		-0.0462*** (0.0068)
Bar .25 - .50 x Black + Latino		-0.0273*** (0.0074)

Bar .50 - .75 x Black + Latino		-0.0121*
		(0.0058)
Bar .75 - 1 x Black + Latino		-0.0051
		(0.0046)
Observations	34,860	34,860
Adjusted R-Squared	.7171	.7178

* significant at 10%, ** 5%, and *** 1% level

[1] Standard errors clustered at the year of sale level

1.A Appendix

Appendix 1.A.1: Full regression results without interaction effects or commercial controls[1]

Variable (lnprice)	High School FE & Year of Sale Dummies	Separate Time Trends by High School District	Neighborhood FE & Year Dummies	NBHD FE & High School Time Trends
Distance to CBD (000's ft.)	0.0002 (0.0023)	-0.0001 (0.0028)	0.0040 (0.0033)	0.0055 (0.0043)
Distance to CBD^2	0.0001*** ($4.32 \cdot 10^{-5}$)	0.0001* ($5.28 \cdot 10^{-5}$)	$9.28 \cdot 10^{-5}$ ($6.03 \cdot 10^{-5}$)	$6.97 \cdot 10^{-5}$ ($7.21 \cdot 10^{-5}$)
Distance to closest Park (000's ft.)	-0.0004 (0.0060)	-0.0005 (0.0059)	-0.0085 (0.0060)	-0.0102* (0.0054)
Distance to closest Park^2	-0.0022** (0.0012)	-0.0024*** (0.0011)	0.0021 (0.0015)	0.0021 (0.0012)
Stories	0.0124* (0.0064)	0.0111* (0.0060)	0.0288*** (0.0065)	0.0267*** (0.0057)
Home Age	-0.0022*** (0.0006)	-0.0024*** (0.0006)	-0.0048*** (0.0007)	-0.0049*** (0.0008)
Home Age^2	$-2.51 \cdot 10^{-6}$ ($3.55 \cdot 10^{-6}$)	$-2.36 \cdot 10^{-6}$ ($3.44 \cdot 10^{-6}$)	$1.52 \cdot 10^{-5}$*** ($3.91 \cdot 10^{-6}$)	$1.55 \cdot 10^{-5}$*** ($4.30 \cdot 10^{-6}$)
Home Size	0.0004*** ($2.05 \cdot 10^{-5}$)	0.0004*** ($2.00 \cdot 10^{-5}$)	0.0003*** ($2.09 \cdot 10^{-5}$)	0.0003*** ($1.97 \cdot 10^{-5}$)
Bedrooms	-0.0126** (0.0046)	-0.0118** (0.0044)	0.0004 (0.0041)	0.0012 (0.0041)
Full Bathrooms	0.0661*** (0.0132)	0.0629*** (0.0127)	0.0690*** (0.0114)	0.0648*** (0.0110)
Half Bathrooms	0.0449*** (0.0106)	0.0399*** (0.0107)	0.0443** (0.0094)	0.0364*** (0.0095)
Fireplace	0.0731*** (0.0079)	0.0719*** (0.0074)	0.0559*** 0.0071	0.0554*** (0.0066)
Air Conditioning	0.0696*** (0.0067)	0.0678*** (0.0061)	0.0584*** (0.0062)	0.0566*** (0.0056)
Powder Room	0.0198 (0.0151)	0.0226 (0.0155)	0.0158 (0.0137)	0.0213** (0.0143)
Basement	0.2481*** (0.0100)	0.2495*** (0.0097)	0.2326*** (0.0087)	0.2345*** (0.0088)
Attic	0.0226** (0.0078)	0.0229*** (0.0070)	0.0330*** (0.0079)	0.0333*** (0.0071)
Attached Garage	0.1088*** (0.0091)	0.1093*** (0.0083)	0.0908*** (0.0083)	0.0914*** (0.0076)
Detached Garage	0.1015*** (0.0057)	0.1006*** (0.0049)	0.0810*** (0.0061)	0.0809*** (0.0052)
Grocery within .25 mi	0.0060* (0.0068)	0.0066* (0.0067)	0.0003 (0.0053)	$2.47 \cdot 10^{-6}$ (0.0048)
Grocery within .25-.5 mi	0.0141*** (0.0040)	0.0133*** (0.0038)	0.0182*** (0.0047)	0.0169*** (0.0025)
Grocery within .5-.75 mi	-0.0017 (0.0032)	-0.0034 (0.0031)	0.0129** (0.0045)	0.0123*** (0.0026)

Grocery within .75-1 mi	-0.0100*** (0.0015)	-0.0120*** (0.0020)	0.0062** (0.0024)	0.0049** (0.0017)
% Black + % Latino census tract	-0.5916*** (0.0309)	-0.5806*** (0.0268)	-0.3301*** (0.0359)	-0.2628*** (0.0256)
Observations	26,912	26,912	26,912	26,912
Adjusted R-Squared	.6836	.6969	.7246	.7389

* significant at 10%, ** 5%, and *** 1% level

[1] Standard errors clustered at the year of sale level

Appendix 1.A.2: Full results for preferred specification (neighborhood FE with separate high school time trends) and commercial controls[1]

Variable Table (lnprice)	Model 1	Model 2	Model 3
Distance to CBD (000's ft.)	0.0046 (0.0043)	0.0042 (0.0044)	$-4.06 \cdot 10^{-5}$ (0.0045)
Distance to CBD^2	$7.72 \cdot 10^{-5}$ ($7.22 \cdot 10^{-5}$)	$8.54 \cdot 10^{-5}$ ($7.56 \cdot 10^{-5}$)	0.0001* ($7.55 \cdot 10^{-5}$)
Distance to closest Park (000's ft.)	-0.0097* (0.0051)	-0.0098* (0.0051)	-0.0100* (0.0055)
Distance to closest Park^2	0.0022* (0.0011)	0.0023* (0.0011)	0.0023* (0.0011)
Stories	0.0270*** (0.0056)	0.0271*** (0.0056)	0.0270*** (0.0055)
Home Age	-0.0050*** (0.0008)	-0.0049*** (0.0008)	-0.0051*** (0.0008)
Home Age^2	$1.60 \cdot 10^{-5}$*** ($4.26 \cdot 10^{-6}$)	$1.60 \cdot 10^{-5}$*** ($4.26 \cdot 10^{-6}$)	$1.66 \cdot 10^{-5}$*** ($4.24 \cdot 10^{-6}$)
Home Size	0.0003*** ($1.95 \cdot 10^{-5}$)	0.0003*** ($1.95 \cdot 10^{-5}$)	0.0003*** ($1.96 \cdot 10^{-6}$)
Bedrooms	0.0013 (0.0040)	0.0013 (0.0040)	0.0014 (0.0040)
Full Bathrooms	0.0651*** (0.0110)	0.0650*** (0.0109)	0.0645*** (0.0109)
Half Bathrooms	0.0367*** (0.0092)	0.0364*** (0.0092)	0.0366*** (0.0090)
Fireplace	0.0544*** (0.0065)	0.0544*** (0.0065)	0.0540*** (0.0065)
Air Conditioning	0.0562*** (0.0055)	0.0562*** (0.0055)	0.0564*** (0.0056)
Powder Room	0.0207 (0.0139)	0.0209 (0.0140)	0.0204 (0.0137)
Basement	0.2334*** (0.0088)	0.2340*** (0.0089)	0.2333*** (0.0090)
Attic	0.0338*** (0.0070)	0.0338*** (0.0070)	0.0336*** (0.0070)
Attached Garage	0.0903*** (0.0076)	0.0903*** (0.0076)	0.0905*** (0.0075)
Detached Garage	0.0799*** (0.0052)	0.0800*** (0.0052)	0.0800*** (0.0053)
Number of Groceries within .25 mi	0.0037 (.0047)	0.0039 (0.0046)	0.0040 (0.0050)
Number of Groceries within .25-.5 mi	0.0179*** (0.0024)	0.0181*** (0.0023)	0.0185*** (0.0023)
Number of Groceries within .5-.75 mi	0.0124*** (0.0026)	0.0125*** (0.0027)	0.0133*** (0.0024)

Number of Groceries within .75-1 mi	0.0056*** (0.0018)	0.0057*** (0.0018)	0.0060*** (0.0018)
Dist. to other comm.	0.0574*** (0.0077)	0.0577*** (0.0077)	0.0601*** (0.0078)
Comm. Distance^2	-0.0179*** (0.0036)	-0.0180*** (0.0035)	-0.0182*** (0.0035)
Comm. % one mile		-0.0005 (0.0006)	-0.0003 (0.0007)
Number of Bars within .25 mi.			-0.0004 (0.0053)
Number of Bars within .25-.5 mi.			-0.0031 (0.0034)
Number of Bars within .5-.75 mi.			-0.0069*** (0.0022)
Number of Bars within .75-1 mi.			-0.0078*** (0.0025)
% Black + % Latino census tract	-0.2632*** (0.0250)	-0.2647*** (0.0248)	-0.2556*** (0.0274)
Observations	26,912	26,912	26,912
Adjusted R-Squared	.7396	.7396	.7401

* significant at 10%, ** 5%, and *** 1% level
[1] Standard errors clustered at the year of sale level

Appendix 1.A.3: Full results for preferred specification with interaction effects[1]

Variable (lnprice)	Model 1	Model 2	Model 3
Distance to CBD (000's ft.)	0.0058 (0.0041)	0.0045 (0.0043)	0.0015 (0.0043)
Distance to CBD^2	$6.21 \cdot 10^{-5}$ ($7.00 \cdot 10^{-5}$)	$7.78 \cdot 10^{-5}$ ($7.36 \cdot 10^{-5}$)	0.0001 ($7.32 \cdot 10^{-5}$)
Distance to closest Park (000's ft.)	-0.0108* (0.0057)	-0.0103* (0.0054)	-0.0108* (0.0058)
Distance to closest Park^2	0.0021 (0.0012)	0.0023* (0.0011)	0.0011** (0.0023)
Stories	0.0267*** (0.0058)	0.0270*** (0.0057)	0.0262*** (0.0056)
Home Age	-0.0049*** (0.0008)	-0.0049*** (0.0008)	-0.0050*** (0.0008)
Home Age^2	$1.56 \cdot 10^{-5}$*** ($4.36 \cdot 10^{-6}$)	$1.60 \cdot 10^{-5}$*** ($4.32 \cdot 10^{-6}$)	$1.66 \cdot 10^{-5}$*** ($4.31 \cdot 10^{-6}$)
Home Size	0.0003*** ($1.97 \cdot 10^{-5}$)	0.0003*** ($1.95 \cdot 10^{-5}$)	0.0003*** ($1.94 \cdot 10^{-5}$)
Bedrooms	0.0011 (0.0041)	0.0012 (0.0040)	0.0015 (0.0041)
Full Bathrooms	0.0648*** (0.0110)	0.0650*** (0.0109)	0.0645*** (0.0110)
Half Bathrooms	0.0366*** (0.0095)	0.0365*** (0.0091)	0.0364*** (0.0087)
Fireplace	0.0555*** (0.0066)	0.0545*** (0.0064)	0.0542*** (0.0064)
Air Conditioning	0.0566*** (0.0056)	0.0562*** (0.0055)	0.0565*** (0.0056)
Powder Room	0.0212** (0.0142)	0.0209 (0.0139)	0.0208 (0.0134)
Basement	0.2333*** (0.0091)	0.2328*** (0.0092)	0.2319*** (0.0091)
Attic	0.0337*** (0.0071)	0.0341*** (0.0070)	0.0339*** (0.0070)
Attached Garage	0.0914*** (0.0078)	0.0903*** (0.0077)	0.0909*** (0.0076)
Detached Garage	0.0810*** (0.0053)	0.0802*** (0.0052)	0.0806*** (0.0052)
Number of Groceries within .25 mi	-0.0051 (0.0106)	0.0002 (0.0107)	-0.0055 (0.0103)
Number of Groceries within .25-.5 mi	0.0119*** (0.0037)	0.0145*** (0.0037)	0.0108** (0.0036)
Number of Groceries within .5-.75 mi	0.0163** (0.0057)	0.0165*** (0.0057)	0.0149** (0.0057)
Number of Groceries within .75-1 mi	-0.0086 (0.0048)	-0.0068 (0.0050)	-0.0081 (0.0051)

Grocery .25 x Black + Latino	0.0096 (0.0203)	0.0058 (0.0205)	0.0169 (0.0197)
Grocery .25 - .50 x Black + Latino	0.0125 (0.0097)	0.0090 (0.0055)	0.0163*** (0.0044)
Grocery .50 - .75 x Black + Latino	-0.0083 (0.0092)	-0.0085 (0.0091)	-0.0032 (0.0093)
Grocery .75 - 1 x Black + Latino	0.0301*** (0.0082)	0.0278*** (0.0082)	0.0320*** (0.0084)
% Black + % Latino Census Tract	-0.2908*** (0.0358)	-0.2878*** (0.0347)	-0.2535*** (0.0430)
Distance to Other Commercial		0.0561*** (0.0076)	0.0612*** (0.0075)
Comm. Distance^2		-0.0175*** (0.0036)	-0.0183*** (0.0036)
Commercial % 1 mile		-0.0005 (0.0006)	-0.0003 (0.0006)
Number of Bars within .25 mi			0.0252*** (0.0064)
Number of Bars within .25-.5 mi			-0.0097* (0.0053)
Number of Bars within .5-.75 mi			-0.0032 (0.0041)
Number of Bars within .75-1 mi			-0.0084** (0.0029)
Bar .25 x Black + Latino			-0.0475*** (0.0080)
Bar .25 - .50 x Black + Latino			-0.0287*** (0.0070)
Bar .50 - .75 x Black + Latino			-0.0032 (0.0085)
Bar .75 - 1 x Black + Latino			-0.0023 (0.0049)
Observations	26,912	26,912	26,912
Adjusted R-Squared	.7392	.7399	.7408

* significant at 10%, ** 5%, and *** 1% level
[1] Standard errors clustered at the year of sale level

Appendix 1.A.4: Full results including recession year sales[1]

Variable (lnprice)	Average Effects	Interaction Effects
Distance to CBD (000's ft.)	0.0012 (0.0034)	0.0029 (0.0032)
Distance to CBD^2	$9.26 \cdot 10^{-5}$* ($6.45 \cdot 10^{-5}$)	$5.78 \cdot 10^{-5}$ ($6.21 \cdot 10^{-5}$)
Distance to closest Park (000's ft.)	-0.0094* (0.0046)	-0.0101* (0.0048)
Distance to closest Park^2	0.0021** (0.0010)	0.0022** (0.0010)
Stories	0.0219*** (0.0056)	0.0213*** (0.0057)
Home Age	-0.0050*** (0.0006)	-0.0050*** (0.0006)
Home Age^2	$1.65 \cdot 10^{-5}$*** ($3.43 \cdot 10^{-6}$)	$1.66 \cdot 10^{-5}$*** ($3.45 \cdot 10^{-6}$)
Home Size	0.0003*** ($1.61 \cdot 10^{-5}$)	0.0003*** ($1.60 \cdot 10^{-5}$)
Bedrooms	0.0027 (0.0035)	0.0028 (0.0035)
Full Bathrooms	0.0639*** (0.0086)	0.0638*** (0.0087)
Half Bathrooms	0.0417*** (0.0081)	0.0414*** (0.0079)
Fireplace	0.0527*** (0.0052)	0.0530*** (0.0051)
Air Conditioning	0.0565*** (0.0044)	0.0568*** (0.0045)
Powder Room	0.0124 (0.0118)	0.0128 (0.0116)
Basement	0.2334*** (0.0088)	0.2223*** (0.0088)
Attic	0.0293*** (0.0062)	0.0297*** (0.0062)
Attached Garage	0.0966*** (0.0080)	0.0974*** (0.0081)
Detached Garage	0.0870*** (0.0067)	0.0878*** (0.0067)
Number of Groceries within .25 mi	-0.0008 (0.0048)	-0.0076 (0.0076)
Number of Groceries within .25-.5 mi	0.0125** (0.0052)	0.0060 (0.0044)
Number of Groceries within .5-.75 mi	0.0089** (0.0031)	0.0097* (0.0052)
Number of Groceries within .75-1 mi	0.0040** (.0018)	-0.0108** (0.0037)

Grocery .25 x Black + Latino		0.0156 (0.0145)
Grocery .25 - .50 x Black + Latino		0.0142* (0.0104)
Grocery .50 - .75 x Black + Latino		-0.0009 (0.0093)
Grocery .75 - 1 x Black + Latino		0.0338*** (0.0056)
% Black + % Latino Census Tract	-0.2650*** (0.0254)	-0.2628*** (0.0368)
Distance to Other Commercial	0.0640*** (0.0069)	0.0642*** (0.0066)
Comm. Distance^2	-0.0199*** (0.0029)	-0.0199*** (0.0029)
Commercial % 1 mile	-0.0001 (0.0006)	-0.0001 (0.0006)
Number of Bars within .25 mi	0.0045 (0.0043)	0.0247*** (0.0051)
Number of Bars within .25-.5 mi	-0.0027 (0.0027)	-0.0092* (0.0044)
Number of Bars within .5-.75 mi	-0.0071*** (0.0021)	-0.0009 (0.0032)
Number of Bars within .75-1 mi	-0.0087*** (0.0021)	-0.0104*** (0.0026)
Bar .25 x Black + Latino		-0.0462*** (0.0068)
Bar .25 - .50 x Black + Latino		-0.0273*** (0.0074)
Bar .50 - .75 x Black + Latino		-0.0121* (0.0058)
Bar .75 - 1 x Black + Latino		-0.0051 (0.0046)
Observations	34,860	34,860
Adjusted R-Squared	.7171	.7178

* significant at 10%, ** 5%, and *** 1% level

[1] Standard errors clustered at the year of sale level

CHAPTER 2

Estimating Consumer Preferences for Food Access in a Residential Sorting Model of Milwaukee, WI

ABSTRACT

A central point of contention with the ongoing discussion of food insecurity is whether food insecurity rises as the result of rational consumer decision making, or broader institutional inequalities. This debate has significant implications for policy discussions as to how to improve food access in these contexts. This paper contributes to this dialogue by estimating a structural model of consumer preferences in residential markets in Milwaukee, WI. With the horizontal residential sorting model, I am able to estimate heterogeneous consumer preferences for housing characteristics, including local food access. Using fourteen years of publically available land use and residential transaction data, I find that households of color, namely African-American households, have a significantly higher willingness-to-pay for an additional grocery store than white households, after controlling for income. Further, I find that poor households of color have a willingness-to-pay for an additional grocery store roughly equivalent to those of white households in the richest neighborhoods in the city. These results suggest that existing inequalities are not a reflection of consumer preferences alone. Counterfactual policy analysis then suggests that increasing the number of grocery stores in neighborhoods with a large concentration of poor households or households of color yield significant welfare increase for affected households, in contrast with negligible gains for the same policies in whiter and/or richer households.

2.1 Introduction

Urban and environmental economics have long been interested in the estimation of consumer preferences for local public goods. Over the last decade, a class of locational sorting models has become increasingly prominent as a tool to estimate demand for these amenities. In contrast with other models used to characterize housing decisions, in particular the hedonic model, sorting models have the ability to directly estimate structural consumer preferences. Furthermore, these same models allow for the estimation of heterogeneous preferences across space and household characteristics. This allows for more refined evaluation of how policy interventions affect households differently across space.

To date, no work in economics has investigated how food access serves as a local public good influencing how households make sorting decisions in residential markets. Outside of economics, the importance of food access has become a point of interest over the last several decades. In particular, researchers have focused on the concept of food deserts, which seeks to characterize inadequacies and inequalities in food access, especially in urban spaces. Despite the level of attention, food access and food deserts remain an ill-defined concept, with definitions and points of emphasis varying significantly across disciplines and scholars.

One common thread throughout the existing literature is the connection between socioeconomic inequality and food access. For low-income households, living in close proximity to a quality food source is likely important, due to their lack of transportation options. However, their limited income may prevent them from being able to afford homes in areas with sufficient healthy food outlets, as higher income households may also prefer the convenience of nearby food outlets, and can better afford it. Similar arguments apply to households of color, particularly black and Latino households, both due to their elevated rates of poverty and also their often limited housing choices as a result of residential segregation, both past and present. Underlying these claims is an argument of heterogeneous preferences for food access: due to factors constraining their ability to access quality food sources, lower-income households and households of color may have a stronger preference to locate in neighborhoods with higher food access. With the development of these locational sorting models, economists are well positioned to quantify the extent to which these heterogeneous preferences exist.

In this paper I use a residential sorting model to provide the first quantitative estimates of household preferences and marginal willingness to pay (MWTP) for food access in the form of large, full-service grocery stores, in the city of Milwaukee. I also estimate how these preferences vary across household income and race. Leveraging publicly available residential and land-use data, I find significant evidence of heterogeneity in the MWTP for food access, particularly along racial lines. For households of mean income, I estimate that an African-American household is willing to make a one-time payment of over $5000 for an additional grocery store within 1 mile of their residence, compared to a MWTP of less than $2000 for other households earning the same income. The results also suggest that an African-American household at the poverty line is willing to pay approximately the same amount as a non-black household earning over $90,000 a year. These results support findings suggesting that inequalities in food access are due to outside factors such as income inequality or housing segregation, and not due to weaker preferences for healthy food on the parts of poor households or households of color. They also support the growing evidence in the urban and environmental economics literature that the spatial location of land use and other public policies is vital to consider to ensure maximum welfare benefits.

Using these estimates, I perform counterfactual analysis to consider the site-specific impact of improving food access in different neighborhoods across the city. I find that policies which introduce new grocery stores to underprivileged communities based on income and race yield nearly identical estimates of welfare benefits, reflecting the conflation of poverty and race in Milwaukee and many other large urban areas. Both policies provide significantly higher benefits, both to those affected and those unaffected, than policies that increase food access in more affluent areas, where homeowners are found to be largely indifferent to the presence of additional grocery stores. These results confirm the results of previous research indicating the importance of location when enacting policies aimed at altering the levels of a local public good.

The rest of the paper proceeds as follows. Section two provides a review of the relevant work in both the food access and urban and environmental economics literature. Section three describes the theoretical sorting model and its econometric implementation. Section four describes the data and the construction of the dataset used in this analysis. Section five presents the estimates from the sorting model. Section six provides robustness checks for the sorting model results. Section seven investigates counterfactual policies and section eight concludes.

2.2 Literature Review

Academic interest in food deserts and food access date back to the 1960’s. However, the term “food desert” was not coined until the 1990’s, which coincided with the emergence of food access as a prominent and interdisciplinary topic within the academy. Despite the wide attention given to the topic, no common definition is used across the literature, leading some to question whether or not food deserts actually exist (Hendrickson et al. 2006). Despite the lack of a consensus definition, most scholars have centered on access in describing a food desert, with a food desert often representing either a lack of food outlets within a close proximity of homes within a neighborhood, or relatively high prices for healthy food options, such as fresh fruits and vegetables (see, for instance, Giang et al 2008). The term has also been subject to much critique and debate, with some scholars and particularly food justice advocates arguing that the term places the blame of low food access on those harmed by low food access[19]. As such, while I draw from the literature characterizing food deserts, I use the term "food (in)security" moving forward in defining inadequacies and inequalities in food access for this study, as it lacks these troubling connotations, and more easily accommodates the comparative analysis used here.

Another point of common inquiry across the literature is the connection between socioeconomic status and food access. More specifically, regardless of the specific metric used, many have considered whether poor households and minorities face lower food access than their counterparts. Almost universally, researchers agree that socioeconomic inequality is reflected in food access. Studies such as those performed by Chung and Myers (1999) and Powell et al (2007) indicate that poor and minority households have fewer grocery stores in their vicinity, and thus typically have to travel further to reach a food outlet. Once they arrive at a storefront, these households typically face fewer product options (Morris et al 1992) than others, and ultimately pay more for food that is often of a lower quality (Zenk et al 2006), a result borne out of both statistical analysis and household assessment (Hendrickson et al 2006).

Beyond attempting to specify what defines food access, a central debate within the literature has been (and remains) the cause of existing inequalities in food access along socioeconomic lines. Many scholars have argued that food insecurity is a reflection, in part, of a lack of demand for healthy food options among the poor and minorities. This lack of demand could be due to a lack of the income necessary to purchase fresh produce (Bitler and Haider

[19]http://www.wellhousegr.org/wordpress/wp-content/uploads/2014/05/What-is-Food-Justice.pdf

2011), or a lack of education or other cultural factors which diminish the understanding of the importance of eating fresh produce (Bitler and Haider 2011, Short et al 2007). In turn, minimal demand discourages grocery stores from locating in these areas, leading to low food access/food insecurity. These hypotheses are supported by findings suggesting that even in the presence of an additional grocery store, consumption of fresh fruits and vegetables remain largely unchanged (for instance, see Elbel et al 2015). Many similar studies in public health and economics have estimated demand elasticities for fruits and vegetables for under-privileged groups, finding that even reducing prices for these products would not encourage enough consumption for these households to meet the dietary recommended amount (for instance, see USDA 2009).

Other scholars have argued that food insecurity is instead the result of societal inequalities more broadly, rather than a lack of demand. While many acknowledge the role of income in inhibiting demand, scholars have argued against the notion that poor households and households of color inherently have low demand due to cultural factors. Zepeda et al (2004) argue that, after increasing their awareness of organic products, African-Americans actually expressed a higher willingness-to-pay for organic food than their Caucasian counterparts. Other researchers have also challenged the notion that reducing prices for fresh fruit and vegetables would have a negligible impact on their consumption. Weatherspoon et al (2013), for instance, conclude that given their higher sensitivity to price, introducing additional sources of fresh food, and thus competition in the food market, would be a key aspect to addressing obesity and poor diets in areas of low food access. Rather than a lack of demand then, these scholars argue that other structural factors, including income inequality, explain the existence of food insecurity along with any perceived lack of demand for quality food sources. As an example, several scholars have pointed to the practice of redlining, both in residential and commercial markets, as a mechanism which has created food insecurity which disproportionately affects people of color. Through a lack of access to low-interest loans, restrictive housing covenants, and other forms of housing segregation, people of color are unable to purchase homes in areas with ample sources of fresh food (ACLU 2006). These problems are then intensified by similar practices on the commercial side, preventing minority populations from being able to open their own businesses within their neighborhoods, such as insurance companies refusing to insure businesses in areas with high concentrations of people of color (Eisenhauer 2001, ACLU 2006). It is here that the current project enters and contributes to the food access literature. By testing for heterogeneity in preferences for food access, I provide evidence as to whether or not households in under-privileged communities have comparable demand for food outlets as their counterparts.

Within the economics literature, this research contributes to a growing body of work which utilizes sorting models to estimate consumer demand for local public goods. These models are largely derived from Tiebout's (1956) seminal paper in public finance on sorting in housing markets. Subsequent theoretical work by Epple and Zelenitz (1981), Epple et al (1984, 1993), and Benabou (1993, 1996), among others, further laid the theoretical foundation for this model. These sorting models have two variants seen in the empirical literature: the vertical and horizontal sorting model. The horizontal sorting model, which I use in this project, builds on the random-utility model (RUM), as seen in McFadden (1978) and Berry et al (1994). However, it departs from the standard RUM in two keys ways. First, the horizontal sorting model includes a set of alternative specific constants, aspects of utility observable to the agents, but not the modeler. Second, the model allows for the inclusion of interaction terms between household characteristics and alternative specific variables. These interaction terms characterize heterogeneous preferences for any selection of alternative specific variables.

Compared with the hedonic model, which focuses on the price equilibrium in residential markets, and the vertical sorting model, which assigns a housing type to each household in the market, the horizontal sorting model uses a supply and demand framework in defining its locational equilibrium. As a result, the horizontal sorting model is able to derive a new price equilibrium after an exogenous change to the amenity space a given market. This allows for counterfactual analysis, estimating the impact of changes to the levels of amenities across housing alternatives, and the resulting changes to housing prices, on households. This empirical power comes at the cost of stronger assumptions on household preferences and error structure, however, in comparison to the hedonic model. To date, this framework has been used to estimate household preferences for a wide class of public goods and amenities, such as school quality (Bayer and McMillan 2012), socioeconomic characteristics of one's neighbors (Bayer et al 2009), and heterogeneous open space types (Klabier and Phaneuf 2010). This paper represents the first attempt to extend this model to consider food access as an amenity in this way, particularly as one that household have heterogeneous preference over, representing the second major contribution of this paper to the literature.

2.3 Model

2.3.1 Theoretical Model

Households, denoted by i and with individual characteristics I, choose between housing types h, which in this application are defined by their location, j, and the time of sale, t, so that $h = (j, t)$. Each housing type is characterized by its structural characteristics, S_h, the characteristics of its neighborhood, N_h, its price, p_h, a component of unobserved characteristics, ξ_h, and an idiosyncratic error term ε_h, all of which are allowed to vary across the j and t

dimensions, but are constants across individuals. For a given a housing type $h = (j, t)$, consumer preferences are defined with the following indirect utility function:

$$U_{jt}^{i} = U_{h}^{i} = V\left(S_h, N_h, \widehat{I^{\iota}}, p_h, \xi_h\right) + \varepsilon_h^{i}, \qquad (1)$$

In this project, I use $\widehat{I^{\iota}}$ to denote individual household characteristics, as I do not have access to household-level demographics, and am forced to aggregate up to the census block level. I discuss the implications of this aggregation later in this and the following sections. I follow the previous sorting literature by using a linear specification for utility:

$$U_h^i = \alpha_S^i S_h + \alpha_N^i N_h + \alpha_P p_h + \xi_h + \varepsilon_h^i, \qquad (2)$$

A key element of the horizontal sorting model is the inclusion of interaction terms between alternative and household characteristics, which allow for preferences for housing to vary across households. In this context, I allow preferences to vary for housing type characteristics $X \in \{S, N\}$ across households with individual characteristics *I* such that:

$$\alpha_X^i = \alpha_{0X} + \textstyle\sum_{k=1}^{K} \alpha_{kX} \widehat{I_k^{\iota}}, \qquad (3)$$

where α_{0X} represents a mean component common to all households, and *K* denotes the number of household characteristics. Of particular interest in this research are the interactions between the number of grocery stores in a neighborhood, and household characteristics.

Households choose a housing type to maximize utility. That is, household *i* will choose housing type $h^* = (j^*, t^*)$ if and only if:

$$U_{h^*}^i \geq U_h^i, \forall h \neq h^* \qquad (4)$$

Given the utility specification in (2), and an assumption on the distribution for ε_h^i, the probability that a household chooses alternative *h* can be defined as a function of both the observed and unobserved components of all the alternatives in their choice set:

$$Pr_h^i = f_h\left(\boldsymbol{S}, \boldsymbol{N}, \boldsymbol{P}, \boldsymbol{\xi}, \boldsymbol{I}, \widehat{I^{\iota}}\right), \forall i, h \qquad (5)$$

Aggregating equation (5) across the entire market then yields the demand for housing type *h*:

$$\sigma_h^d = \int f_h(\boldsymbol{S}, \boldsymbol{N}, \boldsymbol{P}, \boldsymbol{\xi}, \widehat{I^i})\, g(\widehat{I^i}) d\widehat{I^i} \qquad (6)$$

where $g(\widehat{I^i})$ is the distribution of individual characteristics. Given supply share σ_h^s, which is assumed exogenous in this project, a market-clearing equilibrium is defined when demand share equals supply share for each housing type in the market:

$$\sigma_h^d = \sigma_h^s, \forall h \quad (7)$$

2.3.2 Econometric model

Estimation of this model occurs in two stages, similar to that seen in Berry et al (1995), and more directly seen in a sorting model application in Bayer et al (2004). The discussion here also closely follows those seen in Bayer et al (2007) and Klaiber (2008). To expedite the description of this process, I rewrite the utility function in (2) as:

$$V_h^i = \Theta_h + \Gamma_h^i + \epsilon_h^i, \qquad (8)$$

where:

$$\Theta_h = \alpha_{0S} S_h + \alpha_{0N} N_h + \alpha_p p + \xi_h \qquad (9)$$

and:

$$\Gamma_h^i = \left(\textstyle\sum_{k=1}^{K} \alpha_{kS} \widehat{I_k^i}\right) S_h + \left(\textstyle\sum_{k=1}^{K} \alpha_{kN} \widehat{I_k^i}\right) N_h \qquad (10)$$

In equation (9), the $\Theta_h{}'s$ represent the mean indirect utility of housing choice *h*, common to all households, and of which there is one for each housing alternative in the choice set. Equation (10), defining Γ_h^i, represents the heterogeneous component of utility, unique to each household. Notice here that I choose not to allow the price coefficient to vary. While the price coefficient can also vary across households, in this application I find that it only serves to complicate the calculation of welfare effects, which I discuss in further detail later. To test the robustness of my results, I also performed the two stages of analysis discussed here while allowing the price coefficient to vary. In doing so, I obtained qualitatively identical results. The results were also very similar quantitatively. As such, I move forward presenting only the analysis with a constant price coefficient.

In the first stage of estimation, a maximum likelihood estimator (ML) is used to recover the mean indirect utilities Θ_h and the heterogeneous parameters α_{kX} in Γ_h^i. It is worth noting the difference in the dimensions of the parameters estimated at this stage. As previously mentioned, there is a single Θ for each housing alternative, so this first stage recovers h Θ. As for the heterogeneity parameters, if there are k individual household characteristics, and $s + n$ structural and neighborhood characteristics for each alternative in the choice set, there are a maximum of $k \cdot (s + n)$ potential heterogeneity parameters to identify at this stage. Functionally, the researcher distinguishes which heterogeneity parameters are likely to be significant, so the actual number of parameters estimated here are almost certainly a subset of this maximum total.

The ML estimator functions by choosing parameters to maximize the probability that it correctly matches each household with its housing choice. Given an assumption about the distribution of ϵ_h^i, this probability, characterized in equation (5), has a closed form. In this application, I assume that ϵ_h^i takes an i.i.d. extreme value distribution, leading to the familiar multinomial logit form seen in many discrete choice applications:

$$Pr_h^i = \frac{e^{\Theta_h + \Gamma_h^i}}{\sum_{h'} e^{\Theta_h + \Gamma_h^i}} \qquad (11)$$

Maximizing the probability that each home makes the correct choice yields the following log-likelihood function:

$$\ell\ell = \sum_i \sum_h Y_h^i \ln(Pr_h^i) \qquad (12)$$

where Y_h^i is an indicator variable which equals one if agent i chooses housing type h and equals zero otherwise.

The intuition behind the estimation of the heterogeneity parameters α_{kX} is straightforward. If it is observed, for instance, that more affluent households consistently choose neighborhoods with additional grocery stores, a higher value for the interaction term between local grocery stores and household income will increase the likelihood higher income households will correctly be matched to neighborhoods with greater food access. In this way, the estimator leverages the aggregate decisions made by households with similar demographic profiles to pin down the heterogeneity parameters.

The logic for the estimation of the indirect utility parameters Θ_h is less immediate. Examining the derivative of the log-likelihood function with respect to Θ_h, which represents a first order condition, helps illuminate this reasoning:

$$\frac{\partial \ell\ell}{\partial \Theta_h} = \sum_{i \in h} \frac{\partial \ln(Pr_h^i)}{\partial \Theta_h} + \sum_{i \notin h} \frac{\partial \ln(Pr_h^i)}{\partial \Theta_h} = \sum_{i \in h}(1 - Pr_h^i) + \sum_{i \notin h}(-Pr_h^i) = 1 - \sum_i (Pr_h^i) = 0 \quad (13)$$

This equation represents the tradeoff made during the first stage estimation. The above condition must hold for all housing types h. By increasing an individual Θ_h, it increases the indirect utility of that housing type for all households and increases the chance that each household chooses that housing type, and lowers the probability that they chooses any other housing type. Thus, the first stage estimation process identifies the combination of parameters that matches each household with its housing choice, while balancing these tradeoffs. This is accomplished in this study through the use of a contraction mapping, as seen in Berry (1994). Specifically, the contraction for this process is:

$$\Theta_h^{t+1} = \Theta_h^t - \ln \sum_i \widehat{P_h^i} \qquad (14)$$

Where *t* represents the number of contractions. Using this contraction mapping allows for the quick estimation of a consistent vector of the mean indirect utilities, even with a large number of elements.

In the second stage, the mean indirect utilities Θ_h are decomposed into its observable and unobservable components as seen in equation (9), which bears a significant resemblance to its equivalents in the IO literature. Absent any endogeneity concerns, equation (9) could be estimated via ordinary least squares, regressing the Θ of each housing type *h* against its structural and neighborhood characteristics. However, it is certainly the case that the sorting decision by households is influenced by features unobservable to the researcher. These features, in turn, are likely correlated with price, which is included in the second stage estimation. For instance, higher quality locations are likely to have higher prices, in part due to unobservable features, leading to the correlation with price. To address this, I use an instrument developed within this sorting literature which closely follows the approach used in the IO literature.

Within the IO literature, creating an instrument for price relies on the use of characteristics of close substitutes for each product. The intuition here is that the price of a given product will not only be influenced by its own quality, but also the quality of its close competitors. As an example, take the market for automobiles studied in the Berry et al (1995) paper typically cited in discrete choice applications. More specifically, consider minivans produced by Mazda and Toyota. Both vehicles operate in the market for minivans, and are in direct competition with one another. As such, for any individual seeking to purchase a minivan, their decision will not be determined by the attributes of the vehicle produced by Toyota or Mazda in isolation, but by the full set of alternatives available to them with both minivan options. As such, the price of either vehicle will be dependent in part on the features of the other, since they are competing with one another. However, it is unlikely that the price of a Toyota minivan is directly correlated with unobserved features of a Mazda which might affect consumer utility, since they are produced by different companies. This allows the observable features of competing options in a given market a valid instrument, given that there is sufficient “distance” (i.e. different companies) between the competitors.

The instrument used in the residential sorting, first developed by Bayer et al (2004), uses a similar logic, exploiting the spatial structure of the housing market. Two homes of identical quality command different prices based on their locations within a given market: one may be located in a higher quality neighborhood, based on unobservable characteristics of the two neighborhoods the homes are located in. Regardless of their distance from each other, however, both homes are located in the same housing market. As such they are linked by a common equilibrating process, since households will consider the attributes of these and other similar homes within the market when making their housing choice. This indicates that home prices across the market are determined by a common set of observable characteristics. At the same time, the physical distance between homes at opposite ends of the market provides a plausible case that the observable characteristics of one such home is correlated with the unobservable characteristics that affect household utility of its distant competitor. The physical distance in this application is the analog of the metaphorical distance between products in the market for minivans. In other words, while the size of a home on the north end of Milwaukee is likely to be correlated with the price of a home on the south end, it is unlikely that the square footage of a home on the north side is correlated with the unobserved characteristics impacting the quality of a neighborhood on the south side. This requires the identification and assumption of an appropriate distance beyond which homes in a given market meet the criteria of independence from unobserved features of competing homes. Then, by controlling for the observable characteristics influencing home price of housing alternatives within a set radius of a given neighborhood, the exogenous attributes of alternatives outside of this range can be used as an instrument.

Following Klaiber and Phaneuf (2010), the process for creating an instrument proceeds in the several steps. To begin, equation (9) is rearranged as follows:

$$\Theta_h - \alpha_p p = \alpha_{0S} S_h + \alpha_{0N} N_h + \xi_h \qquad (14)$$

An initial guess for α_p, α_p^*, is made, and equation (14) is estimated via OLS, including additional neighborhood regressors from nearby neighborhoods. In this project, I formed concentric circles from the center of each neighborhood with diameters of .5, 1, 1.5, and 2 miles. For unique neighborhoods contained in each ring, I averaged the observable characteristics included in analysis here for all homes sold within a given year in these neighborhoods. These characteristics are then included in equation (14), to be estimated by OLS:

$$\Theta_h - \alpha_p^* p = \alpha_{0S} \widetilde{S_h} + \alpha_{0N} \widetilde{N_h} + \widetilde{\xi_h} \qquad (15)$$

where $\widetilde{N_h}$, $\widetilde{S_h}$, and $\widetilde{\xi_h}$ reflect the inclusion of the additional regressors. By including these additional variables, the residual variation from the estimation of equation (15), $\widetilde{\xi_h}$, thus arises from neighborhoods outside of the two-mile radius of each home, which is assumed to be independent of the unobserved features of central neighborhood. The estimates from this regression are then used to construct a new vector of mean indirect utilities $\widetilde{\Theta}_h$, where $\widetilde{\xi_h} = 0$, so that $\widetilde{\Theta}_h = \hat{\alpha}_{0s} \widetilde{S_h} + \hat{\alpha}_{0n} \widetilde{N}_h + \alpha_p^* p$. The instrument is then generated by setting and determining the vector of prices p^{iv} that satisfies the market-clearing condition:

$$\sigma_h^s = \frac{1}{N} \sum_{n=1}^{N} \frac{\exp(\widetilde{\Theta}_h + \widehat{\Gamma}_h^i)}{\sum_{q=1}^{J} \exp(\widetilde{\Theta}_q + \widehat{\Gamma}_q^i)} \qquad (16)$$

where the $\hat{\Gamma}_h^i$ are constructed using the first stage interaction estimates. The resulting constructed price vector p^{iv} is an aggregator of the exogenous variation of these distant neighborhoods which affects housing prices. By assumption, the variation in p^{iv} only comes from the exogenous variation of the neighborhoods outside of this two-mile radius. This assumption is necessary to state that the instrument manages to adequately explain how these features are likely to affect the equilibrium price of a single home choice.

This instrument is then used in the traditional two-stage IV regression. Finally, to ensure the instrument is not influenced by the initial guess for α_p^*, the process is then repeated using the estimated α_p from the IV regression as the new guess for α_p^*. As in Klaiber and Phaneuf (2010), I find that in practice the instrument does not change after five iterations of this process.

Finally, it is useful to discuss the asymptotic properties of the estimator outlined here. I do not go into detail in this paper, but these properties are given a more thorough treatment in Berry (2004), and I also draw from Klaiber (2010) and Bayer et al (2012). In the case that the true vector of alternative specific constants $\mathbf{\Theta}$ is known, then the consistency of the second stage estimates is given if the number of housing types approaches infinity. However, in this project as is generally the case, this vector is estimated and not known. As such, additional assumptions are necessary in order to ensure the consistency of the second stage estimates. Specifically, it is necessary that the number of agents *N* must grow large relative to the number of housing types *H*, a condition which is met if $\frac{H \ln H}{N}$ approaches zero as $H \rightarrow \infty$. In addition, asymptotic normality at rate $\sqrt{H}$ occurs as long as $\frac{H^2}{N}$ is bounded. As discussed in the next section, the choice set available to households contains a distinct number of housing types, one which is greatly outnumbered by the number of agents observed. Thus it is reasonable to assume that these assumptions are met in this study.

2.4 Data

2.4.1 Study Area

My study area is the City of Milwaukee, the largest city in the state of Wisconsin at 96.8 square miles, with a population of 595,064 as of the 2010 Census. While it has not received the same level of attention as larger urban areas in economic housing analyses or food access studies, the USDA does find that several census tracts within the city qualify as “low income” and “low access”[20]. In addition, the limited academic research to date done in Milwaukee confirms the presence of limited accessibility and affordability in areas across the city (Johnson et al 1996, Pothukuchi 2004). Both the city government and outside nonprofits, such as Growing Power, have undertaken initiatives to address these issues over the last decade, though challenges persist[21,22]. Milwaukee makes for an interesting case study for a number of reasons. First, it bears several similarities to other urban, industrialized cities such as Chicago and Detroit (McMillen 2001), which have received much greater notoriety for food inequality within the academic literature and beyond (for example, see Zenk et al 2005). Much like those cities, Milwaukee is home to significant levels of racial and income inter and intra-city segregation (see, for instance, Massey and Denton 1989). This level of segregation eases the challenges presented by the absence of household-specific demographic data (see section 4.6), and the stark contrast between those living within the city versus those living outside of it allows me to consider the city of Milwaukee as a distinct housing market, diminishing concerns of bias from incorrectly specifying the scale of the market within my study. Finally, the city provides a wide selection of

---

[20] http://www.ers.usda.gov/data-products/food-access-research-atlas/go-to-the-atlas.aspx

[21] http://www.biztimes.com/2016/07/11/feeding-milwaukee/

[22] http://www.superiorsites3.com/NNSp09GrowingPower.pdf

publicly available land use and residential data. This, in conjunction with publicly available Census data, are the inputs for my data set.

2.4.2 Defining the Choice Set

Estimating a residential sorting model requires that the study area be divided into complete and mutually-exclusive neighborhoods. These neighborhoods necessarily have a spatial component, though additional levels of aggregation along temporal or structural lines may also be used in order to increase the available variation for analysis. For this project, I define a choice by its location and year of sale. For the locational aspect of the choice set, I use neighborhood definitions created by the Milwaukee Neighborhood Identification project, which divided the city into 190 neighborhoods. In partitioning the city into neighborhoods, those working with the Milwaukee Identification Project listed six criteria for defining a neighborhood: subdivisions, major streets and other physical/natural barriers, community group participation, common housing attributes, historic areas, and resident opinions. These factors fit naturally with the factors a household likely considers when choosing a neighborhood, and thus are an ideal locational axis for the choice set. The city provides a shapefile characterizing these neighborhoods for free to the public[23]. Figure 2.1 illustrates these neighborhoods.

The second component of the choice set is the time at which households decide to make a purchase. This study considers all home sales in Milwaukee during the time periods of 2002-2015, and I treat each year as a separate period during which households may choose to enter the market and purchase a home. One potential source of concern with this time frame lies with the recent recession of 2007-2010, particularly as it relates to the housing market. While the residential sorting model does not rely on the same type of equilibrium assumptions as the

[23] http://city.milwaukee.gov/DownloadMapData3497.htm#.V4VYRbgrKhc

hedonic model, one potential concern is that a dramatic shift in home prices and consumer behavior during those years might confound the estimates of household preferences. To address the potential concern, I later estimate this model with those years removed for comparison. With 190 neighborhoods across 15, this generated a choice set with 2660 options. Removing those alternatives with no or few observed sales reduced this number to 1199. This final total represented an unbalanced panel, with 133 neighborhoods represented across 15 years. Incorporating a time element to the choice set raises questions as to how to incorporate future expectations of prices when households make their choices.

In this paper, I simplify the estimation process by making three key assumptions. First, I assume that population growth is constant in the study area and driven by outside trends, and that external migrants arrive in an exogenously given time period. Second, internal migrants also leave and re-sort within the study area at an exogenously given time based on individual shocks. These two assumptions imply that there is an equilibrium for each of the 10 years over which the study takes place. Finally, the equilibria are linked under the assumption that preferences are constant over time. To test the robustness of this third assumption, I also run the same model for each year individually, and obtained qualitatively similar results to those presented here.

2.4.3 Residential Transactions

Milwaukee's Assessor office provides data characterizing land sale transactions dating back to 2002. These transactions include single-family homes in addition to condominiums, vacant lots, and commercial properties. In addition to the sale year and price, the data provide information on basic structural characteristics, including the number of bedrooms and bathrooms, lot and home size, and home style[24]. After removing incomplete and erroneous

---

[24] http://assessments.milwaukee.gov/

records, I also trim 2.5% of the records based on price from both tails of the distribution. Finally, after removing those sales associated with alternatives which were removed, I end up with 34,885 records used in this analysis. In order to determine the attributes of the housing alternatives, I take the average characteristic of homes sold within a neighborhood during a given year, including price. This process is similar to that of Chay and Greenstone (2005), who differ primarily in their use median home prices. I present the distribution of a select number of attributes in Figure 2.2. A summary of the structural and neighborhood characteristics of each housing type, along with the individual household characteristics, is also found in table 2.1

2.4.4 Land-Use Data

To characterize the broader landscape around each home, including food access, I leverage Milwaukee's Master Property Record (MPROP) and parcel shapefiles, which provide annual spatial and tabular data describing every parcel in the City of Milwaukee dating back to 2001[25]. These records provide structural and ownership information for each of the parcels in the city. The city also provides a parcel shapefile, which illustrates each parcel in a city-wide map. Each parcel has a unique taxkey, contained in both the parcel shapefiles and MPROP. This allows me to match each parcel's structural characteristics with its location within the city. The same taxkey is also present in the home sales transaction data, enabling me to isolate home sales on the map, and construct the variables describing the features of their surroundings.

The structural features found in the MPROP contain both physical characteristics of the buildings themselves, including additional variables not found in the land-sale transactions data such as the presence of an attic or air conditioning, and also of the land-use characteristics of the parcel. The MPROP contains two distinct land-use codes, which provide a fairly robust

[25] http://city.milwaukee.gov/mapmilwaukee

categorization of the city's parcels. The first of these is a four-digit code based on the Standard Industrial Classification code (SIC), which identifies the activity done on a given parcel. Among these codes is a code for grocery stores, which is my basis for defining food access. A second land use variable separates the parcels into eight larger catch-all categories, such as commercial, residential, and manufacturing, which I use to construct broader land use variables.

2.4.5 Defining Food Access and other Neighborhood Amenities

To define food access, I draw from the existing literature on food security, which typically defines food access using large, full-service[26] grocery stores. While many stores, including convenience stores, offer various food products, full-service grocery stores typically have lower prices (Kaufman et al 1997) and a wider variety (Glanz et al 2007) of options, typically offering broad selections of fresh meat and produce, often not found in one location with convenience stores.

From this position, two issues present themselves: defining large, full-service grocery stores in the context of Milwaukee, and deriving a metric of food access around this definition. To begin, I sorted all of the locations coded as grocery stores present in the MPROP by their lot size. In examining the locations, a clear break in grocery store size occurred at 10000 square feet, with the largest grocery store across the ten years under that size reporting at 8916 square feet, and the next largest reporting at 10406 square feet[27]. I then cross-checked the addresses provided by the MPROP for each of these stores with Google results, listings in the yellow pages, and store websites when available. This confirmed that stores above this size consisted of only full-

[26] I use the rubric provided by the District of Columbia in defining a food desert: http://abra.dc.gov/page/full-service-grocery-stores

[27] Two stores were listed at approximately 10000 square feet across the study time period, but both were Family Dollars, which were not considered for this study

service grocery stores (such as Whole Foods or Pick N' Save), with stores below this cutoff largely consisting of convenience stores (of a similar style to Seven-Eleven), along with a small number of corner stores, none of which met the full-service criteria. As a result, I used 10000 square feet as the cutoff between "large" grocery stores included in this analysis, and smaller stores excluded from the analysis.

In addition to these stores, I added a number of locations that were not coded as grocery stores, but serve the same role, in the MPROP. Most notably, Super Walmarts and Targets, listed under "department stores" in their land-use codes, but also meet the criteria of a full-service grocery store, were also considered grocery stores for this study. Finally, a smaller number of grocery stores listed under "mixed-use commercial" were also added by changing their land-use codes to that used for grocery stores. For these, I again used a combination of Google searches along with business listings to determine which addresses belonged to grocery stores. I then manually changed their codes as well to match the one for grocery stores.

While I am able to identify the locations of grocery stores throughout my sample period, the data does not provide any description of the name of each store, so with the exception of a small subset of stores found while cross-checking locations, I am generally unable to identify which grocery store is at each location. This prevents me from distinguishing between grocery stores based on some metric of quality of affordability, or identifying stores that close but are replaced by another outlet at the same location. Over the course of the fourteen year sample period, the number of grocery stores within the city varies moderately, with a low of 44 and a high or 55 grocery stores within the city during the study period. In 11 of the 14 years, however, the number of grocery stores observed in the city are between 47 and 50. Figures 2.3 – 2.5

illustrate the location of grocery stores within the city in the years 2015 (min. year), 2010 (max year), and 2003 (49 stores).

To define food access, I draw from the definition of a food desert provided by the USDA, which defines one as a census tract where 33% of households to not have access to a quality food source within one mile of their residence.[28] This definition captures the importance of "walkability" for food access in an urban area, where many households, particularly low-income households, do not own a car. In my hedonic study of how grocery stores are capitalized into home prices, Warsaw (2016) used a modified version of this definition in defining food access as the number of grocery stores in quarter-mile rings up to a one-mile radius of each home sale. In that study, I found evidence of a positive average effect for grocery stores between .25 – .50 miles of a home, and stores between .50 – .75 miles of a home. I also found evidence of a positive effect on home prices for grocery stores within .75 – 1.0 miles of a home in neighborhoods with a combined proportion of African and Latino-Americans above .29.

Drawing from this work, I define food access for this study as the number of grocery stores within one mile of one's home. Additionally, as a robustness check, I estimate the same model using a reduced radius of .75 miles in the definition of food access, as only stores within that range were estimated to have a universally positive average effect on home prices. As with the other amenities in this study, in order to aggregate food access, I average the number of grocery stores within one mile of homes sold within a given neighborhood in a given year. In contrast with the hedonic study, where identifying how grocery stores impacted home prices lent itself to considering access at different distances (similar to the work done by Pope 2012 and Walls et al 2015), the same does not appear to be necessary for this study. When choosing a

[28] http://americannutritionassociation.org/newsletter/usda-defines-food-deserts

neighborhood (as opposed to a house), it is more intuitive to think of home owners considering the number of local grocery stores within a particular neighborhood (or in bordering neighborhoods), as opposed to attempting to determine the number of stores at different distances from the typical home in a given neighborhood. Only after choosing a neighborhood, once households are choosing a specific home, would the locations of local grocery stores likely affect household decisions, in a type of sub-sorting process not modeled here.

In addition to grocery stores, I use the parcel shapefiles and MPROP to identify several other features within the city to address the potential for omitted variable bias. Omitted variable bias in the sorting model arises from structural or neighborhood characteristics that households value in their sorting decisions, which are not included in the estimation process, but correlated with observable features. As a result, most studies control for the presence of local parks, and proximity to the central business district, at a minimum, regardless of the amenity of focus. In addition to these, for this study I include two other controls for commercial land use to address this concern. As a retail location, grocery stores typically appear in close proximity to other retail storefronts. If their presence is not accounted for, it is possible that this variation could be captured within the coefficient for grocery stores. Since commercial could act as either an amenity to households (additional shopping options) or a disamenity (congestion), the direction of the bias is unclear. However, to address this potential source of bias, I include as a variable the distance to the nearest non-grocery commercial parcel.

I also use the average number of bars within one mile of a residence within a neighborhood as a proxy for commercial density within that neighborhood. As discussed in Warsaw (2016), in the context of Milwaukee, bars serve as a useful control for commercial density due to their relative abundance of bars relative to other commercial locations, including grocery stores. In addition, they have a practical advantage over other potential proxies for commercial controls such as coffees shops due to their receiving a unique land-use code within the MPROP, making it feasible to identify them in contrast with other potential controls. Finally, previous research has indicated that they appear in high frequency even in areas with diminished food access (Moorland 2002), providing a direct tie to the research agenda here. Their density in Milwaukee allows them to serve as a useful proxy of commercial density. In Warsaw (2016), I use both bars and the percentage of land within a mile of a home marked as commercial-use as controls for commercial density, and argue that the count of bars is a more robust and useful measure of commercial density in that study. As such I choose to solely use the number of bars as a control for commercial density in this study.

2.4.6 Census Data

To construct neighborhood level-demographic variables for the choice set, along with describe household characteristics, I use publically available census data from the years 2000, and 2007-2014. I define the racial characteristics of each neighborhood using data at a tract level. In particular, I use the combined proportion of African-Americans and Latinos as my control for neighborhood race. While not a perfect match, Census tracts and Milwaukee neighborhoods share a significant level of overlap, allowing for a fairly close mapping of tract characteristics to each neighborhood. Each home sale transaction was spatially joined to its appropriate census tract, and then matched with the Census data of the closest year. From here, I took the average proportion across all homes sold in a neighborhood for a given year as its racial composition. This allows for a slight weighting of the demographics of each census tract, to cover for the imperfect overlap of tracts and neighborhoods.

For the individual household characteristics, I approximate the same process outlined above, with publically available block and block-group Census data. The characteristics used for this study are a household's yearly income, available at the block group level, it size, and whether the household is black or Latino, all of which are available at the block level. Without access to household-level data, these characteristics are averages of those households in its block or block group. In particular, for the race attributes, the averages represent the probability that a household is black or Latino. Because I am aggregating individual household characteristics here, additional assumptions are necessary to ensure unbiased coefficients in the first stage of estimation, in which the heterogeneity parameters are identified. Specifically, it is necessary to assume that households do not jointly select neighborhood demographics at the census block jointly with neighborhood level food access. I argue that this assumption is justified to the vast differences in scale between Census blocks and neighborhoods. The city of Milwaukee contains 190 neighborhoods, as determined by the Milwaukee Neighborhood Identification Project, and overlaps 8,038 census blocks and 690 census block groups. Since the block and block-group geographies are significantly smaller than our neighborhoods, there is still a significant level of demographic variation within out choice set. However, given the unrefined nature of the income variable, it is not used in any budget constraint for the consumer's choice.

Additionally, as Klaiber and Phaneuf (2010) discuss, aggregating household characteristics in this way does reduce the precision of estimates of in the first stage. Consistency could also be a concern in smaller samples, as it increases the possibility that the sampled households will have characteristics that vary greatly from the observed average, though the sample size here is large enough to minimize those concerns.

2.5 Results

2.5.1 First Stage Estimation Results

In the first stage, I estimate the mean indirect utilities for the housing bundles, along with the interaction parameters for heterogeneous housing preferences. The interaction terms are reported in table 2.2. Of particular interest to this research are the interactions between grocery stores and household demographics. Omitted variable bias in this stage poses a concern to accurate estimates of the interaction terms. To account for this, I present several models using varying sets of commercial controls. Without any other commercial controls, grocery stores are estimated to be an inferior good, with a negative and significant interaction term between household income and the number of grocery stores within a neighborhood. With the addition of controls for commercial activity the estimate increases, however, and is positive in the model containing controls both for commercial proximity and the number of bars within a one mile radius. This shift indicates the presence of negative bias in the absence of those controls and the robustness of the claim that grocery stores are a normal good in the broader food access literature.

Further, the models provide significant evidence of heterogeneous preferences for food access by race, particularly for African-Americans, with a positive and significant interaction for these households across all four models presented. In contrast with the income interaction, the parameter decreases in magnitude with the introduction of commercial controls, suggesting that when estimated alone, the interaction with race picks up a broader preference for commercial access. The interaction between grocery stores and Latino is also positive, but also decreases with the introduction of commercial controls, and is no longer significant in the preferred specification. Still, the results corroborate previous research in the food access literature which suggest that households of color, particularly African-Americans, have a stronger preference for food access than their counterparts despite (or in part because of) their elevated levels of poverty, and low access to modes of transportation. Given the levels of income inequality along racial lines within Milwaukee, a point of inquiry is how the marginal utility of grocery stores compares across poorer households of color and richer, non-black or Latino (i.e. white) households. I consider this in more detail below.

The commercial controls are also all significant at the one percent level. However, interpreting their signs are difficult, especially in the case of commercial proximity. The first stage includes three interaction effects, and the second stage includes both a linear and squared term for proximity to the closest non-grocery commercial lot. This combination of variables makes interpreting the economic meaning of the willingness-to-pay for a marginal change in proximity to a commercial lot very difficult to ascertain. In the controls for the average number of bars within a mile of a home, while calculating of the MWTP is easier, the interpretation is complicated by the use of the bars in part as a control for commercial density. The same is the case for the estimates for the interaction terms with bars. As stated earlier, the primary role of these regressors is to serve as a control for commercial density. However, the number of bars within a given neighborhood likely also contains an amenity value which influences household decisions, and thus it is possible that this is conflated with these coefficients as well. Given that the focus of this study is on grocery stores, and the lack of necessary refinement to distinguish the effects with these controls, I refrain from interpreting them here.

Finally, I include an interaction between home size and household size. Intuitively, larger households should have stronger preferences for additional space to accommodate them. In line with this intuition, the coefficients are estimated as positive and significant across all of the specifications.

2.5.2 Second Stage Estimation Results

The second stage estimation decomposes the mean indirect utility into observable and unobservable components as described in equation 7, and recovers the parameters for those observable components. As mentioned in the estimation section, were there no concerns about endogeneity of price, equation 7 could be estimated via OLS. The results of a naïve OLS regression of equation 7 are in table 2.3[29]. Several of the results generated via OLS are counterintuitive, suggesting the possibility of price endogeneity. The price parameter is estimated to be negative but insignificant. Additionally, the parameter for home size is negative and significant. Even when considering the positive interaction term for home size, the overall estimate for home size remains negative for all households of three or fewer people. In addition, both lot size and the number of bathrooms are estimated to have negative marginal utilities. While it is not uncommon for estimates for lot size coefficients in hedonic and housing choice models to report as negative, this typically only occurs when a home size term is estimated as positive and significant (see, for instance, Klaiber and Phaneuf 2010 or Warsaw 2016).
These results point to the likely existence of endogeneity in alternative prices. As such, I follow equations (11) – (13) in section 3.2 to construct an instrument for price. In doing so, I introduce additional alternative characteristics for neighborhoods within half mile rings up to two miles. For this process I include all of the variables used in the naïve OLS regression. I report the second stage regression IV regression results in table 2.4. While I do not include the first stage results, of note is the high R-squared, 0.9040, underscored by a very high correlation between the

[29] Full results for regressions in this section are included in the appendix

constructed instrument and actual alternative prices of 0.9389. This indicates that the constructed instrument meets the criteria of relevance as it relates to alternative price.

The second stage IV estimation produces results much more in line with intuition and the existing literature on preferences for homes and neighborhoods. The estimated parameter for price remains negative but increases in magnitude by a factor of nearly eight. In addition, the estimates for the structural housing characteristics more in line with previous research on home choice. The coefficient on home size now reports as positive and significant. Several of the non-size related home characteristics, such as the existence of a basement or air conditioning, also report as positive and significant.

The parameters for grocery stores reports as negative and significant at the one percent level. However, along with the term for home size, it cannot be considered in isolation when considering householder WTP for these attributes. Instead, they must be considered in conjunction with the interaction terms, and thus household composition, which I consider in the next section.

2.5.3 Heterogeneity in Household Preferences

Table 2.5 illustrates the marginal WTP for grocery stores, house size, and bedrooms of six representative families in Milwaukee. As mentioned in the discussion of the first stage results, I do not consider the WTP for the commercial controls, proximity to the closest non-grocery commercial lot or the average number of bars within a mile of a home. The WTP for an additional square food of home size ranges from $49.72 to $61.74 for households of two and five people, respectively.

Of particular interest for the current project is the WTP for additional grocery stores within 1 mile of a home. Representative families 2, 4, and 6 examine the WTP for an additional grocery store for a household likely to be black, Latino, and “white” (neither black or Latino), respectively, at approximately the mean income. This comparison reveals a significant gap in the MWTP among these households. Black and Latino households have a MWTP of $5,931 and $2,328, respectively, compared to a MWTP of $1,748 for white households at the same level of income. Since the racial characteristics used here are an expectation, the reported MWTP are subject to a level of arbitrariness. However, it is worth noting that the number of observed households that met the criteria for the “white” household (less than five percent expectation of being black and less than five percent expectation of being Latino) is 6,253 (of 32,431 observed households), a number which increase to 9,888 when considering all households with less than a 10 percent chance of being black or Latino. Similarly, the number of families with at least a 60 percent chance of being black is 6,547, again representing a significant portion of the sample, suggesting that these profiles are fairly representative families.

For white households earning $90,000, at the upper end of observed income, the MWTP for white households rises to $5,429, which, depending on the construction of an “African-American” households, is roughly equivalent to such a household earning significantly less. For instance, note that of those observed households earning $20,000 a year or less, over 71% lived in census blocks where 80 percent or higher of households are African-American. Comparing the MWTP of the white household earning $90,000 to as African-American household with the profile outlined above, the MWTP for the white household is slightly less than the $5,993 MWTP of their African-American counterpart. Again, one must consider the use of expectations here, as along the coarseness of the income variable. Many households exist in Milwaukee that earn over $90,000, and adjusting either the income or racial expectations can increase this number. That said, these results provide a compelling picture suggesting that households of color do in fact have strong preferences for food access. These results provide evidence suggesting that the lack of quality food outlets in poor neighborhoods, particularly poor neighborhoods of color, is not a reflection of a lack of demand or preferences for them. Figure 2.3 also illustrates the MWTP based on the local demographics of households that have chosen that neighborhood in 2005. As expected, those neighborhoods with a higher concentration of African-American households have a much higher than their counterparts throughout the city. At the highest, the MWTP reaches $8,077 in one neighborhood, though with the level of aggregation here this result should be interpreted with extreme caution.

Further, these results help provide some context as to the actual causes of food insecurity, particularly in the context of the broader literature. On the supply-side, firms are constrained in the number of locations they can choose from by a number of factors, such as the availability of land and zoning restrictions. In addition, the results here suggest that grocery stores are in fact a normal good. Thus, if firms primarily (or only) consider income in their location decisions, they are likely to locate in areas of higher income for their stores, areas which are likely to attract other forms of retail for similar reasons, as indicated by past research (see, for instance, Scheutz et al 2012). Regardless of how one characterizes a “white” family in this context of using expectations, the average income for such a family is significantly above the average income of the city as a whole. This higher income (and, by extension, likely higher budget constraint), allows them to sort in areas with a higher level of their desired public goods (such as school quality or other urban amenities), areas which are likely to have grocery stores for the aforementioned reasons. Poor households, which in urban spaces typically consist of households of color (such as in Milwaukee), are then unable to afford to purchase homes in these areas despite a stronger preference for grocery stores, likely due to the levels of other amenities increasing the price. However, these results suggest the potential for significant welfare improvements by increasing food access in low income areas, a possibility I consider in section seven.

2.6 Robustness Checks

2.6.1 Excluding Recession Year Sales

One concern for the robustness of the results presented here are the potential influence of the recession year sales. The primary includes the recession year sales, but it is possible that the volatility in the housing market could obscure the underlying preferences driving locational decisions. Thus, it is useful to compare these results to a model without the inclusion of recession years, to determine if any significant shift occurred during that time. Tables 2.6-2.9 contain these results. Qualitatively, the results are identical in both the first and second-stage estimation. The MWTP for an additional grocery does rise for each of the representative families with the exception of the highest earning “white family”, which sees a drop in MWTP from approximately $5000 to $4000. The MWTP for a square feet also rises for each of the representative families. These results suggest the slight possibility of a drop in MWTP during the recession years, although not uniformly so. However, the qualitative consistency, combined with fairly minor shifts in MWTP suggests that the model is robust even to the inclusion of recession year sales.

2.6.2 Defining food access at .75 miles

For this study, I chose to focus on grocery stores within 1 miles of a home in defining food access, following the convention used by other scholars and organizations within the literature. However, in his hedonic analysis of implicit prices for grocery stores in Milwaukee, Warsaw (2016) finds that only grocery stores within .75 miles of a home have a universally positive effect on home prices. From .75 to 1 mile, grocery stores only increase home prices in neighborhoods with a combined black and Latino proportion of .29. As I estimate the model again, comparing all four models in the first stage, and the preferred specification for the second stage, only considering grocery stores (and the corresponding control for bars) within a .75 radius of each home. Tables 2.10-2.13 present the results. Qualitatively, the results follow the same pattern as the previous two models. However, in contrast with those models, here the interaction term between Latino and grocery stores is estimated to be both positive and significant in the preferred specification.

The second stage results also mirror the previous models qualitatively, though the larger and significant interaction in the first stage leads to different results in the estimation of MWTP for an additional grocery store. For the representative African-American households, the MWTP for an additional grocery store increases but remains close to $6000. However, the MWTP increases significantly for Latino households, to approximately $3000 for the two families characterized. In addition, the estimated MTWP for “white” households fall, to less than $500 for a family earning approximately the mean wage, and to only $4000 for a family earning $90000 a year.

2.7 Counterfactuals and Welfare Analysis

One of the primary advantages of the sorting model over the hedonic model is the ease with which it can accommodate welfare analysis of alternative changes. In particular, the sorting model here naturally accommodates both partial and general equilibrium analysis, while the hedonic model is typically only able to perform partial equilibrium analysis. In both partial and general equilibrium, the measure used to determine the welfare benefits is the same:

$$\Delta E\left(CS^i\right) = \frac{1}{\alpha_p}\left[\ln\left(\sum_h e^{V_h^1}\right) - \ln\left(\sum_h e^{V_h^0}\right)\right] \quad (14)$$

where V_h^0 represents initial utility, and V_h^1 represents utility after the policy. What changes from partial to general equilibrium is the calculation of the V_h^1. In partial equilibrium, households are not allowed to resort into new locations after the policy. As such the new utility is simply calculated using the updated value of the affected amenity, and welfare is calculated using equation (14). In general equilibrium, houses are allowed to choose a new location after the policy shift. As such, it is necessary to calculate a new equilibrium price to satisfy the market clearing condition. This is accomplished using a similar process to that outlined in the generation of the price instrument. Initially prices are assumed to be constant, and utility is calculated for all households with the updated amenity levels. Then, using equation (13), the shares are calculated for all of the housing alternatives. For those alternatives where this calculated share exceeds the observed market share (i.e. demand greater than supply), the price is increased slightly, and it is decreased slightly for those where the calculated share is less than the observed share (supply greater than demand). The process is then repeated with the updated price until the calculated shares converge on the observed shares.

For this study I compare a three counterfactual scenarios of increasing food access, two on income and one on race. For the first, I increase the number of grocery stores neighborhoods with a combined proportion of African-Americans and Latinos above .75. In the second, I increase the number of grocery stores neighborhoods with an average household income below $30000. Finally, I increase the number of grocery stores in the neighborhoods with a combined income above $60000. Table 2.14 presents the both the partial and general equilibrium welfare measures for the three policies, and table 2.15 highlights the price changes to the housing alternatives in general equilibrium.

As expected, the welfare estimates are similar for the low income and race policies. In both cases the partial and general equilibrium welfare for those directly affected by the policy is significantly greater than for those. In addition, the general equilibrium average welfare is higher than the partial equilibrium welfare measure, indicating that the decrease in prices in those neighborhoods not affected by the policy offsets the increase in prices of those that were, illustrated in part by the drop in average price of an alternative for these policies. This reflects the conflation of race and income in Milwaukee, a conflation which occurs in most urban spaces with unequal food access. Figures 2.4 and 2.5 illustrate the neighborhood affected by the race policy and the income policy, respective. Of note is the overlap between the neighborhoods affected by both counterfactual changes. These results also highlight the valuable nuance of considering both income and race simultaneously when considering questions inequality in public good provision. Under the assumption that grocery stores are correctly identified as a normal good, there would be no way to generate this result without a consideration of race, as poor families (who are largely households of color) would necessarily have a lower MWTP for an additional grocery stores, unless they were incorrectly identified as an inferior good. Likewise, the inclusion of race but not income misses variation within races, as not all households of color in urban spaces are poor.

The policy targeting high-income households presents a starkly different narrative than the other alternatives. Despite affecting a comparable number of alternatives, the number of households affected and the welfare measures, both in partial and general equilibrium, are significantly lower than those policies targeting low-income neighborhoods or those with a high concentration of households of color. As with the first two policies, these results likely reflect the conflation of race and income in Milwaukee. Of those households earning over $60,000, the majority have a less than a 10 percent expectation to be either black or Latino, and over 96 percent of these households have a less than 50 percent expectation to be black or Latino, a reversal of the structure of low-income households. By considering race and income in conjunction, this allows the estimated impact of an additional grocery store in high-income neighborhoods, which are largely “white”, to be positive, but also dwarfed by the impact of policies targeted at low-income, largely black and Latino, neighborhoods, or policies targeting race directly. The analysis presented here not only highlights the importance of considering location-specific policies in urban planning, but one that consider the demographics of those impacted on multiple planes.

2.8 Conclusion

Across the social sciences, understanding the nature and consequences of declining food access in urban spaces, along with investigating potential remedies, remains a top priority. Given the correlation between limited food access and ill-health effects, such as high morbidity and obesity rates, the presence of food insecurity in urban spaces is of significant concern to public health scholars. These concerns connect this body of research with a broader literature investigating the cause and interconnectedness of social injustices more broadly (see, for instance, Jyoti et. al 2001). However, an open point of contention to this date has been whether low food access in poor urban areas are a reflection of consumer preferences, especially on the part of poor households of color.

In this paper, I used a residential sorting model to estimate consumer preferences for large, full service grocery stores in their neighborhoods. My results show that while grocery stores are indeed a normal good, households of color, particularly African-America households, also have strong preferences for food access in their neighborhoods, with a MWTP for an additional grocery store that even eclipses white households in much higher income brackets. I then used these estimates to perform a number of simple counterfactual examples to highlight the heterogeneity in welfare impacts of placing grocery stores in different neighborhoods by income and race.

Together, these results help quantify the benefit of increasing food access in low income communities, especially those with a large minority population. While *ability* to pay remains a concern, these results suggest a significant level of demand for food access in these neighborhoods, and supports the significant existing literature suggesting that food inequalities stem not from lack of preferences on the part of those affected, but of larger structural inequalities which must be addressed in order to remedy food inequality.

This current project represents one of the first attempts to use residential markets to address questions of food access. The results presented here could be extended in several important ways. First, obtaining more refined data on household characteristics would open several avenues of research unavailable here. For instance, the inclusion of household-level race characteristics would enable the estimation of household preferences for the racial composition of their neighbors, as seen in work such as Bayer and McMillan (2012), alongside preferences for food access. This would remove one of the primary sources of potential bias not addressed in this project. In addition, obtaining household-level income data would allow for the creation of a budget constraint for use in the second stage, limiting the homes which households could choose after an exogenous change to the amenity space in a residential market. This would bring the tension of gentrification in urban renewal into focus in developing solutions to unequal access to public goods in urban spaces. One concern many scholars have in addressing food insecurity or other similar inequalities is that many policies would cause housing prices to rise, pushing the poor households the policies were designed to help out of their homes. This logic could be seen in how home prices rose in those neighborhoods that received an additional grocery store. While the impact was not very large, adding additional grocery stores in addition to other public goods in limited quantities could lead prices to rise to a point where homes are no longer able to remain.

Future research might also consider the tradeoffs between different land-use types in consumer housing decisions. In this paper, because I only considered preferences for grocery stores, my counterfactuals look at the impact of an additional grocery store in isolation. However, in most urban areas land is very limited. As such, the rise of a new commercial lot, especially one as large as a full-service grocery stores, will often require the sacrifice of other amenities, such as other retail options or open space. These changes then potentially add an environmental component to the trade of, in the loss of limited open space and an increase in carbon emissions. Being able to characterize heterogeneous consumer preferences for both urban and environmental amenities would allow for a richer and more complete assessment of the impact of altering urban structure.

Finally, the model could be extended to consider a wider class of public goods in urban spaces. While research in environmental economics has begun to consider how a wider class of environmental amenities affect sorting decisions (such as heterogeneous open space types), work within urban economics has continued to focus primarily on school quality as a primary driver of household decisions. While school quality remains a key driver of household decisions, the results here also show the value of considering a broader class of urban amenities, such as access to different types of commercial properties. This logic likely extends to other urban amenities, such as local environmental quality (local air quality for instance). Extending the model to consider these amenities would provide valuable tools to urban planners and policy makers to simultaneously consider the distributional impacts of policies alongside its other effect

2.9 References

Bayer, P., Ferreira, F. and McMillan, R., 2007. A unified framework for measuring preferences for schools and neighborhoods. Journal of political economy 115(4),.588-638

Bayer, P., McMillan, R., 2012. Tiebout sorting and neighborhood stratification. Journal of Pubilc Economics 96, 1129-1143

Bayer, P., McMillan, R. and Rueben, K.S., 2004. What drives racial segregation? New evidence using Census microdata. Journal of Urban Economics 56(3), 514-535

Bayer, P. and Timmins, C., 2005. On the equilibrium properties of locational sorting models. Journal of Urban Economics, 57(3), 462-477

Bayer, P., Timmins, C., 2007. Estimating Equilibrium Models of Sorting across Locations. The Economic Journal 117(518), 353-374

Benabou, R., 1993. Search market equilibrium, bilateral heterogeneity, and repeat purchases. Journal of Economic Theory, 60(1), 140-158

Benabou, R., 1996. Heterogeneity, stratification, and growth: macroeconomic implications of community structure and school finance. The American Economic Review 86(3), 584-609

Berry, S.T., 1994. Estimating discrete-choice models of product differentiation. The RAND Journal of Economics, 242-262

Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. Econometrica: Journal of the Econometric Society 63(4), 841-890

Bitler, M., Haider, S.J., 2011. An economic view of food deserts in the United States. Journal of Policy Analysis and Management 30(1), 153-176

Chay, K.Y. and Greenstone, M., 2005. Does air quality matter? Evidence from the housing market. Journal of political Economy, 113(2), pp.376-424

Chung, C., Myers, S.L., 1999. Do the poor pay more for food? An analysis of grocery store availability and food price disparities. The Journal of Consumer Affairs 33(2), 276-296

Dong. D. Lin, B., 2009. Fruit and Vegetable Consumption by Low-Income Americans. USDA Economic Research Report Number 70

Eisenhauer, E., 2001. In poor health: Supermarket redlining and urban nutrition. GeoJournal 53(2), 125-133

Elbel, B., Moran, A., Dixon, L.B., Kiszko, K, Cantor, J., Abrams, C., Mijanovich, T. Assessment of a government-subsidized supermarket in a high-need area on household food availability and children's dietary intake. Public Health Nutrition 18)15), 2881-2890

Epple, D., Filimon, R., Romer, T., 1993 Existence of voting and housing equilibrium in a system of communities with property taxes. Regional Science and Urban Economics 23, 585-610

Epple, D., Sieg, H., 1999. Estimating equilibrium models of local jurisdictions. Journal of Political Economy 107(14), 645-681

Epple, D. and Zelenitz, A., 1981. The implications of competition among jurisdictions: does Tiebout need politics? The Journal of Political Economy, 1197-1217

Giang, T., Karpyn, A., Laurison, H.B., Hillier, A., Perry, R.D., 2008. Closing the grocery gap in underserved communities: the creation of the Pennsylvania fresh food financing initiative. Journal of Public Health Management and Practice 14 (3), 272–279

Glanz, K., Sallis, J.F., Saelens, B.E., Frank, L.D., 2007. Nutrition environment measures survey in stores (NEMS-S). American Journal of Preventive Medicine 32 (4), 282–289

Hendrickson, D., Smith, C., Eikenberry, N., 2006. Fruit and vegetable access in four low-income food deserts communities in Minnesota. Agriculture and Human Values 23, 371–383

Johnson, K., Percy, S., Wagner, E., 1996. Comparative study of food pricing and availability in Milwaukee. A report prepared for The Food System Assessment Project of The Hunger Task Force. Milwaukee, WI: Urban Research Center, University of Wisconsin-Milwaukee

Jyoti, D.F., Frongillo, E.A. and Jones, S.J., 2005. Food insecurity affects school children's academic performance, weight gain, and social skills. The Journal of nutrition, 135(12), 2831-2839

Kaufman, P.R., MacDonald, J., Lutz, S., Smallwood, D., 1997. Do the Poor Pay More for Food? Item Selection and Price Differences Affect Low-Income Household Food Costs. AER 759, USDA-ERS

Klaiber Jr, H.A., 2008. Valuing open space in a locational equilibrium model of the Twin Cities. North Carolina State University. Dissertation

Klaiber, H.A., Phaneuf, D.J., 2010. Valuing open space in a residential sorting model of the Twin Cities. Journal of Environmental Economics and Management, 60(2), 57-77

Massey, D.S. and Denton, N.A., 1989. Hypersegregation in US metropolitan areas: Black and Hispanic segregation along five dimensions. Demography, 26(3), 373-391

McFadden, D., 1978. Modeling the choice of residential location. Transportation Research Record 673, 72-77

McMillen, D.P., 2001. Polycentric Urban Structure: The Case of Milwaukee. Economic Perspectives, Federal Reserve Board of Chicago 25(2): 15-27

Morris, P.M., Neuhauser, L. Campbell, C. 1992. Food security in rural America: a study of the availability and costs of food. Journal of Nutrition Education 24(1), 525-585

Pope, D.G., Pope, J.C., 2012. Crime and property values: evidence from the 1990's crime drop./ Regional Sciences & Urban Economics 42(1), 177-188

Pothukuchi, K., 2004. Community Food Assessment: A First Step in Planning for Community Food Security. Journal of Planning Education and Research 23, 356-377

Powell, L.M., Slater, S., Mirtcheva, D., Bao, Y., Chaloupka, F.J., 2007. Food store availability and neighborhood characteristics in the United States. Preventive Medicine 44, 189–195

Schuetz, J., Kolko, J., Meltzer, R., 2012. Are poor neighborhoods "retail deserts"? Regional Science and Urban Economics 42, 269-285

Short, A., Guthman, J., Raskin, S., 2007. Food deserts, oases, or mirages? Small markets and community food security in the San Francisco Bay Area. Journal of Planning Education and Research, 26(3), 352-364

The New York Law School Racial Justice Project, 2006. Unshared Bounty: How Structural Racism Contributes to the Creation and Persistence of Food Deserts.

Tiebout, C.M., 1956. A pure theory of local expenditures, Journal of Political Economy 64, 416–424

Walls, M., Kousky, C., Chu, Z. Is What You See What You Get? The Value of Natural Landscape Views. Land Economics 91(1), 1-20

Warsaw, P., 2016. The economics of food security: investigating price inequalities for food access in Milwaukee, Wisconsin. Working Paper

Weatherspoon, D., Oehmke, J., Dembele, A., Weatherspoon, L., 2015. Fresh vegetable demand behavior in an urban food desert. Urban Studies 52(5), 960-979

Zenk, S.N., Schulz, A.J., Israel, B.A., James, S.A., Bao, S., Wilson, M.L., 2005. Neighborhood racial composition, neighborhood poverty, and the spatial accessibility of supermarkets in Metropolitan Detroit. American Journal of Public Health 95 (4), 660–667

Zenk, S.N., Schulz, A.J., Israel, B.A., James, S.A., Bao, S.M., Wilson, M.L., 2006. Fruit and vegetable access differs by community racial composition and socio-economic position in Detroit, Michigan. Ethnicity & Disease 16, 275-280

Zepeda, L., Chang, H., Leviten-Reid, C., 2004. Organic food demand: A focus group study involving Caucasian and African-American shoppers

2.10 Tables and Figures

Figure 2.1: Spatial element of the choice set (Milwaukee neighborhoods)

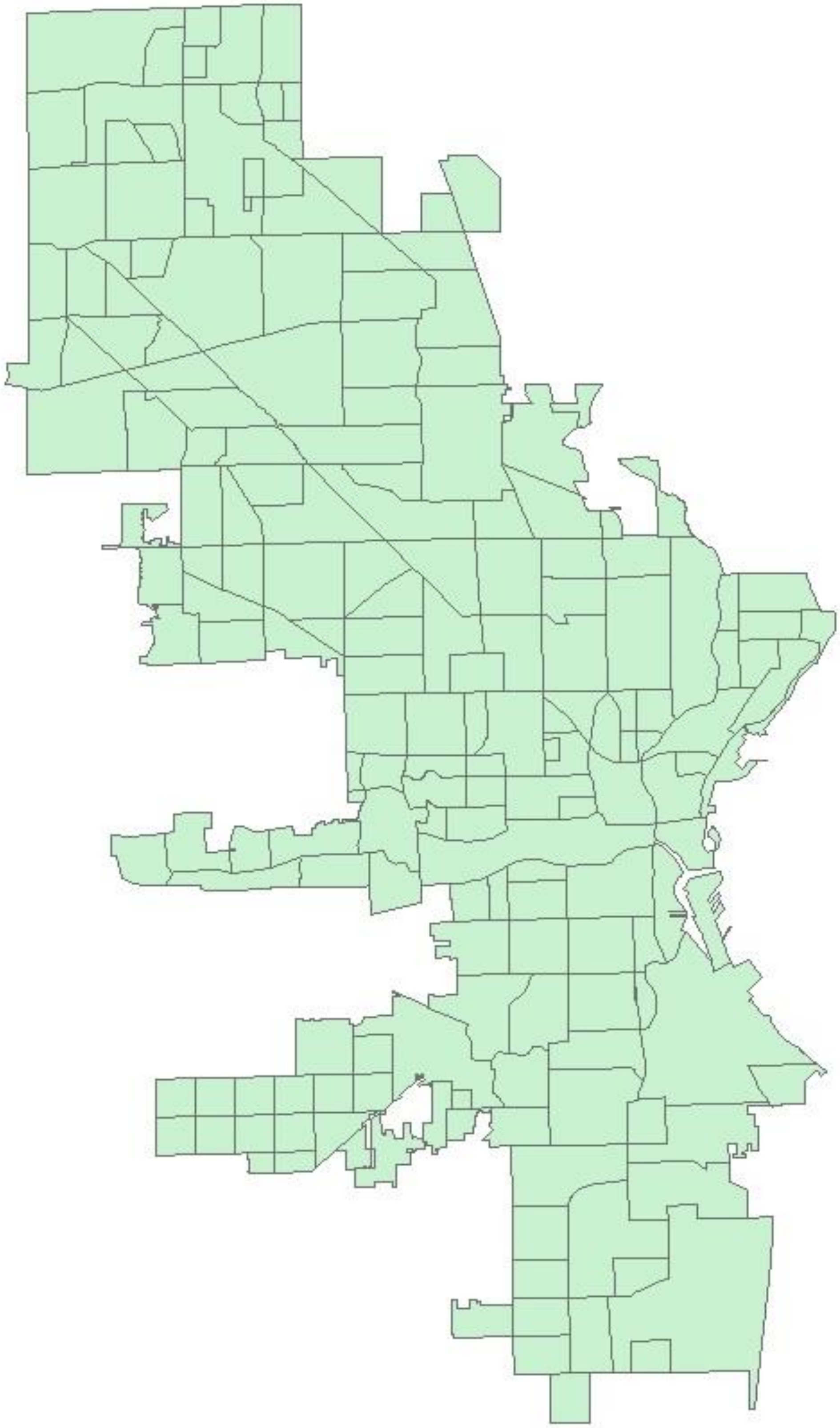


Figure 2.2: Distribution of selected neighborhood attributes

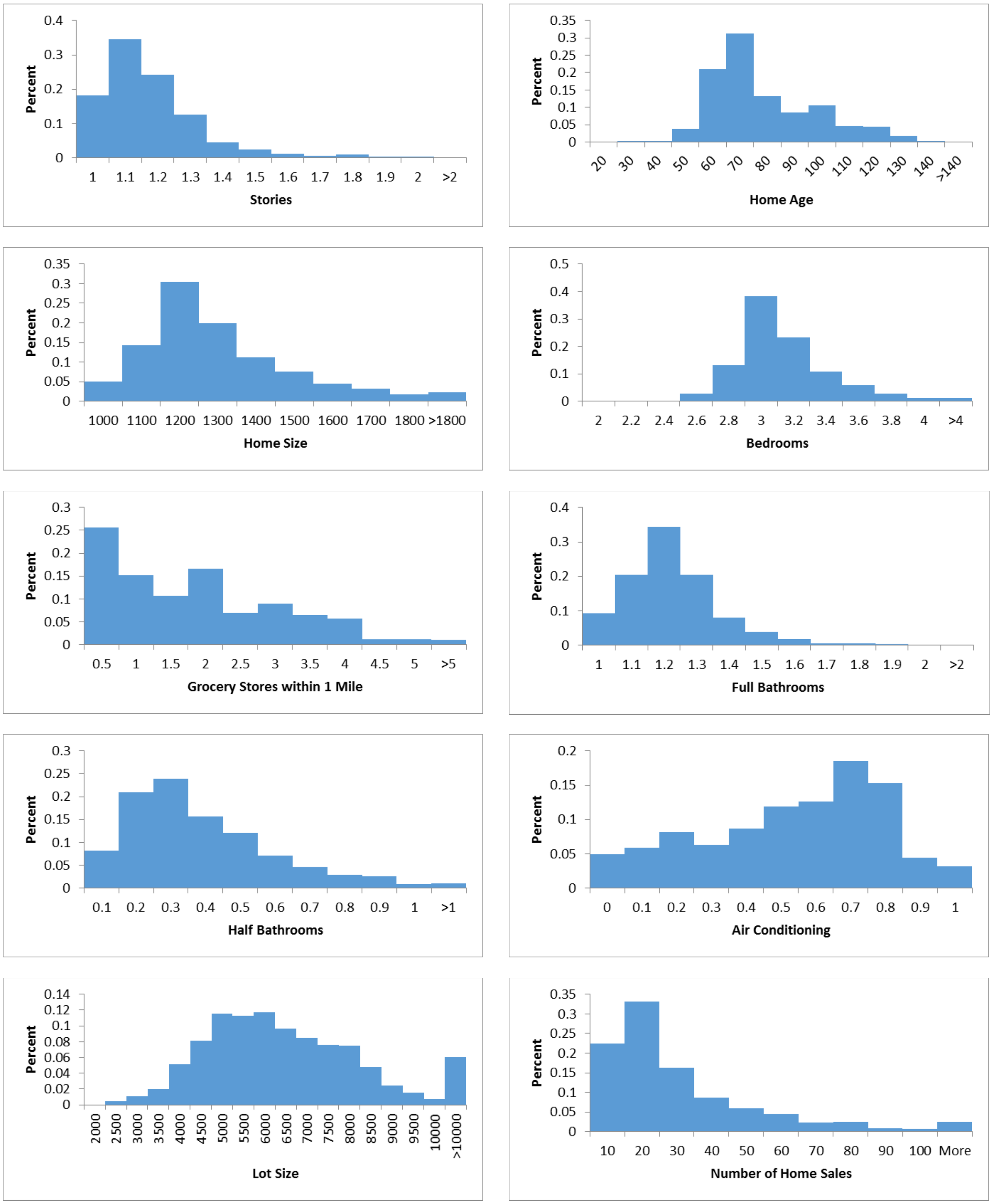


Table 2.1: Summary statistics

Variable Name	Mean	Std. Dev.	40th percentile
Household Characteristics (I = 32,421)			
Income (0000's)	4.193	1.345	3.789
Black (Expectation)	0.251	0.166	0.028
Latino (Expectation)	0.104	0.320	0.030
Household Size	2.587	0.556	2.38
Alternative Variables (N = 1199)			
Price (000's)	108.520	32.232	77.818
Stories	1.139	0.156	1.02
Home Age	73.592	19.250	59
Home Size	1253.351	209.984	1102.215
Bedrooms	3.055	0.300	2.838
Full Bathrooms	1.186	0.149	1.077
Half Bathrooms	0.340	0.215	0.169
Lot Size	6316.207	2112.268	4649.111
Fireplaces	0.155	0.171	0
Air Conditioning	0.584	0.258	0.321
Powder Rooms	0.336	0.214	0.162
Basement	0.974	.0621	0.958
Attic	0.383	0.268	0.1
Attached Garage	0.109	0.174	0
Detached Garage	0.779	0.216	0.636
Groceries within 1 mile	1.590	1.307	0.067
Bars within 1 mile	2.539	2.611	0.543
Black + Latino proportion (census tract)	0.326	0.292	0.052
Distance to the CBD (000s feet)	25.745	9.984	16.536
Distance to closest commercial lot (000s ft)	0.604	0.829	0.342
Distance to the nearest park (000s feet)	1.340	0.833	0.730

Figure 2.3: Location of Milwaukee Grocery Stores in 2015

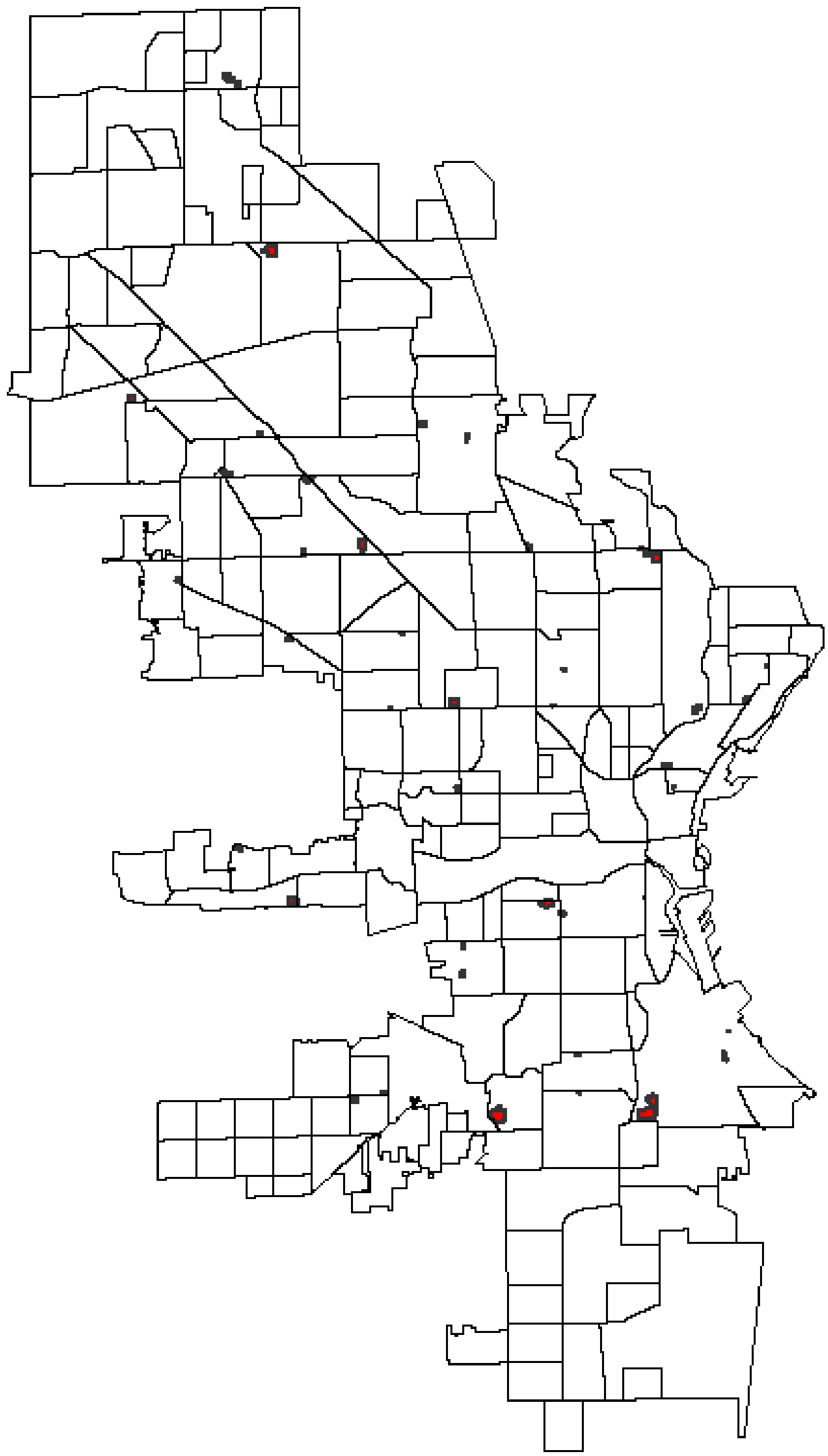


Figure 2.4: Location of Milwaukee Grocery Stores in 2010

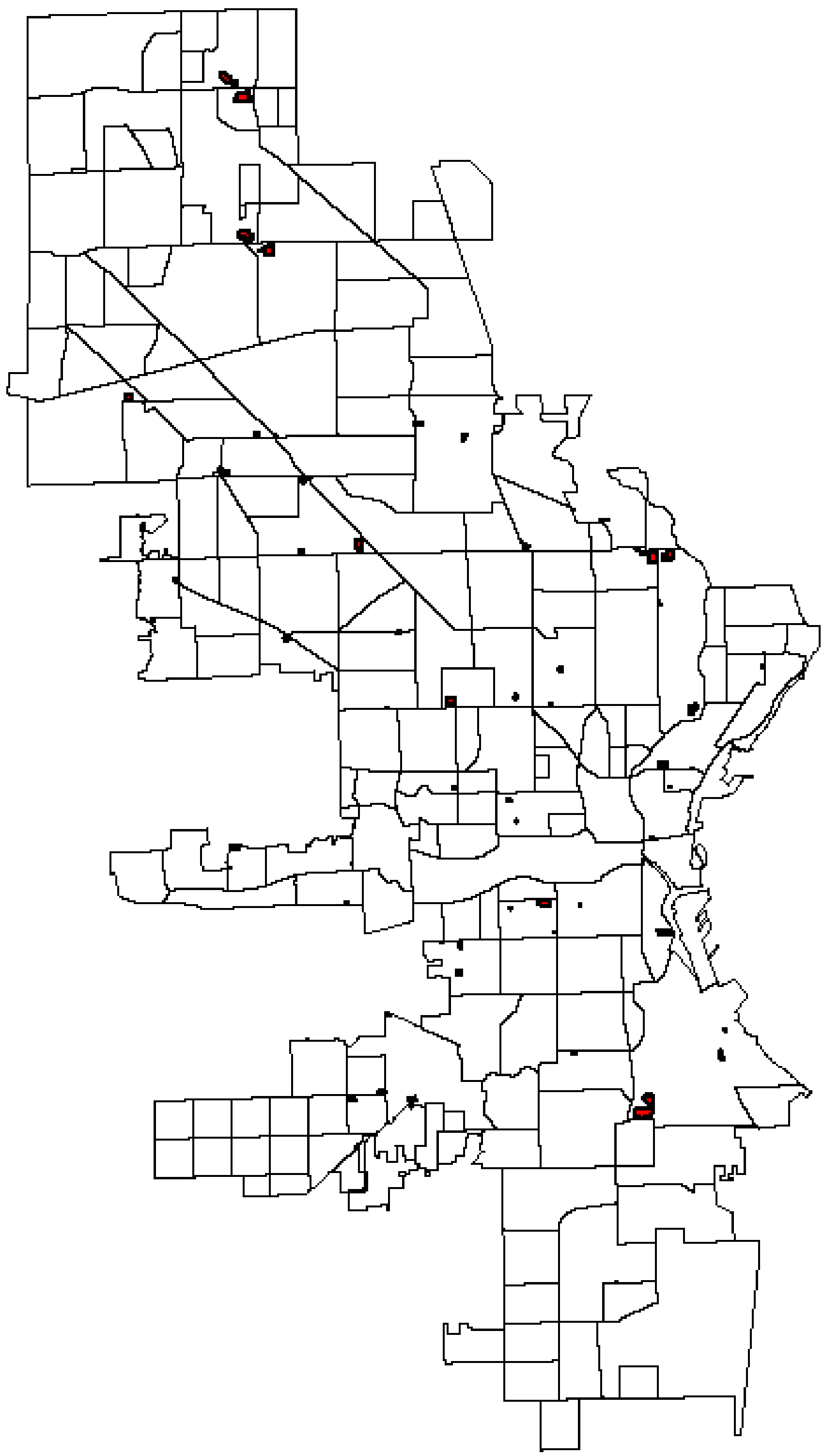


Figure 2.5: Location of Milwaukee Grocery Stores in 2003

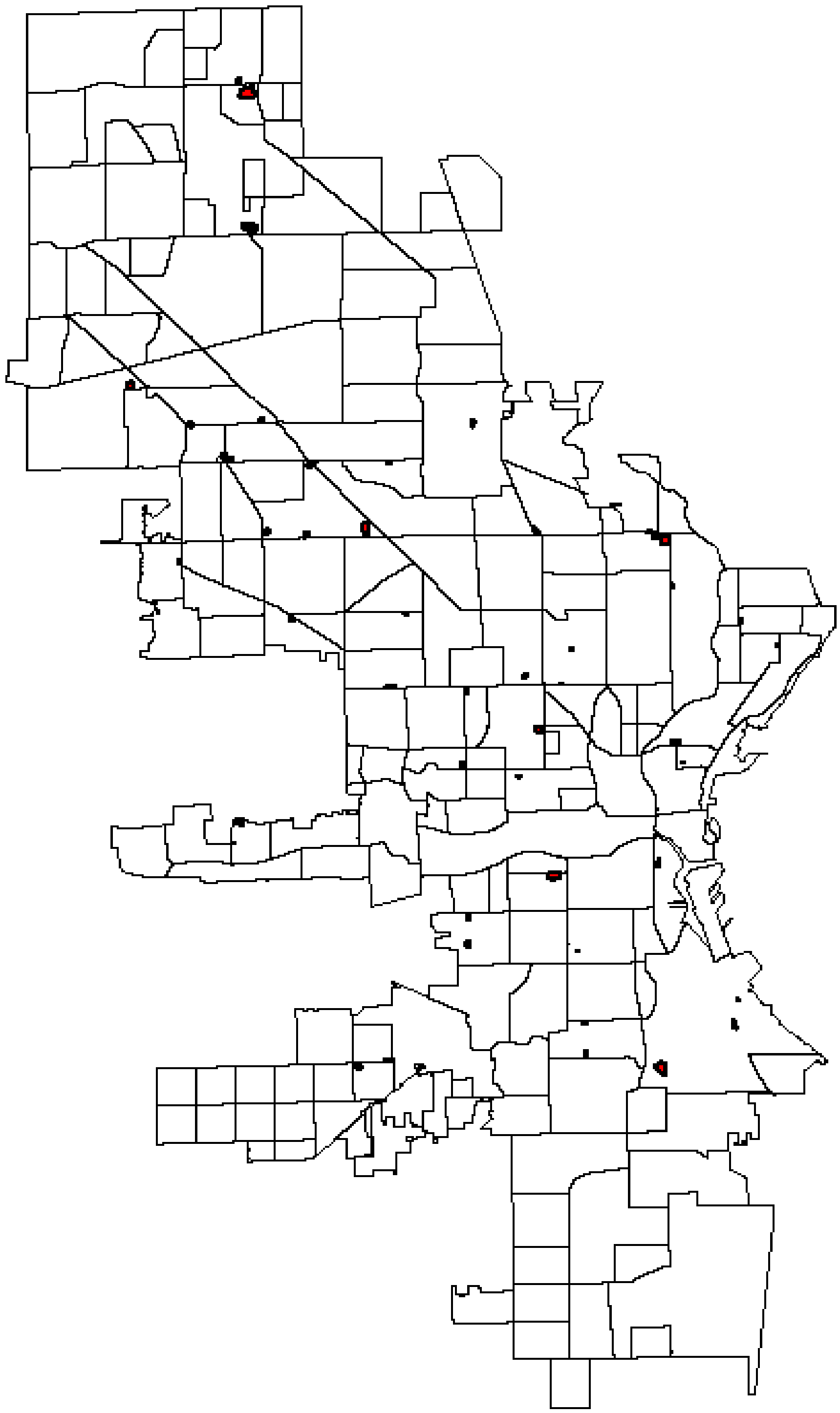


Table 2.2**:** First stage results

Variable Name	Model 1	Model 2	Model 3	Model 4
House Size x Family Size	$7.3811 \cdot 10^{-4}$*** ($5.0642 \cdot 10^{-5}$)	$3.7278 \cdot 10^{-4}$*** ($6.1874 \cdot 10^{-5}$)	$3.6615 \cdot 10^{-4}$*** ($5.9862 \cdot 10^{-5}$)	$2.7855 \cdot 10^{-4}$*** ($6.2588 \cdot 10^{-5}$)
Number of Grocery Stores x Income[1]	-0.0726*** (0.0015)	$-7.9640 \cdot 10^{-4}$*** (0.0016)	0.0377*** (0.0017)	0.0517*** (0.0018)
Number of Grocery Stores x Latino	0.9267*** (0.0255)	0.4030*** (0.0327)	0.3245*** (0.0299)	0.0310 (0.0339)
Number of Grocery Stores x Black	0.8962*** (0.0124)	0.7921*** (0.0128)	0.6140*** (0.0136)	0.5364*** (0.0137)
Number of Bars x Income			-.1343*** (0.0014)	-0.1015*** (0.0014)
Number of Bars x Latino			1.0475*** (0.0118)	0.6927*** (0.0131)
Number of Bars x Black			0.5002*** (0.0081)	0.4682*** (0.0085)
Distance to nearest Comm x Income		0.7559*** (0.0051)		0.5853*** (0.0052)
Distance to nearest Comm x Latino		-14.2287*** (0.2053)		-11.7028*** (0.2334)
Distance to nearest Comm x Black		-2.6758*** (0.0809)		-1.2190*** (0.0849)

***Significant at a 1% level

[1] Income in $10000s

Table 2.3: Naïve OLS second stage results

Variable Name	Coefficient	Standard Error[a]
Price	-0.0035	0.0040
Stories	-0.0757	0.3681
Home Age	0.0472	0.0152***
Home Age^2	-0.0004	$9.45 \cdot 10^{-5}$***
Home Size	-0.0008	0.0005
Bedrooms	-0.1322	0.2079
Full Bathrooms	-0.2856	0.2817
Half Bathrooms	0.2399	0.5556
Lot Size	-0.0001	$3.32 \cdot 10^{-5}$***
Fireplace	0.2580	0.2661
Air Conditioning	-0.0411	0.2578
Powder Room	-0.4392	0.6190
Basement	-0.1085	0.4748
Attic	0.2465	0.2474
Attached Garage	0.8770	0.3938**
Detached Garage	1.2299	0.3093***
Proportion of Census tract which is Black or Latino	-0.3829	0.2973**
Distance to the CBD (000's ft)	-0.0500	0.0195**
Distance to the CBD^2	0.0008	0.0003***
Distance to the nearest park (000's ft)	-0.0191	0.1031
Distance to the nearest park^2	0.0094	0.0183
Distance to the nearest non-grocery commercial lot (000's ft)	-0.8602	0.8187
Distance to the nearest non-grocery commercial lot^2	-0.5302	0.2965*
Groceries within 1 mile	-0.1513	0.0370***
Bars within 1 mile	0.0488	0.0361

* significant at 10%, ** 5%, and *** 1% level

[a] Standard Errors clustered at the neighborhood level

Table 2.4: Second stage results using IV regression

Variable Name	Coefficient	Standard Error[a]
Price (000s dollars)	-0.0701	0.0137***
Stories	0.6498	0.5812
Home Age	0.0221	0.0257
Home Age^2	-0.0003	0.0002
Home Size	0.0029	0.0011***
Bedrooms	-0.9724	0.3254***
Full Bathrooms	0.3607	0.5311
Half Bathrooms	1.0401	1.0660
Lot Size	$4.81 \cdot 10^{-5}$	$5.10 \cdot 10^{-5}$
Fireplace	2.0274	0.5182***
Air Conditioning	1.3126	0.4724***
Powder Room	-0.6254	1.1414
Basement	2.2439	0.6996***
Attic	-0.2080	0.3398
Attached Garage	1.0803	0.7444
Detached Garage	2.1122	0.6409***
Proportion of Census tract which is Black or Latino	-4.6105	0.8781***
Distance to the CBD (000's ft)	-0.1119	0.0543**
Distance to the CBD^2	0.0016	0.0007*
Distance to the nearest park (000's ft)	-0.0067	0.1383
Distance to the nearest park^2	-0.0026	0.0338
Distance to the nearest non-grocery commercial lot (000's ft)	-0.2295	0.9455
Distance to the nearest non-grocery commercial lot^2	-0.7064	0.3317**
Groceries within 1 mile	-0.1124	0.0600*
Bars within 1 mile	0.0462	0.0514***

* Significant at 10%, ** 5%, and *** 1% level

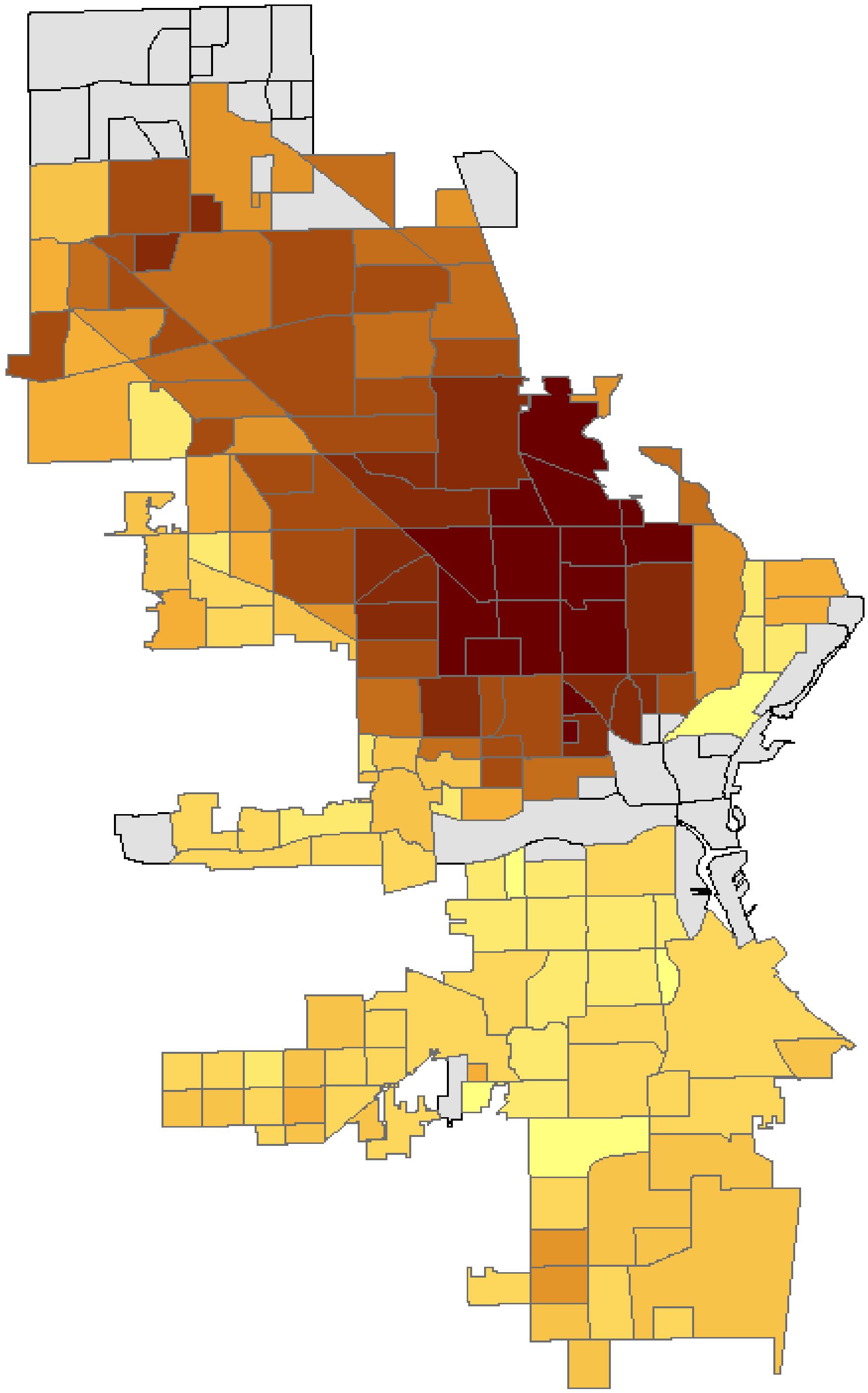
[a] Standard Errors are clustered at the neighborhood level

Table 2.5: Heterogeneity in MWTP for housing and neighborhood characteristics[1]

	Family Characteristics						
Characteristic	Mean	1	2	3	4	5	6
Annual Income	41,931	20,000	40,000	30,000	40,000	90,000	40,000
Black (Expectation)	0.251	.8	.6	0	.1	.05	.05
Latino (Expectation)	0.104	0	0	.6	.5	.05	.05
Household Size	2.587	3	2	3	2	5	4
Variable (Change)	Marginal Willingness-to-pay						
Square Feet (1 additional sq. ft.)	$52.15	$53.79	$49.82	$53.79	$49.82	$61.74	$57.76
Grocery Stores (1 additional store)	$3,455	$5,993	$5,931	$871	$2,328	$5,429	$1,748
Basement	$31,989	$31,989	$31,989	$31,989	$31,989	$31,989	$31,989
Attached Garage	$15,401	$15,401	$15,401	$15,401	$15,401	$15,401	$15,401

[1] Willingness-to-pay measures represent a one-time payment made by households for the amenity of choice

Figure 2.6: Spatial heterogeneity in MWTP for an additional grocery store within one mile



Shade					
WTP ($)	96 - 842	842 - 1393	1393 - 1778	1778 - 2253	2253 - 2686
Shade					
WTP (S)	2686 - 3341	3341 - 4402	4402 - 5657	5657 - 6797	6796 - 8077

Table 2.6**:** First stage results excluding recession year sales

Variable Name	Model 1	Model 2	Model 3	Model 4
House Size x Family Size	$5.7759 \cdot 10^{-4}$*** ($5.8744 \cdot 10^{-5}$)	$2.5896 \cdot 10^{-4}$*** ($6.5493 \cdot 10^{-5}$)	$2.8344 \cdot 10^{-4}$*** ($6.8173 \cdot 10^{-5}$)	$2.0312 \cdot 10^{-4}$*** ($7.3799 \cdot 10^{-5}$)
Number of Grocery Stores x Income[1]	-0.1155*** (0.0018)	-0.0509*** (0.0019)	0.0037 (0.0021)	0.0244*** (0.0021)
Number of Grocery Stores x Latino	0.9890*** (0.0291)	0.2790*** (0.0419)	0.3752*** (0.0350)	0.0681 (0.0427)
Number of Grocery Stores x Black	0.8752*** (0.0143)	0.7544*** (0.0144)	0.5820*** (0.0155)	0.5564*** (0.00154)
Number of Bars x Income			-.1470*** (0.0016)	-0.1220*** (0.0016)
Number of Bars x Latino			1.0628*** (0.0134)	0.5903*** (0.0151)
Number of Bars x Black			0.4973*** (0.0090)	0.4576*** (0.0093)
Distance to nearest Comm x Income		0.7198*** (0.0064)		0.5936*** (0.0066)
Distance to nearest Comm x Latino		-19.2496*** (0.2782)		-16.2445*** (0.3112)
Distance to nearest Comm x Black		-2.5398*** (0.0941)		-0.6350*** (0.0961)

[1] Income in $10000s

Table 2.7**:** Naïve OLS second stage results excluding recession year sales

Variable Name	Coefficient	Standard Error[a]
Price	-0.0014	0.0036
Stories	-0.1016	0.4347
Home Age	0.0516	0.0202**
Home Age^2	-0.0004	0.0001***
Home Size	-0.0008	0.0006
Bedrooms	-.1971	0.2393
Full Bathrooms	-0.1768	0.3612
Half Bathrooms	0.2528	0.6379
Lot Size	-0.0001	$4.31 \cdot 10^{-5}$***
Fireplace	0.0775	0.3011
Air Conditioning	-0.4020	0.2797
Powder Room	-0.1506	0.6719
Basement	0.1579	0.5339
Attic	0.4500	0.3129
Attached Garage	1.0752	0.5368**
Detached Garage	1.1811	0.4216***
Proportion of Census tract which is Black or Latino	-0.1686	0.3285
Distance to the CBD (000's ft)	-0.0433	0.0568*
Distance to the CBD^2	0.0006	0.0009
Distance to the nearest park (000's ft)	-0.3492	0.2384***
Distance to the nearest park^2	0.0976	0.0665***
Distance to the nearest non-grocery commercial lot (000's ft)	1.3351	0.5223***
Distance to the nearest non-grocery commercial lot^2	-1.3814	0.2229***
Groceries within 1 mile	-0.0310	0.0444***
Bars within 1 mile	0.1283	0.0378***

* Significant at 10%, ** 5%, and *** 1% level

[a] Standard errors clustered at the neighborhood level

Table 2.8**:** IV regression second stage results excluding recession year sales

Variable Name	Coefficient	Standard Error[a]
Price	-0.0616	0.0127***
Stories	0.6304	0.6838
Home Age	0.0666	0.0356*
Home Age^2	-0.0006	0.0002***
Home Size	0.0031	0.0013**
Bedrooms	-1.2628	0.3663***
Full Bathrooms	0.3723	0.6849
Half Bathrooms	1.3853	1.2394
Lot Size	-0.0001	$4.34 \cdot 10^{-5}$***
Fireplace	1.7619	0.5024***
Air Conditioning	0.8911	0.4653*
Powder Room	-0.7651	1.2766
Basement	1.8665	0.7273***
Attic	-0.1995	0.4792
Attached Garage	1.8848	0.9393**
Detached Garage	2.4571	0.7749***
Proportion of Census tract which is Black or Latino	-4.4345	0.9632***
Distance to the CBD (000’s ft)	-0.2871	0.0979***
Distance to the CBD^2	0.0044	0.0015***
Distance to the nearest park (000’s ft)	-0.2789	0.3315
Distance to the nearest park^2	0.0611	0.0813
Distance to the nearest non-grocery commercial lot (000’s ft)	2.0370	0.7547***
Distance to the nearest non-grocery commercial lot^2	-1.5268	0.3042***
Groceries within 1 mile	-0.0478	0.0639
Bars within 1 mile	0.1298	0.0540**

* significant at 10%, ** 5%, and *** 1% level

[a] Errors clustered at the neighborhood level

Table 2.9: Heterogeneity in MWTP for housing and neighborhood characteristics excluding recession year sales

	Family Characteristics						
Characteristic	Mean	1	2	3	4	5	6
Annual Income	41,931	20,000	40,000	30,000	40,000	90,000	40,000
Black (Expectation)	0.251	.8	.6	0	.1	0.05	0.05
Latino (Expectation)	0.104	0	0	.6	.5	0.05	0.05
Household Size	2.587	3	2	3	2	5	4
Variable (Change)	Willingness-to-pay						
Square Feet (1 additional sq. ft.)	$58.35	$59.71	$56.42	$59.71	$56.42	$66.31	$63.01
Grocery Stores (1 additional store)	$3,652	$7,241	$6,228	$1,077	$2,266	$3,300	$1,317
Basement	$30,289	$30,289	$30,289	$30,289	$30,289	$30,289	$30,289
Attached Garage	$30,586	$30,586	$30,586	$30,586	$30,586	$30,586	$30,586

Table 2.10**:** First stage results defining food access at .75 miles

Variable Name	Model 1	Model 2	Model 3	Model 4
House Size x Family Size	$7.6347 \cdot 10^{-4}$*** ($5.2604 \cdot 10^{-5}$)	$3.4788 \cdot 10^{-4}$*** ($6.1584 \cdot 10^{-5}$)	$5.5535 \cdot 10^{-4}$*** ($5.7271 \cdot 10^{-5}$)	$3.4124 \cdot 10^{-4}$*** ($6.0267 \cdot 10^{-5}$)
Number of Grocery Stores x Income[1]	-0.1117*** (0.0021)	0.0087*** (0.0022)	0.0051*** (0.0023)	0.0401*** (0.0023)
Number of Grocery Stores x Latino	0.9109*** (0.0358)	0.7116*** (0.0426)	0.6167*** (0.0381)	0.3781*** (0.0432)
Number of Grocery Stores x Black	0.8389*** (0.0175)	0.8412*** (0.0181)	0.6934*** (0.0186)	0.5922*** (0.0187)
Number of Bars x Income			-.1490*** (0.0018)	-0.1105*** (0.0019)
Number of Bars x Latino			1.3197*** (0.0170)	0.7596*** (0.0186)
Number of Bars x Black			0.5708*** (0.0115)	0.4534*** (0.0122)
Distance to nearest Comm x Income		0.7548*** (0.0050)		0.6239*** (0.0051)
Distance to nearest Comm x Latino		-14.5331*** (0.2026)		-12.5175*** (0.2278)
Distance to nearest Comm x Black		-3.2604** (0.0800)		-2.4115*** (0.0844)

[1] Income in $10000s

Table 2.11: Second stage naïve OLS results defining food access at .75 miles

Variable Name	Coefficient	Standard Error[a]
Price (000's dollars)	-0.0008	0.0034
Stories	0.0971	0.3396
Home Age	0.0379	0.0160**
Home Age^2	-0.0003	$9.40 \cdot 10^{-5}$***
Home Size	-0.0009	0.0005*
Bedrooms	-.2894	0.2040
Full Bathrooms	-0.1891	0.2807
Half Bathrooms	0.1742	0.5524
Lot Size	-0.0001	$3.25 \cdot 10^{-5}$***
Fireplace	0.1529	0.2665
Air Conditioning	0.0845	0.2578
Powder Room	-0.4743	0.6140
Basement	-0.1627	0.4616
Attic	0.3714	0.2437
Attached Garage	0.7637	0.4064*
Detached Garage	1.1147	0.3191***
Proportion of Census tract which is Black or Latino	-0.1506	0.2323
Distance to the CBD (000's ft)	-0.0451	0.0183**
Distance to the CBD^2	0.0008	0.0003***
Distance to the nearest park (000's ft)	0.0509	0.1014
Distance to the nearest park^2	-0.0058	0.0317
Distance to the nearest non-grocery commercial lot (000's ft)	-0.8023	0.9260***
Distance to the nearest non-grocery commercial lot^2	-0.5639	0.3321
Groceries within .75 miles	-0.2330	0.0452***
Bars within .75 miles	0.1265	0.0436***

* significant at 10%, ** 5%, and *** 1% level
[a] Standard Errors clustered at the neighborhood level

Table 2.12**:** Second stage IV regression results defining food access at .75 miles

Variable Name	Coefficient	Standard Error[a]
Price	-0.0551	0.0116***
Stories	0.4793	0.5031
Home Age	0.0147	0.0234
Home Age^2	-0.0003	0.0001*
Home Size	0.0021	0.0009**
Bedrooms	-0.9817	0.2784***
Full Bathrooms	0.5280	0.4659
Half Bathrooms	0.8372	0.9412
Lot Size	$-7.14 \cdot 10^{-5}$	$4.71 \cdot 10^{-5}$
Fireplace	1.5656	0.4522***
Air Conditioning	1.2034	0.4542***
Powder Room	-0.6259	1.0211
Basement	1.6012	0.6211***
Attic	-0.0383	0.3181
Attached Garage	0.8675	0.6654**
Detached Garage	1.7643	0.5510***
Proportion of Census tract which is Black or Latino	-3.5265	0.7596***
Distance to the CBD (000's ft)	-0.0965	0.0450**
Distance to the CBD^2	0.0014	0.0007**
Distance to the nearest park (000's ft)	0.0594	0.1273
Distance to the nearest park^2	-0.0144	0.0306
Distance to the nearest non-grocery commercial lot (000's ft)	-0.2890	1.0147
Distance to the nearest non-grocery commercial lot^2	-0.7087	0.3568***
Groceries within .75 miles	-0.1884	0.0722***
Bars within .75 miles	0.0927	0.0588*

* Significant at 10%, ** 5%, and *** 1% level

[a] Standard errors clustered at the neighborhood level

Table 2.13: Heterogeneity in MWTP for housing and neighborhood characteristics defining food access at .75 miles

	Family Characteristics						
Characteristic	Mean	1	2	3	4	5	6
Annual Income	41,931	20,000	40,000	30,000	40,000	90,000	40,000
Black (Expectation)	0.251	.8	.6	0	.1	0.05	0.05
Latino (Expectation)	0.104	0	0	.6	.4	0.05	0.05
Household Size	2.587	3	2	3	2	5	4
Variable (Change)	Willingness-to-pay						
Square Feet (1 additional sq. ft.)	$54.82	$57.38	$51.18	$57.38	$51.18	$69.76	$63.57
Grocery Stores (1 additional store)	$3,044	$6,634	$5,943	$2,883	$4000	$4,018	$375
Basement	$29,056	$29,056	$29,056	$29,056	$29,056	$29,056	$29,056
Detached Garage	$32,014	$32,014	$32,014	$32,014	$32,014	$32,014	$32,014

* Willingness-to-pay represents a one-time payment made by households

Table 2.14: Welfare measures of grocery store policies

	Number Affected		Partial Equilibrium Average Welfare ($)			General Equilibrium Average Welfare ($)		
Policy*	Alternatives	Households	All	Affected	Not	All	Affected	Not
Race	124	3,441	494	1,689	352	1,468	2,159	1,386
Low Income	182	4,654	618	1,647	445	1,500	1,514	1,497
High Income	117	2,304	253	683	220	612	761	601

* Race: increase number of grocery stores by one in neighborhoods with combined black and Latino proportion above .75

Low Income: increase number of grocery stores by one in neighborhoods with average income below $30000

High Income: increase number of grocery stores by one in neighborhoods with average income above $60000

Table 2.15: Price changes due to policies

		Price Change		
Policy	Alternatives Affected	All	Affected	Not Affected
Race	124	-1002	2,444	-1,400
Low Income	182	-871	2,358	-1,449
High Income	117	-29	2,467	-591

Figure 2.7: Milwaukee neighborhoods affected by race-based grocery store policy in 2005

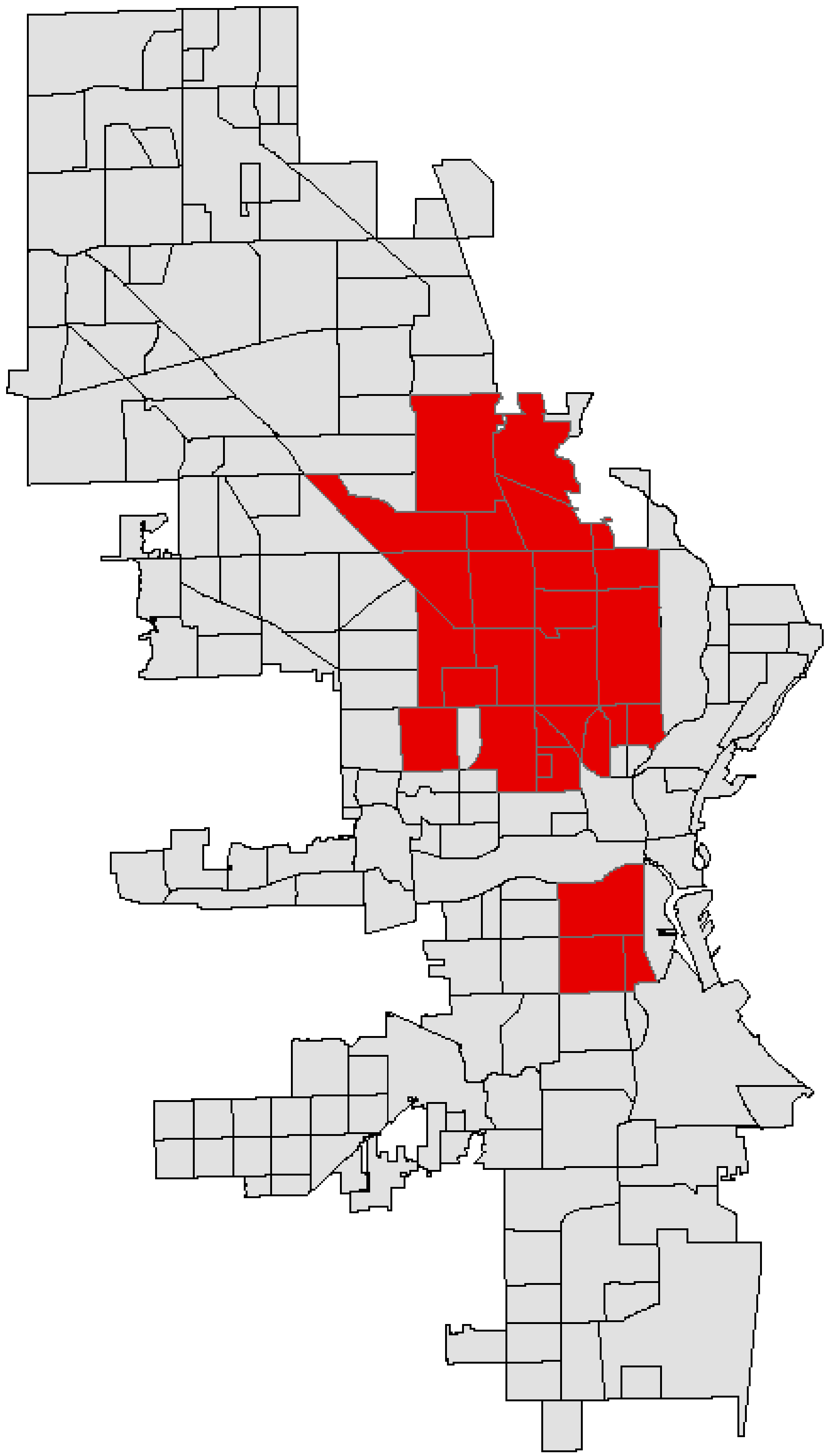
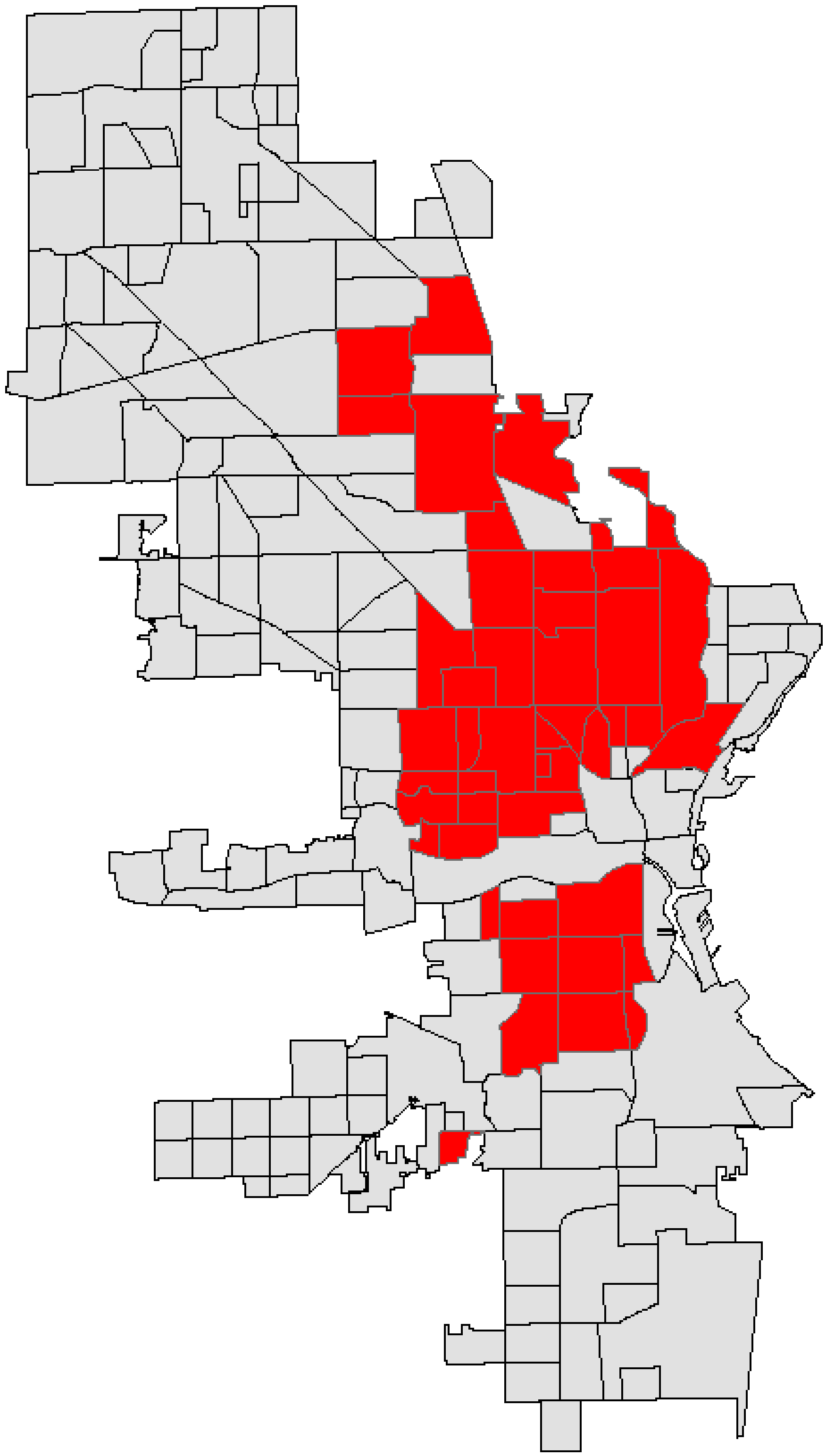


Figure 2.8: Neighborhoods affected by low-income grocery store policy in 2005



2.A Appendix

Appendix 2.A.1**:** Full naïve OLS second stage regression results[a]

Variable Name	Coefficient	Standard Error[b]
Price	-0.0035	0.0040
Stories	-0.0757	0.3681
Home Age	0.0472	0.0152***
Home Age^2	-0.0004	$9.45 \cdot 10^{-5}$***
Home Size	-0.0008	0.0005
Bedrooms	-0.1322	0.2079
Full Bathrooms	-0.2856	0.2817
Half Bathrooms	0.2399	0.5556
Lot Size	-0.0001	$3.32 \cdot 10^{-5}$***
Fireplace	0.2580	0.2661
Air Conditioning	-0.0411	0.2578
Powder Room	-0.4392	0.6190
Basement	-0.1085	0.4748
Attic	0.2465	0.2474
Attached Garage	0.8770	0.3938**
Detached Garage	1.2299	0.3093***
Proportion of Census tract which is Black or Latino	-0.3829	0.2973**
Distance to the CBD (000’s ft)	-0.0500	0.0195**
Distance to the CBD^2	0.0008	0.0003***
Distance to the nearest park (000’s ft)	-0.0191	0.1031
Distance to the nearest park^2	0.0094	0.0183
Distance to the nearest non-grocery commercial lot (000’s ft)	-0.8602	0.8187
Distance to the nearest non-grocery commercial lot^2	-0.5302	0.2965*
Groceries within 1 mile	-0.1513	0.0370***
Bars within 1 mile	0.0488	0.0361
Year 2002 dummy	0.2521	0.1334
Year 2003 dummy	0.4726	0.1337**
Year 2004 dummy	0.5354	0.1363**
Year 2005 dummy	0.7189	0.2450***
Year 2006 dummy	0.7369	0.2416***
Year 2007 dummy	0.6143	0.2325***
Year 2008 dummy	0.3250	0.2626
Year 2009 dummy	0.1350	0.1640
Year 2010 dummy	-0.3336	0.1571**
Year 2011 dummy	-0.3843	0.0968***
Year 2012 dummy	-0.2304	0.0914**
Year 2013 dummy	-0.1264	0.0923
Year 2014 dummy	-0.0492	0.0889
Constant	0.7199	1.1804

* Significant at 10%, ** 5%, and *** 1% level

[a] Year 2015 dummy excluded

[b] Errors clustered at the neighborhood level

Appendix 2.A.2**:** Full second stage IV regression results[a]

Variable Name	Coefficient	Standard Error[b]
Price (000s dollars)	-0.0701	0.0137***
Stories	0.6498	0.5812
Home Age	0.0221	0.0257
Home Age^2	-0.0003	0.0002
Home Size	0.0029	0.0011***
Bedrooms	-0.9724	0.3254***
Full Bathrooms	0.3607	0.5311
Half Bathrooms	1.0401	1.0660
Lot Size	$4.81 \cdot 10^{-5}$	$5.10 \cdot 10^{-5}$***
Fireplace	2.0274	0.5182***
Air Conditioning	1.3126	0.4724***
Powder Room	-0.6254	1.1414
Basement	2.2439	0.6996***
Attic	-0.2080	0.3398
Attached Garage	1.0803	0.7444***
Detached Garage	2.1122	0.6409***
Proportion of Census tract which is Black or Latino	-4.6105	0.8781***
Distance to the CBD (000's ft)	-0.1119	0.0543**
Distance to the CBD^2	0.0016	0.0007*
Distance to the nearest park (000's ft)	-0.0067	0.1383
Distance to the nearest park^2	-0.0026	0.0338
Distance to the nearest non-grocery commercial lot (000's ft)	-0.2295	0.9455
Distance to the nearest non-grocery commercial lot^2	-0.7064	0.3317**
Groceries within 1 mile	-0.1124	0.0600***
Bars within 1 mile	0.0462	0.0514***
Year 2002 dummy	2.3315	0.4499***
Year 2003 dummy	2.9574	0.5454***
Year 2004 dummy	3.5829	0.6530***
Year 2005 dummy	4.3475	0.7735***
Year 2006 dummy	4.4785	0.7935***
Year 2007 dummy	3.8272	0.2325***
Year 2008 dummy	3.0595	0.5742***
Year 2009 dummy	1.8720	0.4053***
Year 2010 dummy	1.1208	0.3814***
Year 2011 dummy	0.2794	0.2061
Year 2012 dummy	0.1608	0.1569
Year 2013 dummy	0.1378	0.1608
Year 2014 dummy	0.1949	0.1427
Constant	1.5114	1.7632

* Significant at 10%, ** 5%, and *** 1% level

[a] Year 2015 dummy excluded

[b] Errors clustered at the neighborhood level

Appendix 2.A.3**:** Full second stage naïve OLS regression results excluding recession year sales[a]

Variable Name	Coefficient	Standard Error[b]
Price	-0.0014	0.0036
Stories	-0.1016	0.4347
Home Age	0.0516	0.0202**
Home Age^2	-0.0004	0.0001***
Home Size	-0.0008	0.0006
Bedrooms	-.1971	0.2393
Full Bathrooms	-0.1768	0.3612
Half Bathrooms	0.2528	0.6379
Lot Size	-0.0001	$4.31 \cdot 10^{-5}$***
Fireplace	0.0775	0.3011
Air Conditioning	-0.4020	0.2797
Powder Room	-0.1506	0.6719
Basement	0.1579	0.5339
Attic	0.4500	0.3129
Attached Garage	1.0752	0.5368**
Detached Garage	1.1811	0.4216***
Proportion of Census tract which is Black or Latino	-0.1686	0.3285
Distance to the CBD (000’s ft)	-0.0433	0.0568*
Distance to the CBD^2	0.0006	0.0009
Distance to the nearest park (000’s ft)	-0.3492	0.2384***
Distance to the nearest park^2	0.0976	0.0665***
Distance to the nearest non-grocery commercial lot (000’s ft)	1.3351	0.5223***
Distance to the nearest non-grocery commercial lot^2	-1.3814	0.2229***
Groceries within 1 mile	-0.0310	0.0444***
Bars within 1 mile	0.1283	0.0378***
Year 2002 dummy	0.0738	0.1318
Year 2003 dummy	0.2892	0.1227**
Year 2004 dummy	0.3520	0.1455**
Year 2005 dummy	0.5239	0.1785***
Year 2006 dummy	0.5557	0.1900***
Year 2011 dummy	-0.4733	0.1011***
Year 2012 dummy	-0.3085	0.0946**
Year 2013 dummy	-0.1676	0.0981
Year 2014 dummy	-0.1131	0.0962
Constant	-0.1502	1.5637

* Significant at 10%, ** 5%, and *** 1% level

[a] Year 2015 dummy excluded

[b] Errors clustered at the neighborhood level

Appendix 2.A.4**:** Full second stage IV regression results excluding recession year sales[a]

Variable Name	Coefficient	Standard Error[b]
Price	-0.0616	0.0127***
Stories	0.6304	0.6838
Home Age	0.0666	0.0356*
Home Age^2	-0.0006	0.0002***
Home Size	0.0031	0.0013**
Bedrooms	-1.2628	0.3663***
Full Bathrooms	0.3723	0.6849
Half Bathrooms	1.3853	1.2394
Lot Size	-0.0001	$4.34 \cdot 10^{-5}$***
Fireplace	1.7619	0.5024***
Air Conditioning	0.8911	0.4653*
Powder Room	-0.7651	1.2766
Basement	1.8665	0.7273***
Attic	-0.1995	0.4792
Attached Garage	1.8848	0.9393**
Detached Garage	2.4571	0.7749***
Proportion of Census tract which is Black or Latino	-4.4345	0.9632***
Distance to the CBD (000's ft)	-0.2871	0.0979***
Distance to the CBD^2	0.0044	0.0015***
Distance to the nearest park (000's ft)	-0.2789	0.3315
Distance to the nearest park^2	0.0611	0.0813
Distance to the nearest non-grocery commercial lot (000's ft)	2.0370	0.7547***
Distance to the nearest non-grocery commercial lot^2	-1.5268	0.3042***
Groceries within 1 mile	-0.0478	0.0639
Bars within 1 mile	0.1298	0.0540**
Year 2002 dummy	0.2044	0.1815
Year 2003 dummy	0.9757	0.2321***
Year 2004 dummy	1.7862	0.3633***
Year 2005 dummy	2.7672	0.5189***
Year 2006 dummy	3.1890	0.5952***
Year 2011 dummy	-0.2209	0.1726
Year 2012 dummy	-0.1589	0.1516
Year 2013 dummy	-0.0808	0.1699
Year 2014 dummy	-0.1121	0.1609
Constant	4.3420	2.2791**

* Significant at 10%, ** 5%, and *** 1% level

[a] Year 2015 dummy excluded

[b] Errors clustered at the neighborhood level

Appendix 2.A.5**:** Full second stage naïve OLS regression results defining food access at .75 miles[a]

Variable Name	Coefficient	Standard Error[b]
Price (000's dollars)	-0.0008	0.0034
Stories	0.0971	0.3396
Home Age	0.0379	0.0160**
Home Age^2	-0.0003	$9.40 \cdot 10^{-5}$***
Home Size	-0.0009	0.0005*
Bedrooms	-0.2894	0.2040
Full Bathrooms	-0.1891	0.2807
Half Bathrooms	0.1742	0.5524
Lot Size	-0.0001	$3.25 \cdot 10^{-5}$***
Fireplace	0.1529	0.2665
Air Conditioning	0.0845	0.2578
Powder Room	-0.4743	0.6140
Basement	-0.1627	0.4616
Attic	0.3714	0.2437
Attached Garage	0.7637	0.4064*
Detached Garage	1.1147	0.3191***
Proportion of Census tract which is Black or Latino	-0.1506	0.2323
Distance to the CBD (000's ft)	-0.0451	0.0183**
Distance to the CBD^2	0.0008	0.0003***
Distance to the nearest park (000's ft)	0.0509	0.1014
Distance to the nearest park^2	-0.0058	0.0317
Distance to the nearest non-grocery commercial lot (000's ft)	-0.8023	0.9260***
Distance to the nearest non-grocery commercial lot^2	-0.5639	0.3321
Groceries within 1 mile	-0.2330	0.0452***
Bars within 1 mile	0.1265	0.0436***
Year 2002 dummy	0.2151	0.1767
Year 2003 dummy	0.3814	0.1766**
Year 2004 dummy	0.4438	0.2004**
Year 2005 dummy	0.6007	0.2260***
Year 2006 dummy	0.6105	0.2237***
Year 2007 dummy	0.5073	0.2127**
Year 2008 dummy	0.1848	0.2775
Year 2009 dummy	0.0167	0.1527
Year 2010 dummy	-0.3179	0.1554**
Year 2011 dummy	-0.4158	0.0914***
Year 2012 dummy	-0.2512	0.0893***
Year 2013 dummy	-0.1384	0.0901
Year 2014 dummy	-0.0451	0.0871
Constant	1.4207	1.1729

* Significant at 10%, ** 5%, and *** 1% level
[a] Year 2015 dummy excluded
[b] Errors are clustered at the neighborhood level

Appendix 2.A.6**:** Full second stage IV regression results defining food access at .75 miles[a]

Variable Name	Coefficient	Standard Error[b]
Price	-0.0551	0.0116***
Stories	0.4793	0.5031
Home Age	0.0147	0.0234
Home Age^2	-0.0003	0.0001*
Home Size	0.0021	0.0009**
Bedrooms	-0.9817	0.2784***
Full Bathrooms	0.5280	0.4659
Half Bathrooms	0.8372	0.9412
Lot Size	$-7.14 \cdot 10^{-5}$	$4.71 \cdot 10^{-5}$
Fireplace	1.5656	0.4522***
Air Conditioning	1.2034	0.4542***
Powder Room	-0.6259	1.0211
Basement	1.6012	0.6211***
Attic	-0.0383	0.3181
Attached Garage	0.8675	0.6654**
Detached Garage	1.7643	0.5510***
Proportion of Census tract which is Black or Latino	-3.5265	0.7596***
Distance to the CBD (000’s ft)	-0.0965	0.0450**
Distance to the CBD^2	0.0014	0.0007**
Distance to the nearest park (000’s ft)	0.0594	0.1273
Distance to the nearest park^2	-0.0144	0.0306
Distance to the nearest non-grocery commercial lot (000’s ft)	-0.2890	1.0147
Distance to the nearest non-grocery commercial lot^2	-0.7087	0.3568***
Groceries within .75 miles	-0.1884	0.0722***
Bars within .75 miles	0.0927	0.0588*
Year 2002 dummy	1.9231	0.4012***
Year 2003 dummy	2.4143	0.4725***
Year 2004 dummy	2.9307	0.5614***
Year 2005 dummy	3.5568	0.6644***
Year 2006 dummy	3.6553	0.6773***
Year 2007 dummy	3.1204	0.6034***
Year 2008 dummy	2.4064	0.4705***
Year 2009 dummy	1.4355	0.3507***
Year 2010 dummy	0.8758	0.3351***
Year 2011 dummy	0.1269	0.1826
Year 2012 dummy	0.0684	0.1430
Year 2013 dummy	0.0732	0.1417
Year 2014 dummy	0.1515	0.1289
Constant	2.3204	1.5734

* Significant at 10%, ** 5%, and *** 1% level
[a] Year 2015 dummy excluded
[b] Errors clustered at the neighborhood level