Location Choice and Food Tradeoffs: Does Local Matter?

Cristina Connolly, Assistant Professor, University of Connecticut, 1376 Storrs Road, Unit 4021, Room 302, Storrs, CT 06269-4021, cristina.connolly@uconn.edu

H. Allen Klaiber, Professor, The Ohio State University, 333 Ag Admin Building, 2120 Fyffe Road Columbus, OH 43210, klaiber.16@osu.edu

JEL Codes: R21 Housing Demand; R31 Housing Supply and Markets

Abstract

We use housing transactions to estimate consumer preferences for living near direct-marketed locally produced food, which has implications for both consumer- and producer-centric land use policies. We estimate a pure characteristics sorting model to recover preferences and conduct counterfactual simulations accounting for price feedback effects following resorting after changes to the local food environment. We find that consumers significantly value healthy food access, including different direct sources of locally produced foods. However, the welfare gains and associated price changes arising from policies to increase local food access depend on the targeting method, a key concern for policymakers worried about affordability.
1. Introduction

Over 11 million acres of farmland were repurposed between 2001-2016, primarily for housing development (AFT, 2016), and urbanization has led to increased farmland rent prices along the rural-urban fringe (Delbecq et al., 2014). However, this heightened pressure also provides an opportunity for diversified farms to market local foods directly to this new consumer base. Urbanization has been shown to positively impact farm income, potentially due to a shift to specialty crop production (Wu et al., 2011). In fact, crop net returns are a larger driver of cropland changes than urbanization alone (Lubowski et al., 2008), suggesting those farms that produce in urbanizing areas are able to capitalize on an increase in consumer demand.

The prevalence of outlets for direct sales of local foods across the urban environment has also grown quickly over the past few decades: the number of farmers’ markets in the U.S. has tripled between 2000 and 2015 (USDA, 2017), the incidence of community supported agriculture (CSAs) increased by 50% from 2001 to 2005 (Martinez, 2010) and local food sales have grown from $1 billion in 2005 to $11.8 billion in 2017 (USDA, 2013; Martinez and Park, 2021). While this growth is presumably reflective of consumers preferences for locally grown food (Tropp, 2014, Feldmann and Hamm, 2015; Berg and Preston, 2017; Fan et al., 2019), as policymakers consider how best to support farmers along the expanding urban-rural fringe, it is important to understand consumer willingness to pay to live near local food. While research has shown that consumers value open space (Blanchette, Lang and VanCeylon, 2021) and are willing to pay a premium for locally produced food (Printezis et al., 2019), information on consumers’ willingness to pay to live near local food outlets is limited. Additionally, studies looking at household valuation of traditional retailers primarily use reduced-form hedonic models (Warsaw and Phaeneuf, 2019;
Caudill et al., 2021), which are ill suited for analysis of housing market changes following expansions of new local food offerings in the urban landscape. In this study we employ a characteristic, or vertical, sorting model to assess whether and to what degree proximity to local food drives household sorting behavior and examine welfare and price adjustments following changes in local foods provision through a series of policy counterfactuals.

Local food access has been touted as an important element of a healthy built community. Local food direct marketing outlets such as farmers’ markets can improve overall community well-being (Brown and Miller, 2008; Warsaw et al., 2021), and research suggests that local food sales can have positive economic benefits for local communities (Hughes et al., 2008; Brown et al., 2014; Hughes & Isengildina-Massa, 2015; Rahe et al., 2017; Rossi et al., 2017). Additionally, policies striving to improve healthy food access have emphasized utilizing local food resources. The Centers for Disease Control and Prevention (CDC) recommendations to reduce obesity include incentivizing healthy food establishments, facilitating consumer procurement of food directly from farmers and encouraging local food production (Khan et al., 2009). However, the availability of healthy food inputs is directly tied to an individual’s location as food access varies widely across space (Larson et al., 2009; Hilmers et al., 2012; Laxy et al., 2015; James et al., 2017). Thus, our research has policy implications not just for agricultural producers, but also consumers.

In the United States more than one-third of children are obese and in 2008 obesity-related medical costs totaled $147 billion. The federal government has responded to this in a variety of ways, including improving the nutritional quality of school meals, providing education on dietary guidelines, and supporting initiatives to increase consumer produce access (USDA, 2017; Lin et al., 2019; Forrestal et al., 2021; USDA, 2021). However, across the country there is a clear
relationship between neighborhood income and food access. Low-income neighborhoods tend to have less healthy food environments (Laxy et al., 2015), including more convenience stores (Hilmers et al., 2012; Lee, 2012; Bower et al., 2014; Richardson et al., 2014), fast food restaurants (Hilmers et al., 2012; Lee, 2012; James et al., 2017) and small groceries (Bower et al., 2014). These small, independent grocery stores tend to have fewer healthy offerings and variety (Jetter et al., 2005), and convenience store density is associated with lower food quality, especially in low-income communities (Rummo et al., 2015). Low-income neighborhoods also display lower food quality and less variety overall (Hendrickson et al., 2006). Thus, policies that aim to connect producers with local food consumers must take into account not only how household location choices and preferences influence patterns of food access, but also the potential impact of policy interventions on housing affordability.

Though there has been only limited research pertaining to the effect of local food on health outcomes, initial results suggest an important role for local foods marketed directly to consumers as a component of a healthy community (Jilcott et al., 2011; Ruelas et al., 2012; Salois, 2012; Minaker et al., 2014; Vasquez et al., 2017). Providing funds to be spent at farmers’ markets to recipients of the Women, Infants and Children (WIC) and Supplemental Nutrition Assistance Programs (SNAP) is associated with a significant increase in consumption of fruits and vegetables (Herman et al., 2008; Savoie-Roskos, 2016; Durward et al., 2019). Quandt et al. (2013) showed similar results for Community Supported Agriculture as low-income families that were offered a membership consumed a significantly higher amount of fruits and vegetables. These findings held for both community gardens, as neighborhood residents who participated in urban agriculture consumed more fresh produce and had improved health and diets (Barnidge et al., 2013; Algert et
al., 2016; Audate et al., 2019), and mobile farm stands (Evans et al., 2012). Additionally, as direct to consumer marketing is associated with improved farm resiliency (Low et al., 2015), these types of initiatives can support producers otherwise unable to maintain profitability in the face of increasing land prices.

Despite the dual benefits to producers and consumers of local food access, little formal research has been undertaken to understand the linkages between housing markets, residents’ location choices and nearby local food amenities. Taken together, this evidence suggests a basis for public policies that use the provision of local foods to improve neighborhood quality. One challenge facing policymakers concerned with improving the local food environment is how to evaluate the potential benefits and costs associated with efforts to improve access to healthy, and increasingly local, food. While increasing the availability of local food outlets could benefit both producers and residents, any policy efforts need to consider potential feedback effects such as increasing prices if community improvements increase demand for those locations through Tiebout (1956) sorting. This concern is likely to be particularly acute in communities with low existing incomes, often the same communities lacking healthy food access. Studies have shown that income segregation can result in disparate neighborhood amenity provision and health outcomes. For instance, low-income residents tend to have reduced access to healthy food retail stores such as supermarkets and live in neighborhoods with a greater number of unhealthy food establishments.

There have been discussions in the policy domain of how best to select neighborhoods for policy targeting to produce equitable outcomes (Cohen, 2018; Friedman, 2022), suggesting the importance of including this factor in policy simulations.

This paper uses revealed preference data on household location decisions to model
preferences for local food amenities, as well as the tradeoffs made with respect to traditional public goods including school quality and park access, using a structural model of location choice. A subsequent general equilibrium welfare analysis then measures the impact of potential food environment policies on current residents and equilibrium price outcomes. The pure characteristics sorting model is an empirical characterization of Tiebout (1956) sorting where households are deciding where to live and take into account the bundle of amenities provided by their neighborhood (Klaiber and Kuminoff, 2014). As households sort across communities, the relative supply and demand for location forms the foundation of income-based sorting which leads to a set of equilibrium prices needed to clear the market (Epple and Sieg, 1999).

A variety of hedonic valuation studies using the resulting equilibrium prices from this sorting process have shown a positive relationship between house price and traditional amenities such as school quality (Black, 1999; Downes and Zabel, 2002; Figlio and Lucas, 2004; Bayer et al., 2007) and open space (Walsh, 2007; Klaiber and Phaneuf, 2010). Research using hedonic models and difference-in-difference estimation suggest households value living near grocery stores (Cerrato Caceres and Geoghegan, 2017; Warsaw and Phaneuf, 2019; Caudill et al., 2021) and community gardens (Voicu and Been, 2008). However, to our knowledge no prior studies have focused on the degree to which healthy food options influence sorting behavior directly using structural modeling or assessed preferences for heterogeneous types of local food provision. In addition to traditional drivers of location such as education quality it is also likely that households will consider resources such as access to different types of food establishments when choosing a neighborhood; local food accessibility, such as nearby farmers’ markets, have been cited as a potential factor consumers take into account when choosing what neighborhoods to move into (Bibbo, 2009; Markham, 2014). By
estimating a structural sorting model, we not only recover preference and marginal willingness to pay estimates for aspects of the local food environment but can also use those estimates to conduct counterfactual simulations and examine changes in equilibrium price outcomes following policy intervention.

Food access can encompass a variety of dimensions including traditional sources such as supermarkets, convenience stores and restaurants as well as more recent efforts to provide locally produced food to urban communities through farmers’ markets and other direct to consumer sales of local produce. Specifically, the food access literature differentiates between the concept of a food desert, where there is limited healthy food availability, and a food swamp, which is characterized by a multitude of unhealthy and healthy food establishments (Hager et al., 2017). For instance, while low-income neighborhoods contain higher rates of unhealthy food establishments, they don’t necessarily have fewer supermarkets than wealthy neighborhoods (Lee, 2012; Richardson et al., 2014). Cooksey-Stowers et al. (2017) found that food swamps were better predictors of obesity than food deserts, which is in line with a USDA report that suggests any food access could have a causal obesity link as consumers may not substitute healthy food for unhealthy items but instead increase overall consumption (Ver Ploeg et al., 2009). However, none of these studies accounted for the impact of local food sources. Additionally, while low-income neighborhoods face higher food prices overall (Hendrickson et al., 2006; Sheldon et al., 2010; Rogus, 2015), there could be a preference for the convenience offered by less healthy food establishments; prices paid at formats such as convenience stores tend to be higher than those charged at supermarkets for the same items (Ver Ploeg et al., 2009; Borja and Dieringer, 2019). Thus, it is possible that preferences for store types could differ across income levels. As such, in
our analysis we employ multiple measures of food availability, including the presence of agricultural direct marketing operations and an index for the overall healthiness of a neighborhood’s food environment, to capture how the food environment impacts household location choice.

Given the increasing public awareness and policy concern surrounding food provision in communities (see, e.g., recent efforts to limit fast food in some municipalities), this paper provides policy relevant and timely information for local policymakers. Our paper presents three primary contributions. First, we assess the value consumers place on a healthy food environment. Second, beyond the healthiness of the food environment we demonstrate a preference for living near local food availability. Third, we model how changes in the local food environment impact consumer welfare and alter prices across the urban landscape, a key concern for policymakers seeking to balance affordability with local goods provision. Previewing our results, we find that households value living in neighborhoods with healthy and local food options. We also demonstrate that a simulated policy to increase food access leads to overall positive welfare gains, but the degree of welfare improvement depends on how neighborhoods are targeted. Finally, we show how a general equilibrium model that accounts for household movement after a policy provides important insights potentially missed in partial equilibrium models as changes in food access are often nonmarginal in nature.

2. Model of household location choice
Rosen’s (1974) first-stage hedonic has been widely used to measure how public goods are capitalized into housing prices. The equilibrium housing prices in the hedonic are assumed to arise
from an equilibrium outcome resulting from sorting by households over the available housing supply. This partial-equilibrium model allows researchers to calculate consumer marginal willingness to pay (MWTP) for selected amenities but ignores general equilibrium effects. Additionally, due to reliance on this equilibrium relationship, the first-stage hedonic is well-known to suffer from limitations in the types and magnitudes of policy counterfactuals suitable for evaluation (Kuminoff et al, 2013). Taking a step back, a class of equilibrium sorting models has been developed to overcome these challenges by recovering the underlying preference structure of household utility giving rise to the sorting behavior underlying Rosen’s (1974) hedonic approach.

The pure characteristics sorting model is one class of structural Tiebout models (Epple and Sieg, 1999). Numerous authors have applied a structural approach to study local public goods in a housing market setting where the role of feedback effects and re-sorting are of potential interest. These studies include topics such as school quality (Bayer et al., 2007), air quality (Sieg et al, 2004), open space (Walsh, 2007; Klaiber and Phaneuf, 2010) and urban noise (Klaiber and Morawetz, 2021). However, to our knowledge there exists no hedonic or sorting model approach examining the role of the local food environment, a setting where land use, affordability and price feedbacks play a particularly important role in policy evaluation.

The pure characteristics model (PCM) of household location choice begins by specifying a constant elasticity of substitution (CES) utility function. Budget constrained households face a mixed discrete-continuous choice set where they choose both a discrete neighborhood and a continuous quantity of housing services. Following the Tiebout logic, households choose a community which confers a bundle of local public goods, $G_j$, which include aspects of the food environment, school quality, open space and an unobserved public goods component. Conditional
on that location decision, a consumer also selects the optimal level of housing services given housing prices in the chosen neighborhood.

Households differ in their preferences ($\alpha$) and income ($y$). Heterogeneity in households is characterized by the joint distribution, $F(\alpha,y)$, which forms the basis for income based sorting (Ellickson, 1971). Household preferences are defined over neighborhood quality/amenities $G_j$, a quantity of housing services $q$ and a composite private good $b$. The CES specification for preferences defines the utility that household $i$ obtains from living in community $j$ as:

$$V_{i,j} = \left\{\alpha_i(G_j) + \left[\exp\left(\left(\frac{y_i}{1-\nu}\right)^{\frac{1}{\rho}} - \frac{\beta_j b^{1+1}}{1+\eta}\right)\right]^\frac{1}{\rho}\right\}$$

[1]

where $F(\alpha,y) \sim \lognormal$. The first term in this CES specification represents the utility households receive from neighborhood amenities while the second term encompasses utility from the private good component of housing. This specification is chosen as the CES parameters are readily interpretable and can be easily compared to estimates from the existing literature.

The index of public goods, $G$, is defined as a linear index of amenities provided by each community

$$G_j = y_1g_{1,j} + \cdots + y_{R-1}g_{R-1,j} + \xi_j$$

[2]

Households agree on a common set of weights for the amenities in the index ($y_1, \ldots, y_{R-1}$) but differ in their overall preferences for amenities relative to the private good components of housing and the numeraire ($\alpha_i$). Of the $R$ amenities in the index, $R-1$ are observable. Then $G_{R,j} = \xi_j$ represents the composite of public goods unobserved by the analyst but observed by households. Note that the “error term” of the model enters into the indirect utility function in a non-additively separable manner. This gives rise to the “pure characteristics” nomenclature as utility is defined
solely over the characteristics of communities and there is no idiosyncratic location-household-specific shock. As our estimates of the public good index $G$ depend on housing prices, any correlation between our unobserved amenity component $\xi_j$ and housing prices introduces endogeneity into our model. Following Epple and Sieg (1999) and Sieg et al. (2004) we use the price ranks of our communities as instruments.

Validity of the price rank instruments requires that the observed component of the public good play a larger role in the location sorting decision than the unobserved component (Epple and Sieg, 1999), which is generally achieved in U.S. contexts by including some measure of school quality (Klaiber and Morawetz, 2021). While there is no official test for the validity of the endogeneity instrument in a vertical sorting model, a necessary and sufficient condition for validity is that the observed component of the neighborhood public good must be larger than the unobserved component in the public goods index. One potential approach to examine this assumption is to use a hedonic regression to assess the degree of explanatory power observed public goods have on community price rank.

For the private good component, community housing prices are a separate subfunction that reflect an equilibrium driven by neighborhood public goods. Households are assumed to share the same elasticity of substitution between amenities and private goods ($\rho$), and the same demand parameters for the private good components of housing: price elasticity of housing ($\eta$), income elasticity of housing ($\nu$), and demand intercept ($\beta$). $\nu$ is expected to be positive as an increase in income will lead to an increase in demand for housing, while $\eta$ should be negative as a higher price should result in reduced demand. $\beta$ is expected to be positive because as price increases demand will decrease, and the formulation above incorporates a negative sign.
To characterize a sorting equilibrium, it must be the case that prices, physical housing characteristics, amenities and location choices are all defined such that no household can improve its utility by moving and each household occupies exactly one house. Using this indirect utility function, Epple and Sieg (1999) derive the necessary conditions for equilibrium that include boundary indifference, increasing bundles and stratification. The increasing bundles property implies that locations with higher prices have better amenities and conditional on taste there will be a positive sorting by income. Boundary indifference defines the income and preference combination \((\alpha,y)\) that makes a household exactly indifferent between neighborhoods \(j\) and \(j+1\). Stratification implies that households in locations with higher rankings of the public good have higher income and stronger preferences for amenities.

Given that a household with income and preference combination \((\alpha,y)\) makes their location decision based on amenity provision, \(G\), and house price, \(p\), Ellickson’s (1971) single crossing condition ensures the sorting restrictions described above hold. Specifically, the model is “vertical” as households agree on the ranking of locations by overall quality and differ only in their preferences for said “quality” relative to the numeraire. Given this assumption, if the slope of an indirect indifference curve in \((G,p)\) space is monotonically increasing in income \((y|\alpha)\) and preferences \((\alpha|y)\) then indifference curves in the \((G,p)\) plane will satisfy single crossing in \(y\) and \(\alpha\). This ensures that households will sort into neighborhoods by income and taste preferences and implies a negative value for rho \((\rho)\), the elasticity of substitution between amenities and private goods. However, there is no requirement that there be a positive or large correlation between income and preferences \((\lambda)\). In fact, while we assume that conditional on income those with a
larger taste for public goods will sort into higher-ranked communities, this model places no restriction on how income impacts taste.

Along with certain constraints on the utility function, the above conditions ensure that a sorting equilibrium can be described by a hedonic price function as equilibrium prices are functionally related to housing characteristics and amenities \( P_{n_j} = P(G_j, h_{n_j}) \). Unlike the traditional hedonic model, there is no requirement that households be free to choose continuous quantities of each amenity nor is the market assumed to be perfectly competitive. Instead, so long as single-crossing holds, Sieg et al. (2002) show that housing expenditures can be expressed as the product of a price index and a quantity index

\[
P_{n_j} = q(h_{n_j}) \cdot p(G_j).
\]  

By taking the log of this function it is possible to generate a hedonic model

\[
\ln P_{n_j} = \ln q(h_{n_j}) + \ln p(G_j),
\]  

allowing the neighborhood-level prices \( P_1, \ldots, P_j \) to be estimated as fixed effects in a hedonic regression using market transactions data.

Estimation proceeds using the simulated two-stage generalized method of moments estimator developed by Sieg et al. (2004). In the first stage, housing price estimates are treated as known constants in order to recover all of the structural parameters

\[
\theta = [\beta, \eta, \nu, \rho, \mu_x, \mu_y, \sigma_x, \sigma_y, \lambda, G_1, y_1, \ldots, y_{R-1}].
\]  

Following Sieg et al (2004) the parameters can be recovered using methods of moments by matching observed and estimated values of public goods, income quartiles and expenditure quartiles. Parameter estimates are continually adjusted to satisfy the following moment conditions:
The first moment condition is based on the level of amenity provision, where the public good is defined using a linear relationship between school quality, open space and distinct measures of food access. Given a value for the lowest public goods community, $G_1$, the sorting behavior implied by vertical differentiation allows $G_2, \ldots, G_j$ to be defined recursively. The predictions for $G_1, \ldots, G_j$ are then used to identify the weights in the amenity index. The residual to the moment condition defines the composite unobserved amenity in each community $(\xi_1, \ldots, \xi_j)$ as the researcher does not perfectly observe $G_j$ but instead $G_j + \xi_j$.

The next three moment conditions are based on the model’s prediction for the distribution of income. Under the maintained assumptions on preferences, the information in $\theta$ can be used to simulate community-specific income distributions. Three of the moment conditions match the 25th, 50th, and 75th quantiles from the simulated distributions of income in each community $(\tilde{y}_j^{25}, \tilde{y}_j^{50}, \tilde{y}_j^{75})$ to their empirical counterparts $(y_j^{25}, y_j^{50}, y_j^{75})$. Income data from the 2010 American Community Survey was used to create income quartiles for each neighborhood. Given that the data included the number of households in a series of income brackets, coefficients from a censored interval regression were used to estimate the 25th, 50th, and 75th quantiles. The last three moment conditions use the simulated income distributions to match predicted and observed quantiles from the distribution of housing expenditures in each community. The expenditure...
moments are obtained by multiplying the demand function by price and taking logs. For details of the mechanisms of the simulated GMM estimator, see Klaiber and Kuminoff (2014).

3. Data
We assembled data covering a five-county region comprising the Cleveland, OH metropolitan area shown in figure 1.

[Insert Figure 1 Here]
This area contains a population of 2.08 million and is the 34th largest metropolitan region in the United States (U.S. Census Bureau, 2022), as well as one of the most impoverished; Cleveland has both the highest child poverty rate (46.1%) and overall poverty rate (30.8%) of all large cities (Campbell, 2020). Contained within the metropolitan area are 599 census tracts, which are defined to represent one of the j=1...599 discrete neighborhoods over which households choose to locate. The results presented here consider years 2011-2014, a period which captures a time of rapid growth in local foods availability (Low et al., 2015) while avoiding the identification issues inherent in transactions data from the housing boom and subsequent bust. The analysis combines data on residential household transactions, firm locations, and local public goods including school quality and measures of the food environment. Each of these sources of data is described in the subsections that follow, but we provide overall summary statistics for our 599 neighborhoods in table 1.

[Insert Table 1 Here]

Our data reveal significant heterogeneity between neighborhoods, an important condition to ensure that there is likely to exist significant location-based sorting, which suggests our model
is appropriate for the selected study area. For example, the price index, which represents consumer valuations of each neighborhood, ranges from a normalized value of 1 to a high of 11.95. This is similarly reflected in the variation present in neighborhood amenity provision. While an average neighborhood is situated in a school district where 76% of students meet the reading standard, in certain areas only 58% of students are deemed proficient readers, compared to 97% in the top district. Similarly, the amount of opens space in a neighborhood ranges from 0 acres to 7,277. Turning to our food measures, in the average neighborhood only 12% of food establishments are classified as healthy, though this ranges from 0 to 100%. Meanwhile, some neighborhoods have as many as 7 local food establishments, while others have none. The average tract has 0.4 farmers’ markets, 0.14 non-farmers’ market local food operations, and is 0.85 miles from the closest local food establishment. For each neighborhood we also identify the median income and house price quartiles. The heterogeneity between quartiles reflects similar differences in neighborhood desirability; for instance, the interquartile range for income in a given tract is $52,090. Thus, in the Cleveland metro area not only do neighborhoods differ in amenity provision, but this appears to be reflected in variation in house prices and resident income.

**Price Index Estimation**

A key requirement of the sorting model is that each census tract be characterized by a price index, which can be estimated as the fixed effect values of a hedonic housing regression. Housing transactions data for the years 2011-2014 was obtained from each county auditor and supplemented with parcel-level GIS data collected for each county in the study area. These transactions contain information on the sale price as well as physical characteristics of homes such as number of bedrooms, number of bathrooms and lot size. Summary statistics for this set of
housing transactions data are shown in Appendix Table A1 and reveal an average sale price of $129,999 with mean square footage of 1,765 and approximately 1.5 baths.

An initial step required for estimation of the structural location choice model described in section 2 is recovery of neighborhood specific, or location specific, price indices which represent the price of a homogenous unit of housing services in each community. For this analysis, we define communities as consisting of one of 599 Census tracts, with each tract containing approximately 1,316 households. We follow the literature (e.g. Epple and Sieg, 1999; Sieg et al, 2004; Klaiber and Smith, 2012) and estimate a hedonic regression that controls for housing attributes and includes a community specific fixed effect for each community. Conditional on controlling for housing attributes, the estimated fixed effects capture the value of purchasing a unit of housing services in each community providing the locational price index needed for estimation of the pure characteristics model. Defining $\ln q(h_n) = \sum \beta_k x_{nk}$, the hedonic equation can be written as:

$$\ln P_n = \sum_{k=1}^{K} \beta_k x_{nk} + \sum_{j=1}^{J} \delta_j + \epsilon_n$$  \hspace{1cm} [7]

where $P_n$ are annualized housing prices, $k = 1 \ldots K$ represents housing attributes, $\delta_j = 1 \ldots J$ are the tract fixed effects that serve as price indices ($\delta_n = \ln p(G_j)$), and $\epsilon_n$ is an idiosyncratic unobservable.

We estimate this regression using sales prices for 70,067 transactions. The resulting estimates for the key housing attributes are shown in Appendix Table A2. Taking the exponential of the estimated tract-level fixed effects provides the price index used in the structural model. A spatial visualization of these prices is shown in figure 2 (panel A), demonstrating low-priced neighborhoods in inner-city Cleveland, with higher prices further from the city center and along
the waterfront. These location values mirror what is expected in this region, and follow the patterns shown in comparable cities (Hodge et al., 2017).

[Insert Figure 2 Here]

**Sorting Model Data**

To characterize a neighborhood’s food environment, we develop two separate food access measures that encompass both local and traditional retailers. To assess the healthiness of conventional food retailers we use the modified retail food environment index (mRFEI). Developed by the CDC’s Division of Nutrition, Physical Activity and Obesity, this index defines a neighborhood’s health index as the percentage of food establishments that are considered healthy. Data was collected from the Dunn and Bradstreet Million Dollar Database, a national database of both publicly and privately held companies. Establishments were identified using North American Industry Classification System (NAICS) codes. In this index, an unhealthy food retailer includes fast food restaurants (NAICS 722211), small grocery stores (NAICS 445110 with 3 or fewer employees) and convenience stores (NAICS 445120), while healthy establishments include supermarkets, large grocery stores, warehouse clubs (452910) and produce markets (445230). The mRFEI is calculated as \[
\frac{\text{# of healthy food establishments}}{\text{Total # of food establishments}}
\]

As consumers may value local and conventional healthy establishments differently, we include information on the local food environment as additional components of a community’s public goods index. Data on local food establishments was purchased from Local Harvest, an
independent database that maintains a list of farms and direct marketing locations in the United States, on September 3, 2014. A direct-marketing farm is included if it operated a CSA, U-Pick enterprise, or farm stand. To ensure we are accurately counting current farms we only incorporated establishments whose entry had been updated or created in 2011 or later. Local Harvest additionally provided a list of farmer’s market locations, which was supplemented with data from the USDA National Farmer’s Market directory downloaded on August 8, 2014.

To construct variables from the direct to consumer local foods data, we first measured the total number of local food operations within one half mile of each neighborhood’s boundaries. We then grouped local foods availability into measures of farmers’ markets and farm-based operations (CSAs, U-Picks and farm stands). For each census tract we also calculated the distance of each home in our sample to its closest local food establishment to develop a house-weighted average local food distance variable. The spatial distribution of our key food related public goods are visualized in panels B and C in figure 2, and reflect the neighborhood pricing patterns seen in panel A.

In addition to our food environment variables, we also include school quality and open space acreage as additional public goods in the public goods index. As these variables are well known, and frequently used, as core drivers of residential location choice in both hedonic and sorting models, it helps both to ensure our observable public goods meet instrument validity requirements and provides a basis for us to compare our estimates to the literature. To examine the validity requirement that the observed public goods are more important than unobserved goods for community ranking, we ran an auxiliary hedonic regression of community price index rank with school quality as the sole independent variable. We found an $R^2$ of 0.59, suggesting our
observed public goods meet this requirement.

Each census tract was assigned to a school district based on the spatial location of the centroid of the census tract and testing data was collected from the Ohio Department of Education website. As raw test scores are not available in Ohio, the school quality performance statistic is calculated as the percentage of 3rd grade students who passed the reading test for the 2006-2007 school year at a state-designated proficiency level, which is a prerequisite for advancement to the 4th grade. Data on local parks was collected from Ohio county GIS offices and we calculated the total acreage of protected open space in each census tract. The final neighborhood variable is household income, which was obtained from the 2010 American Community Survey five-year estimates.

4. Results and Counterfactual Simulations

The initial GMM estimation results are shown in table 2. All signs are as expected: the parameters for \( \nu \), the income elasticity, are positive; \( \eta \), the price elasticity, is negative for all iterations; and \( \beta \), the demand intercept, is consistently positive. Finally, the estimates for \( \rho \), the elasticity of substitution between amenities and private goods, are negative, which indicates the single crossing property holds. However, our neighborhoods do not demonstrate perfect income stratification; our lambda value is essentially zero, implying very little correlation between income and preferences\(^\text{viii}\). All models include both school quality and protected open space area as components of the public goods index to ensure instrument validity is likely met in our context\(^\text{ix}\).

[Insert Table 2 Here]
We present results for a progression of local public goods specifications in table 2. The first model examines the tradeoffs between healthy retail environments, local food, open space and school quality. Focusing on the parameters associated with the public goods index, and recalling that the index for school quality is normalized to 1 as is standard in the literature, we find that the effect of protected open space, captured by $\gamma_{\text{OpenSpace}}$, is positive, which demonstrates a consumer preference for neighborhoods with more open space acreage. Turning to our food access variables of interest, the positive values of $\gamma_{\text{mRFEI}}$ and $\gamma_{\text{Local Food}}$ suggest that the presence of both healthy retail establishments and local foods are an important and positive determinant of households’ decisions about community location.

The second model in table 2 decomposes the local food variable into off-farm farmers’ markets and farm-based operations. Both types of local food establishments are positively valued, though farmers’ markets appear to contribute less to the overall public good index. In the third, model we add distance to the closest local food establishment. The negative sign on the distance variable suggests that consumers prefer more proximate access to local food. Using this third model as our preferred specification, the point estimate for the income elasticity, $v$ of 0.83 is consistent with that seen in previous literature, as is the value for the price elasticity, $\eta$ of -0.3, adding support for the functioning of sorting in our study context (Polinsky, 1977; Sieg et al., 2004).

Using the estimates reported for our preferred model in table 2, we present marginal willingness to pay (MWTP) estimates for each public good in table 3. These MWTP measures represent the value consumers place on a marginal public good change, holding all other amenities and prices constant. We find that consumers are willing to pay $6.84 a month for a one percentage
point increase in the reading pass rate, and $4.83 for an additional 10 acres of open space. A 1% improvement in the healthiness of the retail food environment is associated with a willingness to pay of $17.46 a month. Consumers are willing to pay an additional $27.27 a month to live 0.1 miles closer to a direct-to-consumer local food operation. We find that an additional farmers’ market is associated with a willingness to pay of an additional $82.77 a month; in comparison, an additional farm-based local food operation is valued at $141.64 a month. Given that there are far fewer farm-based local foods operations on average, this large estimate is indicative of how an additional establishment is a significant, and likely non-marginal, change to a local communities food environment.\textsuperscript{5} It also suggests a preference for living near direct-marketing farms themselves, not solely local food retail outlets.

[Insert Table 3 Here]

**Non-marginal Welfare Measures**

A core feature of the sorting model is the ability to calculate general (versus partial) equilibrium welfare measures that result from a proposed policy by comparing the ex-ante and ex-post equilibria. In a partial equilibrium framework households cannot move so house prices remain constant and willingness to pay (WTP) estimates reflect only the amount of money needed to compensate a resident for a change in amenities. This value can be calculated by solving the equation

\[ V_1(\alpha, y - \text{WTP}_{PE}, G_j^*, p_j) = V_0(\alpha, y, G_j, p_j) \] \[ \text{[8]} \]

A general equilibrium approach additionally recognizes not only that households can react to a change in the provision of amenities by moving to a new neighborhood but also allows for a characterization of how this resorting process results in individuals further inducing a change in
neighborhood price levels. A general equilibrium WTP resulting from an *exogenous* amenity change can be described as the value that solves

\[ V_1(\alpha, y - \text{WTP}_{\text{GE}}, G_k, p_k^*) = V_0(\alpha, y, G_j, p_j) \]  

[9]

where the change in subscript from \( G_j \) to \( G_k \) recognizes that households can move to a new location, thus affecting the community price as house prices will adjust due to changes in demand to clear the market. We denote these new prices by \( p_k^* \).

In table 4 we simulate a public policy of adding one local food establishment to the bottom 5% of communities, which corresponds to 30 additional local food operations. We vary this policy exercise by type of establishment and targeting approach. Specifically, we select communities based on either income or housing prices, and add either an on-farm or off-farm establishment, demonstrating how different types of community targeting lead to distinct welfare changes. The targeted communities are shown in figure 3.

[Insert Figure 3 Here]

When farmers’ markets are added to the lowest-income neighborhoods total welfare gains are $45.78 per year per household across the entire sample, although the average WTP is only $31.23 for residents of those communities. When we target the communities with the lowest housing prices the overall welfare gain is similar at $48.39, though community residents only value the change at $12.58. The relative reduction in welfare gains for the residents locating in the targeted communities compared to the average resident in the overall sample demonstrates how increased prices, driven by demand to live in these improved neighborhoods, reduces some of the welfare gains associated with the policy. This same pattern holds when we instead add an on-farm operation to the lowest-income communities, with a higher overall welfare value of $65.09
compared to an impact of $42.02 for households in targeted communities. Similarly, the addition of a farm-based direct marketing establishment increases overall welfare by $69.84 and welfare for households in targeted communities by only $20.51. While households appear to place a higher value on farm-based direct marketing operations, potentially due to the increased recreation options, adding either an on-farm or off-farm establishment generates resorting leading to price increases that reduce the welfare benefits for residents in the targeted communities.

When targeting the lowest income and price communities, the impact of rising prices is likely particularly acute as many residents are likely budget constrained and may have overall lower preferences for public goods if the local food goods are normal goods. To examine this further, we next simulated policies targeting the highest income and price communities. In these cases, the resulting WTP gains were actually larger in the targeted communities as residents in these communities were likely less budget constrained and have greater preferences for public goods, more than offsetting any demand-driven increases in prices.

5. Discussion
Our results suggest that the local food environment plays a key role in household location decisions alongside more traditional public goods measures. After controlling for school quality and open space, there is evidence of significant sorting behavior in response to variation in food access. Specifically, both the percentage of healthy food establishments and the number of local food retailers affect the perceived public good provision of a neighborhood. Put another way, households care if a neighborhood’s traditional food environment (retailers and restaurants) is
healthy and want to live near local food outlets. This is true for both farm-based and off-farm local food establishments, which suggest a space for direct-to-consumer farms along the rural-urban interface.

Interpreting these results, we should note that while heterogenous households sort based on their preferences for amenities, their decisions are still made conditional on income. Thus, there is an economic justice question as low-income households may be in low-ranked communities not because of a lack of preferences for healthy built environments but instead because of a lack of income. In fact, the large value residents of the best communities place on additional local food, relative to those of the lowest amenity neighborhoods, could potentially be explained by this difference. This finding reflects ongoing concerns regarding “green gentrification”.

While local food alternatives such as farmers’ markets have been used by food justice activists to improve food access in under-resourced communities, it is now understood that this can have unintended consequences if it attracts capital investments and higher income residents that lead to increased housing prices (Anguelovski, 2016; Alkon and Cadji, 2018; Cohen, 2022). Though some activists have proposed “just green enough” strategies to impede negative gentrification effects, others have suggested that whether an alternative food outlet plays an overall positive or negative role depends on market characteristics and the degree to which it is integrated into other local community efforts around collective activism and housing equity (Alkon, 2008; Cohen, 2018; Oths et al., 2019; Sbicca, 2019; Rosan, 2020; Fantini, 2023). While our results suggest an overall positive impact of introducing local food, with a lower value for those in targeted low-income and low-priced communities due to increased housing prices, there will potentially be heterogenous impacts in actual communities depending on choices made around
local food operation structure. Thus, any attempt to introduce local food operations as a policy mechanism must be made judiciously and with the support and input of local community members.

Our study has several potential policy implications. First, as these results show that households value local food establishments it is also likely that policies incentivizing local food would appeal to potential consumers; this very conclusion is bolstered by CDC recommendations that communities incentivize local food production (Khan et al., 2009). Second, this is the first study to identify a consumer preference for living in neighborhoods with healthier food establishments, as measured by the CDC’s mRFEI, which should be considered when creating economic development plans that call for new retail and restaurant operations. As a note of caution, while consumers positively value living near healthy and local food establishments, the evidence linking healthy food access to a healthy diet remains inconclusive (Jeffrey et al., 2006; Lopez, 2006; Boone-Heinonen et al., 2011). However, though this research suggests that an initiative to increase the availability of healthy establishments may not be effective if the goal is to improve health, none of these studies considered consumer preferences or included local food establishments. Future work is needed to assess the relationship between consumer preferences for local food access and health outcomes.

There are a variety of ways public policies can influence neighborhood amenities, including zoning regulations which dictate the nature and provision of food establishments. While our results unsurprisingly demonstrate that households take school quality and open space into account when making location decisions, this is not the only public good that influences location choice. Our study shows that households consider healthy and local food when choosing a neighborhood and demonstrates that policies to improve the local food environment positively
impact consumer welfare. This is a novel addition to our understanding of the household’s location decision-making process, and the first to assess the price and welfare impacts of local and healthy food amenities. Significantly, we also demonstrate the crucial role that neighborhood targeting plays in overall welfare outcomes.
References


Campbell, E. 2020. “Cleveland is now the poorest big city in the country.” The Center for Community Solutions. Available at: https://www.communitysolutions.com/cleveland-now-poorest-big-city-country/.


### Tables

#### Table 1: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price index for public goods</td>
<td>Estimated</td>
<td>3.77</td>
<td>1.99</td>
<td>1.00</td>
<td>11.95</td>
</tr>
<tr>
<td>Annualized Price: 25th percentile ($)</td>
<td>Data</td>
<td>7,638</td>
<td>6,240</td>
<td>1,500</td>
<td>33,440</td>
</tr>
<tr>
<td>Annualized Price: 50th percentile ($)</td>
<td>Data</td>
<td>11,140</td>
<td>8,630</td>
<td>1,950</td>
<td>47,707</td>
</tr>
<tr>
<td>Annualized Price: 75th percentile ($)</td>
<td>Data</td>
<td>15,628</td>
<td>11,571</td>
<td>2,140</td>
<td>63,800</td>
</tr>
<tr>
<td>Income: 25th percentile ($/year)</td>
<td>Estimated</td>
<td>24,960</td>
<td>12,430</td>
<td>3,000</td>
<td>98,920</td>
</tr>
<tr>
<td>Income: 50th percentile ($/year)</td>
<td>Estimated</td>
<td>43,630</td>
<td>23,550</td>
<td>6,130</td>
<td>293,430</td>
</tr>
<tr>
<td>School Quality (Reading Pass Rate)</td>
<td>Data</td>
<td>76%</td>
<td>14%</td>
<td>58%</td>
<td>97%</td>
</tr>
<tr>
<td>MRFEI (% Healthy)</td>
<td>Data</td>
<td>12%</td>
<td>11%</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Open Space (100s acres)</td>
<td>Data</td>
<td>1.33</td>
<td>4.26</td>
<td>0</td>
<td>72.77</td>
</tr>
<tr>
<td>Local Food Operations within 1/2 mile</td>
<td>Data</td>
<td>0.52</td>
<td>0.92</td>
<td>0</td>
<td>7.00</td>
</tr>
<tr>
<td>Farm-Based Operations within 1/2 mile</td>
<td>Data</td>
<td>0.14</td>
<td>0.56</td>
<td>0</td>
<td>4.00</td>
</tr>
<tr>
<td>Farmers' Markets within 1/2 mile</td>
<td>Data</td>
<td>0.37</td>
<td>0.68</td>
<td>0</td>
<td>4.00</td>
</tr>
<tr>
<td>Distance to nearest local food operation miles</td>
<td>Data</td>
<td>0.85</td>
<td>0.50</td>
<td>0.08</td>
<td>2.62</td>
</tr>
</tbody>
</table>

#### Table 2: Preference Parameters

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Error</td>
<td>Estimate</td>
</tr>
<tr>
<td>Standard Error of Income</td>
<td>0.8796</td>
<td>0.0108</td>
<td>0.8763</td>
</tr>
<tr>
<td>Mean Alpha</td>
<td>0.4025</td>
<td>0.1235</td>
<td>0.4837</td>
</tr>
<tr>
<td>Standard Error of Alpha</td>
<td>0.1588</td>
<td>0.0550</td>
<td>0.1838</td>
</tr>
<tr>
<td>Lambda</td>
<td>0.0009</td>
<td>0.0004</td>
<td>0.0001</td>
</tr>
<tr>
<td>Income elasticity (ν)</td>
<td>0.8335</td>
<td>0.0079</td>
<td>0.8208</td>
</tr>
<tr>
<td>Price elasticity (η)</td>
<td>-0.3408</td>
<td>0.0741</td>
<td>-0.2888</td>
</tr>
<tr>
<td>Demand intercept (β)</td>
<td>0.5888</td>
<td>0.0640</td>
<td>0.6342</td>
</tr>
<tr>
<td>Elasticity of substitution (ρ)</td>
<td>-0.0151</td>
<td>0.0044</td>
<td>-0.0159</td>
</tr>
<tr>
<td>Baseline public goods (Go)</td>
<td>0.4841</td>
<td>0.0399</td>
<td>0.4672</td>
</tr>
<tr>
<td>γOpen Space</td>
<td>0.0132</td>
<td>0.0041</td>
<td>0.0190</td>
</tr>
<tr>
<td>γMRFEI</td>
<td>0.6108</td>
<td>0.4402</td>
<td>0.4352</td>
</tr>
<tr>
<td>γLocal Food</td>
<td>0.1381</td>
<td>0.0695</td>
<td>--</td>
</tr>
<tr>
<td>γFarm-based Local Food</td>
<td>--</td>
<td>--</td>
<td>0.2150</td>
</tr>
<tr>
<td>γFarmers Markets</td>
<td>--</td>
<td>--</td>
<td>0.1138</td>
</tr>
<tr>
<td>γDistance to Local</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
Note: School quality (reading pass rate) is the baseline public good. Counts for local food, farm-based local food and farmers markets are within .5 miles of each census tract’s boundaries.

Table 3: Monthly Willingness to Pay Estimates

<table>
<thead>
<tr>
<th></th>
<th>MWTP Change</th>
<th>MWTP</th>
<th>Standard Deviation</th>
<th>MWTP for One Std. Dev Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Quality (% Pass Rate)</td>
<td>0.01</td>
<td>$6.84</td>
<td>0.14</td>
<td>$94.72</td>
</tr>
<tr>
<td>Open Space Area (100s acres)</td>
<td>0.1</td>
<td>$4.83</td>
<td>4.26</td>
<td>$205.74</td>
</tr>
<tr>
<td>mRFEI (% Healthy)</td>
<td>0.01</td>
<td>$17.46</td>
<td>0.92</td>
<td>$33.88</td>
</tr>
<tr>
<td>Farm Based Local Operations</td>
<td>1</td>
<td>$141.64</td>
<td>0.56</td>
<td>$79.62</td>
</tr>
<tr>
<td>Farmers Markets</td>
<td>1</td>
<td>$82.77</td>
<td>0.68</td>
<td>$56.05</td>
</tr>
<tr>
<td>Distance to Local (miles)</td>
<td>0.1</td>
<td>-$27.27</td>
<td>0.50</td>
<td>-$136.89</td>
</tr>
</tbody>
</table>

Table 4: Effect of Policy on Price

<table>
<thead>
<tr>
<th></th>
<th>Targeted Communities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bottom 5%</td>
</tr>
<tr>
<td><strong>Add Farmers' Market</strong></td>
<td>Overall</td>
</tr>
<tr>
<td>Income Targeted</td>
<td>$45.78</td>
</tr>
<tr>
<td>Price Targeted</td>
<td>$48.39</td>
</tr>
<tr>
<td><strong>Add On-Farm Operation</strong></td>
<td>Overall</td>
</tr>
<tr>
<td>Income Targeted</td>
<td>$65.09</td>
</tr>
<tr>
<td>Price Targeted</td>
<td>$69.84</td>
</tr>
</tbody>
</table>
Figures

Figure 1: Cleveland, OH Study Counties

Figure 2: Cleveland Neighborhood Price and Amenity Distribution

Figure 3: Policy Targeted Neighborhoods

\[i\] For further review see Alwitt and Donley (1997); Chung and Myers (1999); Morland et al. (2002b); Shaffer (2002); Zenk et al. (2005); Baker et al. (2006); Powell et al. (2007); Moore et al. (2008).

\[ii\] E.g., Cheadle et al. (1991); Morland et al. (2002a); Laraia et al. (2004); Rose and Richards (2004); Bodor et al. (2008); Zenk and Powell, 2008; Hilmers et al., 2012; Ohri-Vachaspati, 2019.

\[iii\] If \( U_i(G, h_n, b, \alpha) \) is continuously differentiable, monotonically increasing in the numeraire, and Lipschitz continuous.

\[iv\] As long as a \( h_n \) enters utility through a separable sub-function that is homogeneous of degree 1.

\[v\] The price index is normalized by the lowest price tract, and our values match closely with those found in other sorting papers in the literature (Klaiber and Morawetz, 2021).

\[vi\] We annualized housing prices using an assumed user cost of 11% (Poterba, 1984).

\[vii\] Open space is associated with increased housing values in other sorting contexts (Walsh, 2007; Klaiber and Phaneuf, 2010) and also represents an additional land use that can either complement or compete with urbanization and agriculture.

\[viii\] Some authors such as Sieg et al. (2004) find a negative value for lambda, which they attribute to the ability of higher income households to substitute private for public goods. Given the neighborhood heterogeneity present in our study region, as in Southern California for Sieg et al. (2004), a low or even negative value is to be expected.

\[ix\] Appendix Figure B1 shows the positive relationship between the public good index and house prices, confirming the expected relationship.

\[x\] Given the low number of local food operations generally, there is a concern that adding an establishment is non-marginal for both farm and non-farm based operations. As such, we also report the impact of a one standard deviation increase in each of these public goods. While we still see a stronger preference for farm-based operations, the difference is less stark.

\[xi\] These characteristics include whether the market is restricted to producers, the availability of culturally relevant products, the degree to which community activism occurs within the market, the pricing schemes employed by vendors and whether the market accepts food nutrition assistance vouchers or funds.