DRAFT WORKING PAPER

NOT FOR CITATION OR DISTRIBUTION WITHOUT AUTHORS' PERMISSION

# A COMPREHENSIVE METHOD FOR IDENTIFYING OPTIMAL AREAS FOR SUPERMARKET DEVELOPMENT

BY

Kennen Gross
Lance Loethen
Ira Goldstein

**The Reinvestment Fund**
718 Arch Street
Suite 300 North
Philadelphia, PA 19106
April, 2011

## Abstract

With scarce public resources being allocated to address inadequate access to supermarkets, it is important to first define what level of access constitutes an "underserved community". This requires a valid and reliable measure of existing access and the subsequent determination of thresholds that identify "underserved communities". Quantitative measurement of the problem is essential to ensure that scarce resources and interventions are distributed based on an objective measure of both need and market viability. In this paper we focus on quantitative measurement in relation to access to healthy food from a fresh food financing policy perspective. Although there exists ample academic, policy and advocacy literature devoted to documenting areas with limited access to retail grocers, wide variations in methodology still exist and, accordingly, there remains a lack of consensus over what constitutes a low-access area. In this paper we first review and critique existing literature. We then present a new methodology that builds upon earlier methods and adds advanced spatial statistical methods. Finally, we present the results of this methodology applied to all block groups in the continental United States. The results allow us to identify relatively small geographic areas as low access areas – rather than Census tracts, zip codes or larger areas - and provide estimates of the prevalence and characteristics of the population living in these areas.

## Introduction

Increasing access to supermarkets in underserved communities is a compelling goal for public policy and a compelling market opportunity for retailers. Research has shown that economies of scale in the retail grocery industry result in lower prices and a wider selection of fresh produce; larger grocery stores are also more likely to serve as anchors for job development, job growth, and attractive quality of life amenities compared to smaller stores. Research has also shown that specific market obstacles exist, limiting the development of supermarkets in low and moderate income areas. These obstacles include higher start up costs and higher operating costs due to workforce training and facility maintenance. Demand sufficient to support stores exist in some underserved areas, thus policy interventions that help overcome higher start-up costs have been shown to be a viable option for these areas.

Policy interventions would provide benefits that extend beyond food access. Research suggests that grocery retailers, in addition to satisfying local demand for food, can bring many economic benefits to the communities they serve.[1] For instance, research in the Philadelphia area has shown that the introduction of a new supermarket is correlated with an immediate boost of 4-7% in the value of homes within close proximity to the stores being studied.[2] Other research shows that supermarkets often employ workers primarily from the area in which they are located, providing a source of income to the community.[3] Additionally, supermarkets may serve as retail anchors for other businesses, spurring revitalization on a larger scale.[4] And by constructing new stores on previously underused land, as well as improving existing store properties, municipalities can realize increased property tax revenues. All of these economic development benefits can be particularly important in low access areas, which often face other economic challenges in addition to having low access to food.

The Pennsylvania Fresh Food Financing Initiative was developed to help address supermarket financing issues. The initiative improves existing, and establishes new, supermarkets or other grocery stores in underserved communities across Pennsylvania by serving the financing needs of current and potential supermarket operators where infrastructure costs and credit needs cannot be filled solely by conventional financial institutions. Similar state wide programs are being developed in New York, Louisiana, and Illinois. The Obama administration has also announced a federal healthy food financing initiative based on the Pennsylvania program.

With scarce public resources being allocated to address inadequate access to supermarkets, it is important to first define what level of access constitutes an "underserved community". This requires a valid and reliable measure of existing access and the subsequent determination of thresholds that identify "underserved communities". Quantitative measurement of the problem is essential to ensure that scarce resources and interventions are distributed based on an objective measure of both need and market viability. In this paper we focus on quantitative measurement in relation to access to healthy food from a fresh food financing policy perspective. Although there exists ample academic, policy and advocacy literature (beginning in the early 1990s) devoted to documenting areas with limited access to retail grocers (areas sometimes referred to as "food deserts"), wide variations in methodology still exist and, accordingly, there remains a lack of consensus over what constitutes a low-access area. In this paper we first review and critique existing literature. We then present a new methodology that builds upon earlier methods and adds advanced spatial statistical methods. Finally, we present the results of this methodology applied to all block groups in the continental United States. The results allow us to identify relatively small geographic areas as low access areas – rather than Census tracts, zip codes or larger areas - and provide estimates of the prevalence and characteristics of the population living in these areas.

*Areas with Low Access to Food:* The first question one is confronted with when investigating questions of food access is what constitutes a low access area. "Food desert" has been used often as a catch-all term in the healthy food access policy arena. A systematic review of the food desert research literature by Walker, Keane and Burke (2010) found that the phrase has been used differently by different researchers. This, they conclude has resulted in a lack of consensus on the definition of food deserts and the measures needed to identify food deserts, thus contributing to debate about the existence of food deserts. In the 2008 Farm Bill, Section 7527 defines a food desert as "an area in the United States with limited access to affordable and nutritious food particularly such an area composed of predominately lower-income neighborhoods and communities" A report entitled "The Public Health Effects of Food Deserts" issued by the National Academy of Sciences (REF) added geography and quantity factors to the Farm Bill definition and defined a food desert as "a geographic area, particularly lower-income neighborhoods and communities, where access to affordable, quality, and nutritious foods is limited." These varying definitions illustrate the lack of consensus on the words that should be used to define limited food access. From a public policy perspective, a useful definition would be one that could be measured quantitatively using existing data. Thus, coupled with the definitional problem is the problem of consensus on valid and reliable measures of each definition.

A review of existing research indicates methodological variations at almost every step of the process, and no consensus has emerged around any one definition or methodology. For instance, different studies consider different sources of food and discover their locations in different ways. Having determined an area's food sources, studies then use different methods to establish whether an area's food offerings are sufficient or not. Studies differ in their geographic focus with geographic size ranging from neighborhoods to cities to the nation, and have been conducted in different areas of the country. Some research has also considered the intersections of access with other variables, such as race or income.

Despite these wide variations in definition and measurement many studies claim strong evidence of significant disparities in access to food in different areas within the United States. A meta-analysis published in the journal of the Center for Disease Control and Prevention investigated food access through a systematic review on the subject.[5] The authors reviewed the findings of 49 studies from five countries, in order to determine "whether access to healthy, affordable food in retail stores varies by area socioeconomic status to the disadvantage of socioeconomically deprived areas" (1-2). Their findings noted that: "Geographic areas with a high proportion of low-income or African American residents were underserved by food retailers compared with more advantaged areas" (3). "Evidence," they say, "is both abundant and robust enough for us to conclude that Americans living in low-income and minority areas tend to have poor access to healthy food" (4). The researchers who conducted this review were comfortable with this finding despite the wide variations in measurement they encountered.

Another systematic review published in the American Journal of Preventive Medicine came to similar conclusions. This review, which considered 54 journal articles from 1985-2008, found that "despite some inconsistencies, several U.S. studies have shown that residents of rural, low-income, and minority communities are most often affected by poor access to supermarkets, chain grocery stores, and healthful food products."[6] Again, these authors acknowledged the wide variations in methodology they studied, yet were comfortable with concluding that disparities in access were real and substantial.

We believe that, although useful for establishing the general existence of a problem, these systematic reviews cannot accurately shed light on the locational prevalence of, characteristics of or the outcomes associated with low access area until measurement issues are addressed. What follows is a review of the literature pertaining to how measurement issues have or have not been addressed in the existing literature.

*Are supermarkets a good proxy for the availability of affordable, nutritious food?:* Much of the research concerned with food access takes the accessibility of large food retail outlets (e.g. "supermarkets," "chain stores," and "midsized or large stores") as a proxy for the availability of affordable, nutritious food. Although the literature contains widespread assumptions that larger stores such as supermarkets sell affordable healthy food, while small stores such as convenience stores sell predominantly unhealthy food, it is essential to establish whether such assumptions are true.

One recent study investigated this assumption by measuring shelf space devoted to fruits and vegetables in different types of stores in the New Orleans and Los Angeles areas.[7] In addition to comparing absolute shelf space devoted to fruits and vegetables, the authors compared fruit and vegetable shelf space relative to total store area and relative to shelf space for unhealthy items. On both of these relative measurements, the authors found that supermarkets and medium-sized food stores sold significantly greater proportions of healthy foods than did small grocery and convenience stores, with supermarkets reporting the highest ratios of healthy to unhealthy foods. The study concluded that "there were large variations between different store types and much smaller variations within individual store types in the absolute and relative shelf length of healthy and unhealthy food items," and that "measurements indicate that store type is a reasonable proxy measure for store contents" (679, 680).

Another recent study of food availability, this one in the Baltimore area, found similar correlations between store type and availability of a nutritious market basket.[8] Using a healthy food availability survey tool, the researchers computed a healthy food availability index (HFAI) for each store, which ranked each store's healthy food offerings on a scale of zero to 27.[9] They then computed averages for each store type: supermarkets, grocery stores, and convenience stores, as categorized in the InfoUSA directory. Results showed that most of the differences in healthy food availability were

accounted for between – not within – store types. For instance, the average HFAI score for supermarkets in neighborhoods of all of the racial and income brackets studied ranged from 18.67 to 24.0, while HFAI scores ranged from 3.85 to 6.17 for grocery stores and 3.78 to 5.25 for convenience stores. Differences in neighborhood-wide HFAI scores were due largely to the presence of different categories of stores in different areas, rather than to differences in offerings within store categories.

Several other studies have corroborated these findings that the availability of nutritious food varies more between than within store type, and that supermarkets tend to have the best selection and highest percentage of affordable healthy food.[10] This is not to say that smaller stores cannot be an adequate source of healthy food in some neighborhoods, especially in relatively prosperous areas[11], immigrant communities[12], or stores that have received targeted assistance to help them provide a wider selection of healthy food items.[13] Therefore, even if a supermarket is not present in a given area, additional research into a neighborhood food environment is necessary before undertaking investments to create or expand grocery retail. Such research could take the forms of, community surveys, interviews, or inspection of smaller food stores in the area.[14]

*Locating Supermarkets:* An accurate and comprehensive list of supermarkets is an essential first step when initiating supermarket access research. A store that exists but is left off the list will result in some areas being designated as low-access areas when they are not, constituting a "false positive." On the other hand, a store that has closed but is left on the list will keep some areas that are in reality low access from being designated as such, resulting in "false negatives." Researchers, then, have a strong interest in collecting supermarket data that is as accurate as possible, to avoid wrongly identifying areas as having or lacking access to supermarkets.

Researchers have used a variety of resources to locate supermarkets. These resources fall into roughly three categories: proprietary sources, public sources, and direct observation. Proprietary data

sources include the Dun & Bradstreet, ReferenceUSA, InfoUSA, and Nielsen Trade Dimensions databases, all of which offer nationwide data on supermarket locations. These proprietary data generally must be purchased, and vary in quality. Studies comparing the accuracy store location data from Dun and Bradstreet, InfoUSA and state government registry found discrepancies between the datasets [15,16]. Public data sources include Yellow Pages, other telephone directories, and state and municipal databases and records. Direct observation can be as simple as driving the study area and recording the locations of food stores, or can involve in-store observation as well.[17]

An additional limitation is the lack of information on store type within some datasets. If the goal of the measurement is to determine limited to full service stores, under the assumption that full service stores have a wider selection of fresh fruits and vegetables, then it is essential for a store dataset to include information on store type. Most of the data sources rely on SIC or NAICS codes to provide store classifications. A NAICS code of 445110 refers to "Supermarkets and Other Grocery (except Convenience) Stores" and is defined as "comprises establishments generally known as supermarkets and grocery stores primarily engaged in retailing a general line of food, such as canned and frozen foods; fresh fruits and vegetables; and fresh and prepared meats, fish, and poultry. Included in this industry are delicatessen-type establishments primarily engaged in retailing a general line of food."[18] Locations with this NAICS code would include both large supermarkets and small corner stores. For studies interested in measuring access to full-service stores, relying on data sources guided by NAICS codes prevents researchers from extracting full-service stores from the dataset.

*Estimating Access:* Having identified the locations of supermarkets in a given area, a researcher must then determine the degree of supermarket access experienced by populations in that area. An honest, accurate and operational definition of low access is essential for distinguishing those areas that are legitimately underserved by supermarkets from those that are not. In most studies, distance to a

supermarket is used as a proxy for food accessibility. Most often, distance is measured directly from point to point in Euclidean fashion, although some studies have measured distance along street grids to obtain the real-life travel or "network" distance.[19]

In using distance to estimate food access, researchers have typically used one or several fixed-distance measures from the food outlet, thus measuring "proximity." However, other measures have been used. For instance, one study used a measure which considered a supermarket accessible to the population within a Census block group only if the store is located within that block group.[20] Others have tried to account for the presence of multiple accessible food retailers in an area ("variety"), either through a simple count of retailers within a fixed radius[21], or by creating a "heat map" which reflects both distance to food retail and the presence of multiple food retail outlets.[22] One study employed a metric called "competition," which looked at the mean distance to supermarkets belonging to different chains and parent companies.[23]

The use of distance as a proxy for food access, although intuitive, presents a significant problem, which is that there is no one distance to a supermarket that is appropriate for all neighborhoods. Neighborhoods vary widely in terms of their built environment and in their residents' access to different types of transportation, such that a reasonable distance to food for residents in one neighborhood might be a prohibitive distance for residents in another. Some past research has recognized this fact, and proposed different distance buffers for walking and driving. Some research takes account of transit access[24] or neighborhood walkability[25] in order to correlate those features with observed distance to food sources. However, no prior research that we know of has systematically varied the appropriate travel distance for different neighborhoods based on both built environment characteristics and residents' access to transportation.

*Relative versus Absolute Measures and Scale:* Store locations, population centers and distance between the two represent the raw materials for measuring food access. What is done with these raw materials requires careful consideration and constitutes the measurement theory aspect of the research. The first of these considerations is whether to use a relative measure or an absolute measure of access.

A relative measure is one that compares the level of access in each geographic area to the level of access for one sub-group of the total study population. This subgroup constitutes a reference group. For example, Powell, Slater et al. (2007) compared the level of access in low income zip codes to the level of access in the high income reference group. Low access in this particular study is therefore relative to the access experienced by the high income group in the area. Absolute measures of low access quantify areas based on a fixed distance threshold. The recent USDA[26] study uses an absolute measure by defining low access as areas where distance to a grocery store is greater than one mile.

The choice of relative or absolute is important because each can provide vastly different low access population estimates. Additionally once this decision has been made, careful consideration of the reference group (in the case of relative measure) and the absolute cut-off (in the case of an absolute measure) as this selection can also greatly impact the population estimates. Once relative versus absolute issues have been decided, the issues of measurement scale must be considered. The low access scale could be nominal or ordinal. A nominal low access scale would be one where an area is identified as low access with a yes or no. An ordinal score is one where higher scores indicate higher levels of low access than low score areas.

*Low Access Areas, Grocery Leakage, and Substantial Markets:* Despite the large amount of research dedicated to identifying areas with low access to food, relatively little research goes beyond the designation of an area as a food desert to a description of the market characteristics within these areas. Understanding the economic landscape within a given low access area is important because this

knowledge can help identify viable strategies for increasing access in this area. Most important from a retail perspective is the local demand that is un- or inadequately met by current retail options. This unmet demand is presumably "leaked" out of the area, resulting in a loss of economic activity, jobs, and perhaps tax revenue (if sales are made in a neighboring jurisdiction), and it is commonly known as "retail leakage."

Different methods have been used to estimate both general retail leakage and leakage broken down by sectors (e.g., grocery leakage).[27] The most widely used method for finding leakage compares sales receipts within a given area with projected spending in that area. Actual sales figures are collected from the U.S. Economic Census[28]. Estimates of potential spending are based on data collected by the U.S. Department of Labor's annual Consumer Expenditures survey, which collects information on the purchasing habits of households in different income categories. These estimates of potential spending are then weighted by population to calculate an overall demand for the area.

Leakage, therefore, compares actual sales figures with potential sales that an area could support. If actual sales are lower than projected sales, spending is presumably leaking out of the area in question, taking jobs, economic activity, and perhaps tax revenue with it.[29] Because leakage measurements assign a dollar amount on unmet demand, they can suggest the precise scale of new retail that is needed in order to satisfy local demand. Leakage figures are especially important in the case of areas with low access to food, because while it is important to ensure that all low access areas are remedied and provided with adequate access to food, not all low access areas have the market potential to support a new full-service supermarket. Leakage provides a way to distinguish between low access areas that could support a full-service supermarket, were it present, and those for which other forms of assistance (such as mobile produce trucks, or incentives for corner stores to stock fresh produce) might be more appropriate.

Understanding the market conditions within low access areas is an essential step for determining appropriate solutions to the access problem in these areas. For those areas which can support a new full-service supermarket, the potential economic benefits that grocery stores can bring to their communities serves as a strong incentive. However, not all areas can support a new full-service supermarket. Leakage calculations can help determine what solutions are appropriate and viable for low access areas, and are an important tool for describing low access areas once they have been identified.

## Methodology

*Definition of areas with low access to food:* This study defines inadequate supermarket access as neighborhoods where the residents must travel significantly further than residents of the same density and car ownership neighborhoods to reach a full service supermarket. What follows is an explanation of how this definition was operationalized into a quantitative low access measure for the United States

*Locating Supermarkets:* Nielsen's *Trade Dimensions* database was used to locate supermarkets nationwide. The only other study we know of that attempted to map areas with low access to food for the entire nation, the USDA's Congressionally-commissioned report on the question of food access, also used the *Trade Dimensions* database to locate supermarkets.[30] *Trade Dimensions* is a preferable source to Dun & Bradstreet, ReferenceUSA, and InfoUSA because, unlike these commercial datasets, it is an industry database specific to food retail outlets (supermarkets, small grocery stores, convenience stores, wholesale clubs, etc.) which gathers data from a variety of sources rather than relying singularly on self-reporting mechanisms. *Trade Dimensions* offers the advantage over the public data sources and direct observation methods that it is a nationwide database, and so the same data can be used to study the entire country. As a professional resource for the grocery industry, *Trade Dimensions* has the most reliable, up-to-date data on store location and type, with over 49,000 food

store locations nationwide. Because it is the most reliable data source available it was used for our nationwide analysis of low access areas.[31] By using nationwide data, the potential error introduced by edge effects was negated since geographic boundaries did not define the inclusion of stores in the study.

Use of *Trade Dimensions* data adds an additional benefit of having each store classified by store types. These types and the Trade Dimensions definitions are as follows:

> **Supermarket – Conventional** *(volume greater than $2 million):* A supermarket is a full-line, self-service grocery store with annual sales volume of $2 million or more. This definition applies to individual stores regardless of total company size or sales, and therefore includes both chain and independent locations. Trade Dimensions utilizes the trade channel definition endorsed by FMI (The Food Marketing Institute) and the leading industry publications. FMI is a nonprofit association of 1,500 food retailers and wholesalers, their subsidiaries and customers. *Examples: Kroger, Food Lion, IGA, Cub Foods*
> **Supercenter:** A supercenter is a retail unit with a full-line supermarket and a full-line discount merchandiser under one roof. May have separate or combined checkouts. Examples: Wal-Mart Supercenter, Meijer Supermarket
>
> **Supermarket – Limited Assortment:** A limited assortment supermarket has a limited selection of items in a reduced number of categories. These stores typically offer every day low pricing. Principal differentiation from a conventional supermarket is often in the reduced size and depth of produce and non-food categories such as Health and Beauty Care (HBC), cleaning supplies, paper products and general merchandise. A limited assortment supermarket has few, if any, service departments, and less product variety than a conventional supermarket. Examples: Aldi Food Store, Save-A-Lot
>
> **Natural/Gourmet Foods:** A natural or gourmet foods supermarket is a self-service grocery store primarily offering natural, organic or gourmet foods. These stores will either focus product offerings around healthy living with fresh produce and natural products, or around gourmet food preparations with upscale oils, spices, cheese, meat and produce. Natural/gourmet foods supermarkets typically have expanded fresh food departments and/or prepared food selections. These supermarkets also typically have a limited selection, if any, of Health and Beauty Care (HBC) and general merchandise. A natural/gourmet foods supermarket does not have over 50 percent of product offerings in one category, as is the case with traditional butcher shops, delis, produce stands or nutritional supplement stores. Note: Ethnic supermarkets are not considered natural/gourmet foods supermarkets. Examples: Trader Joe's, Whole Foods, Dean & DeLuca
>
> **Warehouse Store:** A warehouse store is a grocery store with limited service that eliminates frills and concentrates on price appeal. Items are displayed for sale in their original shipping carton rather than placed individually on shelves. This type of store also sells bulk food and large size items. Examples: Cash & Carry, Smart & Final
>
> **Military Commissary** A commissary is a grocery store operated by the U.S. Defense Commissary Agency within the confines of a military installation. A commissary can fit within any of the grocery formats. Examples: Fort Hood DECA Commissary, Fort Riley DECA Commissary
> **Superette/Small Grocery** (volume less than $2 million): A superette is a grocery store with a sales volume ranging from $1 to $2 million annually. Typically superettes are independent, but many are affiliated with groups like IGA, Inc. Small grocery is defined as a grocery store with sales below $1 million annually. Also known as "Mom & Pop" stores. Examples: Country Market, Superior Markets. (pp 13-14)

This study used Supermarket – Conventional, Supercenter, Natural/Gourmet Foods, and Warehouse Store as proxy for the availability of affordable and nutritious food. Our assumption based on the literature is that "Superette/Small Grocery" and "Supermarket – Limited Assortment" typically do not provide a full line of perishable foods. Stores of this type that do provide fresh fruits and vegetables, it is assumed, do so at prices higher than those offered at larger stores. Military Commissary stores were not included in the analysis since access is only available to the military and including these stores would run the risk of creating access to geographic areas that are proximate to these locations but residents of those areas are not able to access that food retail opportunity.

We believe that the inclusion of the set of stores described above is justified because supermarkets have the greatest likelihood of offering large quantities of healthy food at reasonable prices. We do not include smaller stores that provide adequate access to food because of the difficulty and subjective nature of distinguishing between small stores that provide healthy items and those that do not, and because of the lack of national-level data addressing this question. However, we do recognize that supermarkets are not the only viable source of affordable nutritious food. Accordingly, we describe our methodology as a first step toward identifying areas with low access to food, rather than as a definitive statement. Before planning an intervention in an area that our methodology designates as low-access, local actors should take additional steps to confirm our identification of food inaccessibility in the area, such as interviewing neighborhood residents or conducting a community food assessment.[32]

*Estimating Access:* As discussed above, one limitation of prior food access research is that much of it has taken a "one size fits all" approach when it comes to access, assuming that one distance to food is universally appropriate for all neighborhoods; we assert that this is not the case, and that

reasonable distances will vary based on the population density and ca ownership rate of its residents. To address this problem, this methodology systematically estimates the typical travel distance for Census block groups based on the population density and rates of car ownership in those block groups. It is important to control for population density because, all else being equal, increased population density tends to increase consumer demand in an area and make it possible for an area to support a higher concentration of commercial ventures. In a high-density area, one would expect travel distance to supermarkets to be relatively shorter. Rates of car ownership, meanwhile, are important to control for both because vehicle ownership drastically changes travel time (which is the true indicator of access that distance is trying to account for) and because car ownership can significantly influence food shopping habits.[33]

*Relative Measures:* It is difficult to define poor access unless we have some definition of good access. Therefore the goal of this method was to define "good access" and then create a relative measure of food access while controlling for population density and car ownership. In order to control for population density and car ownership, we divided Census block groups into fourteen classifications based on a combination of these two variables. Density was ranked on a scale of 1 to 7, and car ownership was described as Low, Medium, or High.[34] Once block groups were divided into classifications, those block groups where median income was higher than 120% of Area Median Income (AMI) were separated, and the distance from each of these higher-income block groups' population-weighted centroid[35] to the nearest supermarket was calculated. The average of these distances for higher-income block groups was calculated, and this average was then taken as the benchmark for that category of block groups. For example, for *Suburban, High Car Ownership* block groups with greater than 120% of AMI, the **average** distance to a supermarket was 1.09 miles, but for one selected *Suburban, High Car Ownership* block group where median income was less than 120% of

AMI, the closest supermarket was 3.21 miles from the population-weighted centroid of that block group.

Three miles may not seem like a long way to go in a high-car, suburban neighborhood, but it is worth keeping in mind that these distances are not absolute measures, but rather are for the purposes of comparison. Distance is computed "as the crow flies," and so significantly underestimates actual walking or driving distance, which must occur along roads or sidewalks. In a standard suburban area without the dense street grid that is typical of many cities, the actual driving distance to a supermarket will be significantly farther than the straight-line distance reported by our methodology. Even in a city with a dense street grid pattern, travel distance from a block group centroid to a supermarket will be longer than the distance that our methodology indicates. The distances we report, therefore, are best used for the purposes of comparison, not as perfect descriptions of the distance to supermarkets.

High-income block groups' (block groups with median household income greater than 120% of area median) median distances to supermarkets are taken as an appropriate benchmark based on the assumption that these high income groups are more likely adequately served by supermarkets. We believe this assumption is valid given existing research describing the supermarket industry's intense level of competition. Furthermore, we believe that an empirically validated appropriate distance to supermarkets is more sensitive to real-world and market conditions than seemingly reasonable, but arbitrary, fixed-distance benchmarks, as are used in much of the existing literature.

Block groups are an appropriate resolution at which to consider food access because they capture detailed information about built environment and car ownership, while still representing a spatially concentrated population that could benefit from an intervention in the food environment.

They also happen to be the smallest unit of analysis for which detailed socioeconomic and car ownership data are available.

*Access Scores and Low Access Areas (LAAs):* For each block group an access score is calculated by subtracting the block group's distance to nearest store form the reference group distance for the car ownership/population density classification it is a member of. This difference is then divided by the block group's distance in order to provide the percentage farther a resident of that block group has to travel to get to a supermarket compared reference group residents. Take for example a block group that is assigned to the "6L" density/car ownership classification and has a straight line distance to a supermarket of .25 miles. The "6L" reference distance is .11 miles. The low access score is calculated as (.25-.11)/.25 = .57. This block group's travel distance would have to be decreased by 57% to equal the distance traveled by its non-LMI counterparts. This measure, calculated for all block groups, creates a quantitative measure of the degree to which a block group is underserved by supermarkets. Accordingly, block group low access scores are a ratio scale where zero represents a complete absence of an equity issue (where the distance is less than the reference distance) and higher scores indicate higher inequity compared to the reference group.

Block groups are next grouped into Low Access Areas (LAAs) via a statistical procedure known as the LISA method (Local Indicators of Spatial Association). In this context, LISA takes each block group with an access score and compares its score with that of its four closest neighbors (defined as the four block groups whose population-weighted centroids are closest to the population-weighted centroid of the block group in question), and combines those block groups into a LAA if their access scores are sufficiently similar. This step in the analysis represents a nominal measure of access. That is, we have assigned a yes or no to every block groups based on the block group being a member of spatial cluster

Once Low Access Areas have been defined, they are assigned weighted access scores by taking the population-weighted average of the access scores of the block groups that make up the Low Access Area. As with the block group access scores, these provide a quantitative measure of the relative disadvantage of each Low Access Area. These scores allow us to represent the nominally classified LAA based on the severity of its access inequality.

*Grocery Leakage:* Grocery leakage for each Low Access Area was calculated in order to provide estimates of the amount of unmet grocery demand within each Low Access Area. Household income categories and their respective percentages of income spent on "food at home," provide estimates of total retail grocery demand in each LAA. Total grocery sales occurring within each LAA (from stores classified as "superettes" and "limited assortment stores") were subtracted from demand, resulting in estimates for retail grocery leakage.[36]

Because the access problem is better understood in terms of square feet, leakage dollar values were converted into square feet by using nationwide averages for grocery sales per square foot. The leakage measure, stated in terms of square feet, helps communities understand the extent of their grocery retail shortage, and start to think about the appropriate scale of the solution.

## Results

Table 1 shows the characteristics of each of the population density/car ownership category. Density 4 High Car is the most prevalent category with 26.2% of all block groups and 27.4% of total population. Block groups in this category have a median population in households per square mile of 1,807 and a median car ownership rate of 94%. This category represents suburban type areas. The block groups in this category that have median household income greater than 120% of area median income have a median distance to nearest full service grocery store of 1.12 miles. Thus, this value serves as the reference distance for this category to which the distance for each block group in the

category will be compared as part of the calculation of the block group Low Access Score. As expected the reference distances decline as the population density/car ownership classification goes from rural areas (Density 1 High Car) to urban areas (Density 7 Low Car).

Of all 206,181 block groups with population greater than 250 people in the lower 48 states, 128,583 (62.5%) had distances to the nearest supermarket that were less than the corresponding group distance. That is, these block groups had access distances equal to or better than their reference group, resulting in a negative low access score value. All of these negative values were assigned a low access score of zero. Table 2 shows the distribution of block groups by low access score as well as the number of block groups in each score band that was identified as being part of an LAA during the spatial clustering step of the analysis. For example 5.6% of all block groups received a low access score between 1 and 10, yet less than 1% of these block groups were part of a low access cluster. Stated differently, these are isolated pockets of low access rather than a cluster of similarly low access block groups. While block groups with low access scores ranging from 51-60 represent 3.8% of all block groups identified as low access, they represent 20.7% of all block groups that are part of LAAs. These columns can also be compared to see the conversion rate from access score to LAA designation. For example, of the 649 block groups with a LA score of 91-100, 626 (96.5%) converted into LAAs.

The extent of the low access problem for each LAA can be quantified by calculating the population weighted low access score for each LAA using the access score of each LAA member block groups. Table 3 shows that over 60% of the LAAs have population weighted low access scores greater than 41. However, not all LAAs represent areas that have enough demand to warrant the development of a new store. Leakage can be used to quantify demand for a new store. In Table 3 the leakage threshold is set to the mean leakage for all LAAs such that those LAAs above the mean are the LAAs with high store demand. A total of 916 (27.0%) of all LAAs met this high demand criteria. The last column shows the number of block groups in each score category that met this high-demand criterion.

The largest number of high demand stores come from the 41-50 and 51-60 population weighted LAA category. LAAs from these two score groups represent 78% of all high demand LAAs.

*Table 1: Characteristics of each population density and car ownership category*

| Population density car classification[a] | Total # block groups | % block groups | Total population in households, 2000 | % total population | Median population in households per square mile (Census, 2000) | Median % no car (Census, 2000) | Reference distance[b] |
|---|---|---|---|---|---|---|---|
| Density 1 High Car | 8,643 | 4.3 | 8,341,948 | 3.0 | 9 | 5% | 8.27 |
| Density 2 High Car | 19,700 | 9.7 | 24,659,704 | 8.9 | 32 | 5% | 5.16 |
| Density 3 High Car | 43,625 | 21.6 | 65,199,415 | 23.6 | 245 | 6% | 2.31 |
| Density 4 High Car | 52,941 | 26.2 | 75,773,463 | 27.4 | 1,807 | 6% | 1.12 |
| Density 4 Moderate Car | 5,449 | 2.7 | 5,552,994 | 2.0 | 1,974 | 30% | .79 |
| Density 5 High Car | 34,576 | 17.1 | 45,974,570 | 16.6 | 4,126 | 7% | .73 |
| Density 5 Moderate Car | 6,723 | 3.3 | 7,141,485 | 2.6 | 4,438 | 33% | .61 |
| Density 6 Low Car | 13,724 | 0.7 | 20,685,915 | 0.5 | 10,916 | 63% | .21 |
| Density 6 Moderate Car | 1,469 | 3.6 | 1,398,643 | 3.2 | 9,872 | 31% | .39 |
| Density 6 High Car | 7,330 | 6.8 | 8,889,777 | 7.5 | 8,357 | 10% | .64 |
| Density 7 Low Car | 1,319 | 1.7 | 2,357,655 | 2.0 | 37,947 | 69% | .13 |
| Density 7 Moderate Car | 3,481 | 1.6 | 5,532,423 | 1.90 | 23,164 | 35% | .25 |
| Density 7 High Car | 3,276 | 0.7 | 5,256,968 | 0.9 | 21,032 | 12% | .59 |

| | | | | |
|---|---|---|---|---|
| Total | 206,181 | 100 | 276,764,960 | 100 |

Notes: [a] Only block groups with Census 2000 population >250 were classified. [b] Median straight line distance (miles) to nearest store from population weighted block group centroid for block groups with >120% AMI.

*Table 2 Block groups counts by Low Access Score and by Low Access Area clusters*

| Low Access Score | Block Groups (column percent) | Block groups identified as members Low Access Area clusters (row percent) (column percent) |
|---|---|---|
| 0 | 128,783 (62.5) | 0 (0.0) (0.0) |
| 1 - 10 | 11,541 (5.6) | 81 (0.7) (0.4) |
| 11- 20 | 13,505 (6.6) | 660 (4.9) (3.3) |
| 21-30 | 13,113 (6.4) | 1,828 (13.9) (9.2) |
| 31-40 | 12,287 (6.0) | 2,934 (23.9) (14.8) |
| 41-50 | 10,292 (5.0) | 3,842 (37.3) (19.4) |
| 51-60 | 7,760 (3.8) | 4,109 (53.0) (20.7) |
| 61-70 | 4,673 (2.3) | 3,079 (65.9) (15.5) |
| 71-80 | 2,415 (1.2) | 1,703 (70.5) (8.6) |
| 81-90 | 1,163 (0.6) | 948 (81.5) (4.8) |
| 91-100 | 649 (0.3) | 626 (96.5) (3.2) |
| Total | 206,181 (100) | 19,810 (9.6) (100) |

Ultimately this analysis identifies block groups across the United States that are members of Low Access Areas. These block groups can be aggregated to larger geographies in order to provide overall low access profiles of for the U.S., states or metro areas. Table 4 depicts the results aggregated to nationwide and by Census metro categories. These results show that there are approximately 23 million people in low access areas representing 8.5 % of the U.S. population. Of all four categories "Rural" has the largest percent (12.3%) of its category population in Low Access Areas. However, it is the "Major Metro" category that has the largest amount of leakage at $23.47 billion. In "Major Metro" and "Other Metro" areas residents of LAAs are more likely to be non-White than White. In "Metropolitan" and "Rural" there is basically no difference by race. The disparity by income is more clearly seen in the "Major Metro" category where 33% of the total category population is low income, while 49% of the categories LAA population is low income. "Other Metro" also shows income disparity as 20% of the total population is low income, but 37% of the "Other Metro" LAA population are low income.

Table 5 shows low access data aggregated to the state level. In order to provide a summary of the low access problem by income, each state was ranked on % total population in LAA, % LAA population in low income block groups, and % LAA population below 200% of poverty. A composite ranking was created by averaging the three rank values and then ranking the states by average value calculation.

The District of Columbia ranks 4$^{th}$ in the percent of its total population that reside in low access areas and 1$^{st}$ in both ranked income categories. This indicates that the low access problem in the District is experienced more by low income populations than by non-low income populations. The extent of this difference can be seen by examining the District's percentages. Of all of the block groups

classified as LAA, 92% of the LAA population lives in block groups where the median household income is less than 80% of area median income. Additionally, while 36% of the District's population lives in areas below 200% of poverty, 49% of the population in LAAs is below 200% of poverty. This indicates low income populations experience 13% more low access than expected.

*Table 3: Low Access Area by weighted average low access score and by leakage levels*

| Low Access Area Population Weighted Low Access Score | Low Access Areas leakage less than or equal to $13.7 million | # Low Access Areas with leakage exceeding $13.7 million | Total Low Access Areas |
|---|---|---|---|
| 1 – 10 | 0 | 0 | 0 |
| 11- 20 | 22 (100%) (0.9%) | 0 (0%) (0%) | 22 (100%) (0.6%) |
| 21-30 | 363 (96.8%) (15.1%) | 12 (3.2%) (1.2%) | 375 (100%) (11.1%) |
| 31-40 | 573 (85.3%) (23.8%) | 99 (14.7%) (10.1%) | 672 (100%) (19.8%) |
| 41-50 | 667 (61.9%) (27.7%) | 410 (38.1%) (41.7%) | 1,077 (100%) (31.8%) |
| 51-60 | 472 (56.7%) (19.6%) | 360 (43.3%) (36.6%) | 832 (100%) (24.6%) |
| 61-70 | 187 (68.5%) (7.8%) | 86 (31.5%) (8.7%) | 273 (100%) (8.1%) |
| 71-80 | 66 (82.5%) (2.7%) | 14 (17.5%) (1.4%) | 80 (100%) (2.4%) |
| 81-90 | 31 (96.9%) (1.3%) | 1 (3.1%) (0.1%) | 32 (100%) (0.9%) |
| 91-100 | 18 (90.0%) (0.7%) | 2 (10%) (0.2%) | 20 (100%) (0.6%) |
| Total | 2,399 (71%) (100%) | 984 (29.0%) (100%) | 3,383 (100%) (100%) |

Notes: The 2,467 low access areas that do not meet the $15 million leakage threshold could be for one of two reasons. The first is that LAA contains no full service stores and the total population is so small that the total dollars spent on food at home does not exceed the threshold. The second reason is that the LAA contains limed assortment and/or superette stores and the money being spent of food is not being leaked out of the area but rather spent in these non-full service stores. Areas meeting this second criterion might not warrant a new store, but might instead warrant the financing of the expansion of limited assortment/superette stores so they can offer a wider assortment of fresh foods.

*Table 4: Low Access Area Figures and Rankings by Metro Area Category*

| | Values | | | | | Percentages | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Metro Category37* | *Population Living in Occupied HH* | *LAA Population* | *Sales Leakage ($billion)* | *Leakage Per LAA Resident ($)* | *Sq Ft Leakage (million)* | *% Total Pop Non-white* | *% LAA Pop Non-white* | *% Total Pop in LAA* | *% of Total Pop in low income area* | *% of LAA Pop in low income area* | *% of Total Pop below 200% of poverty* | *% of LAA Pop below 200% of poverty* |
| Major Metro (1 million or >) | 145,982,206 | 11,839,533 | 23.47 | 1,983 | 36.2 | 31% | 43% | 8.1% | 33% | 49% | 27% | 37% |
| Other Metro (< 1 million) | 80,585,268 | 6,343,807 | 12.55 | 1,978 | 19.3 | 20% | 26% | 7.9% | 20% | 37% | 35% | 38% |
| Micropolitan | 28,416,833 | 2,724,681 | 5.26 | 1,932 | 8.1 | 15% | 14% | 9.6% | 28% | 15% | 31% | 37% |
| Rural | 18,647,141 | 2,291,872 | 4.05 | 1,767 | 6.2 | 14% | 14% | 12.3% | 48% | 49% | 39% | 42% |
| ***Nationwide*** | ***273,631,448*** | ***23,199,893*** | **-** | **-** | **-** | ***25%*** | ***32%*** | ***8.5%*** | ***31%*** | ***42%*** | ***30%*** | ***38%*** |

*Note: Population figures reflect 2000 Census data, aggregated from block groups*

The use of rankings help scale the extreme rural population issue as experienced in, for example, Montana. Montana ranks first among states in the category of percent population in low access areas. How much of this problem is due to Montana's very rural population rather than income disparity? When we factor in the state's ranking for percent LAA population in low income block groups (42%) and percent LAA population below 200% poverty (26%) the overall ranking becomes 22 indicating that the low access issue in Montana is a statewide issue rather than an issue of economic disparity.

As discussion of the previous tables has shown, the results of this analysis can be used to quantify the low access problem on the national or state level. Given that the basic building block of this analysis was the classification of Census block groups as members of low access areas (clusters of block groups with high low access scores); results can also be shown at the very local level and thus used to guide local level decisions.

Figure 1 shows a map of the Low Access Areas in Philadelphia; Table 6 contains data for each of these LAA geographies. A closer look at the data identifies areas that might have a sufficient market demand to support a new store and areas that while classified as underserved might not have sufficient demand to support a new store. An examination of the data for LAA id number 421010180008 (Area A on Figure 1) finds a collection of 8 Census block groups with a leakage rate of approximately $17 million dollars. Since leakage is high and there are no limited assortment stores in this area, a strategy for consideration might be the development of a new full service store in this LAA. LAA id number 421010262004 (Area B on Figure 1) has a different set of circumstances and thus might require a different intervention. This area is comprised of 48 block groups and has a leakage rate of approximately $65 million; however there are 7 limited assortment stores located within this area. An appropriate intervention for this area might be the expansion of one of the limited assortment stores

so that they can expand their selection of fresh foods. LAA id number 421010132001 (Area C on Figure 1) serves as an example of where a new store intervention would not likely be warranted. The LAA is comprised of three block groups with a high average LAA score (65). The leakage amount (approximately $1.8 million) is too small to support a new store. Support for the expansion of the superette in this cluster may be the most viable option.

*Table 5: Low Access Area Figures and Rankings by State*

| | Values* | | | | | Percentages | | | | | Rankings[38] | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *State* | *Population Living in Occupied HH* | *LAA Population* | *Sales Leakage ($million)* | *Leakage Per LAA Resident ($)* | *Sq Ft Leakage (000)* | *% Total Pop in LAA* | *% of Total Pop in low income area* | *% of LAA Pop in low income area* | *% of Total Pop below 200% of poverty* | *% of LAA Pop below 200% of poverty* | *% Total Pop in LAA* | *% LAA Pop in low income block groups*[39] | *% LAA Pop below 200% poverty* | ***LAA Composite Rank*** |
| DISTRICT OF COLUMBIA | 536,373 | 90,588 | 172.9 | 1,909 | 267 | 16.9% | 65% | 92% | 36% | 49% | 4 | 1 | 1 | **1** |
| LOUISIANA | 4,332,884 | 744,320 | 1,290.7 | 1,734 | 1,990 | 17.2% | 32% | 46% | 40% | 48% | 3 | 13 | 2 | **2** |
| PENNSYLVANIA | 11,847,607 | 1,522,561 | 2,992.9 | 1,966 | 4,615 | 12.9% | 28% | 52% | 27% | 38% | 7 | 5 | 21 | **3** |
| ILLINOIS | 12,096,973 | 1,244,366 | 2,217.3 | 1,782 | 3,419 | 10.3% | 32% | 55% | 25% | 38% | 15 | 2 | 19 | **4** |
| SOUTH DAKOTA | 726,335 | 135,962 | 201.7 | 1,483 | 311 | 18.7% | 29% | 34% | 33% | 45% | 2 | 29 | 8 | **5** |
| KENTUCKY | 3,926,781 | 354,965 | 646.4 | 1,821 | 997 | 9.0% | 35% | 44% | 36% | 45% | 20 | 16 | 6 | **6** |
| NEW MEXICO | 1,782,718 | 248,215 | 425.7 | 1,715 | 656 | 13.9% | 30% | 30% | 41% | 48% | 6 | 34 | 3 | **7** |
| NEW YORK | 18,395,693 | 1,566,393 | 3,003.4 | 1,917 | 4,631 | 8.5% | 35% | 53% | 31% | 39% | 23 | 4 | 17 | **8** |
| TEXAS | 20,290,302 | 2,173,310 | 3,990.5 | 1,836 | 6,153 | 10.7% | 36% | 41% | 36% | 43% | 14 | 20 | 11 | **9** |
| WEST VIRGINIA | 1,765,125 | 199,232 | 365.3 | 1,834 | 563 | 11.3% | 30% | 32% | 40% | 46% | 12 | 33 | 4 | **10** |
| OHIO | 11,053,789 | 1,133,182 | 2,188.4 | 1,931 | 3,375 | 10.3% | 28% | 47% | 26% | 38% | 16 | 11 | 23 | **12** |
| RHODE ISLAND | 1,009,430 | 127,986 | 271.2 | 2,119 | 418 | 12.7% | 31% | 50% | 27% | 34% | 8 | 6 | 36 | **11** |
| ARIZONA | 5,020,851 | 417,973 | 797.4 | 1,908 | 1,230 | 8.3% | 36% | 44% | 33% | 43% | 24 | 17 | 10 | **13** |
| TENNESSEE | 5,541,168 | 516,141 | 944.7 | 1,830 | 1,457 | 9.3% | 29% | 35% | 33% | 41% | 19 | 26 | 13 | **14** |
| UTAH | 2,192,656 | 95,559 | 154.5 | 1,617 | 238 | 4.4% | 30% | 48% | 28% | 46% | 45 | 9 | 5 | **15** |
| DELAWARE | 758,963 | 89,544 | 169.1 | 1,888 | 261 | 11.8% | 24% | 48% | 23% | 30% | 9 | 10 | 41 | **16** |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FLORIDA | 15,593,328 | 1,046,019 | 2,289.5 | 2,189 | 3,530 | 6.7% | 33% | 46% | 31% | 40% | 34 | 12 | 16 | **17** |
| MISSISSIPPI | 2,749,237 | 322,645 | 584.1 | 1,810 | 901 | 11.7% | 31% | 23% | 43% | 45% | 11 | 44 | 7 | **18** |
| CALIFORNIA | 33,052,189 | 1,899,947 | 3,514.7 | 1,850 | 5,420 | 5.7% | 33% | 45% | 33% | 41% | 38 | 14 | 14 | **20** |
| CONNECTICUT | 3,297,334 | 242,049 | 509.1 | 2,103 | 785 | 7.3% | 29% | 55% | 19% | 34% | 29 | 3 | 34 | **19** |
| MASSACHUSETTS | 6,127,400 | 721,947 | 1,635.5 | 2,265 | 2,522 | 11.8% | 27% | 45% | 22% | 30% | 10 | 15 | 43 | **21** |
| MONTANA | 877,381 | 772,386 | 1,699.1 | 2,200 | 2,620 | 88.0% | 26% | 26% | 37% | 37% | 1 | 42 | 26 | **22** |
| NORTH DAKOTA | 618,494 | 101,444 | 160.4 | 1,581 | 247 | 16.4% | 25% | 28% | 32% | 38% | 5 | 40 | 24 | **24** |
| OKLAHOMA | 3,338,046 | 287,782 | 534.1 | 1,856 | 824 | 8.6% | 31% | 30% | 37% | 42% | 21 | 36 | 12 | **23** |
| ALABAMA | 4,332,570 | 337,858 | 633.7 | 1,876 | 977 | 7.8% | 31% | 33% | 36% | 41% | 26 | 30 | 15 | **25** |
| WISCONSIN | 5,207,651 | 350,025 | 677.9 | 1,937 | 1,045 | 6.7% | 27% | 49% | 23% | 34% | 33 | 7 | 33 | **26** |
| SOUTH CAROLINA | 3,876,875 | 285,412 | 560.0 | 1,962 | 863 | 7.4% | 29% | 34% | 34% | 38% | 28 | 28 | 20 | **27** |
| ARKANSAS | 2,599,462 | 134,641 | 252.4 | 1,874 | 389 | 5.2% | 30% | 36% | 39% | 44% | 43 | 25 | 9 | **28** |
| MARYLAND | 5,162,381 | 526,211 | 1,080.9 | 2,054 | 1,667 | 10.2% | 31% | 41% | 21% | 30% | 17 | 22 | 42 | **30** |
| MICHIGAN | 9,688,463 | 731,352 | 1,470.6 | 2,011 | 2,268 | 7.5% | 30% | 43% | 25% | 34% | 27 | 19 | 35 | **31** |
| VIRGINIA | 6,846,071 | 316,226 | 665.9 | 2,106 | 1,027 | 4.6% | 31% | 49% | 25% | 35% | 44 | 8 | 29 | **29** |
| NEBRASKA | 1,659,105 | 166,748 | 261.0 | 1,565 | 402 | 10.1% | 30% | 33% | 28% | 35% | 18 | 32 | 32 | **32** |
| KANSAS | 2,606,372 | 177,037 | 330.6 | 1,867 | 510 | 6.8% | 29% | 39% | 27% | 35% | 31 | 23 | 31 | **33** |
| MISSOURI | 5,432,805 | 356,149 | 714.6 | 2,006 | 1,102 | 6.6% | 33% | 41% | 30% | 35% | 35 | 21 | 30 | **34** |
| GEORGIA | 7,952,484 | 433,537 | 816.1 | 1,882 | 1,258 | 5.5% | 33% | 35% | 31% | 38% | 39 | 27 | 25 | **35** |
| COLORADO | 4,198,094 | 219,883 | 450.1 | 2,047 | 694 | 5.2% | 36% | 44% | 24% | 33% | 42 | 18 | 37 | **36** |
| NEW JERSEY | 8,219,436 | 882,851 | 2,046.2 | 2,318 | 3,155 | 10.7% | 24% | 29% | 20% | 24% | 13 | 38 | 47 | **37** |
| WYOMING | 479,676 | 15,091 | 26.9 | 1,784 | 42 | 3.1% | 25% | 30% | 31% | 38% | 46 | 35 | 18 | **38** |
| NEVADA | 1,964,572 | 134,880 | 282.9 | 2,098 | 436 | 6.9% | 35% | 33% | 28% | 31% | 30 | 31 | 39 | **39** |
| NORTH CAROLINA | 7,795,305 | 486,677 | 968.4 | 1,990 | 1,493 | 6.2% | 28% | 30% | 31% | 36% | 37 | 37 | 27 | **40** |
| MINNESOTA | 4,783,531 | 307,790 | 690.2 | 2,243 | 1,064 | 6.4% | 29% | 36% | 22% | 26% | 36 | 24 | 45 | **41** |
| IOWA | 2,821,939 | 190,433 | 370.7 | 1,947 | 572 | 6.7% | 23% | 28% | 26% | 31% | 32 | 39 | 40 | **42** |
| INDIANA | 5,902,164 | 469,347 | 942.5 | 2,008 | 1,453 | 8.0% | 29% | 25% | 26% | 29% | 25 | 43 | 44 | **43** |
| NEW HAMPSHIRE | 1,200,159 | 102,360 | 241.7 | 2,361 | 373 | 8.5% | 21% | 22% | 19% | 20% | 22 | 45 | 49 | **44** |
| OREGON | 3,343,778 | 98,786 | 184.5 | 1,868 | 285 | 3.0% | 26% | 27% | 30% | 35% | 47 | 41 | 28 | **45** |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IDAHO | 1,262,397 | 31,706 | 45.8 | 1,445 | 71 | 2.5% | 23% | 13% | 34% | 38% | 49 | 48 | 22 | **46** |
| MAINE | 1,239,974 | 67,271 | 140.4 | 2,088 | 217 | 5.4% | 26% | 19% | 30% | 31% | 40 | 46 | 38 | **47** |
| WASHINGTON | 5,757,636 | 307,300 | 688.4 | 2,240 | 1,062 | 5.3% | 27% | 14% | 26% | 24% | 41 | 47 | 46 | **48** |
| VERMONT | 588,011 | 15,806 | 33.8 | 2,138 | 52 | 2.7% | 23% | 0% | 27% | 24% | 48 | 49 | 48 | **49** |
| ***Nation*** | ***271,851,968*** | ***23,199,893*** | | | | ***8.5%*** | ***31%*** | ***42%*** | ***30%*** | ***38%*** | | | | |

*Figure 1: Map of Low Access Areas in the City of Philadelphia*

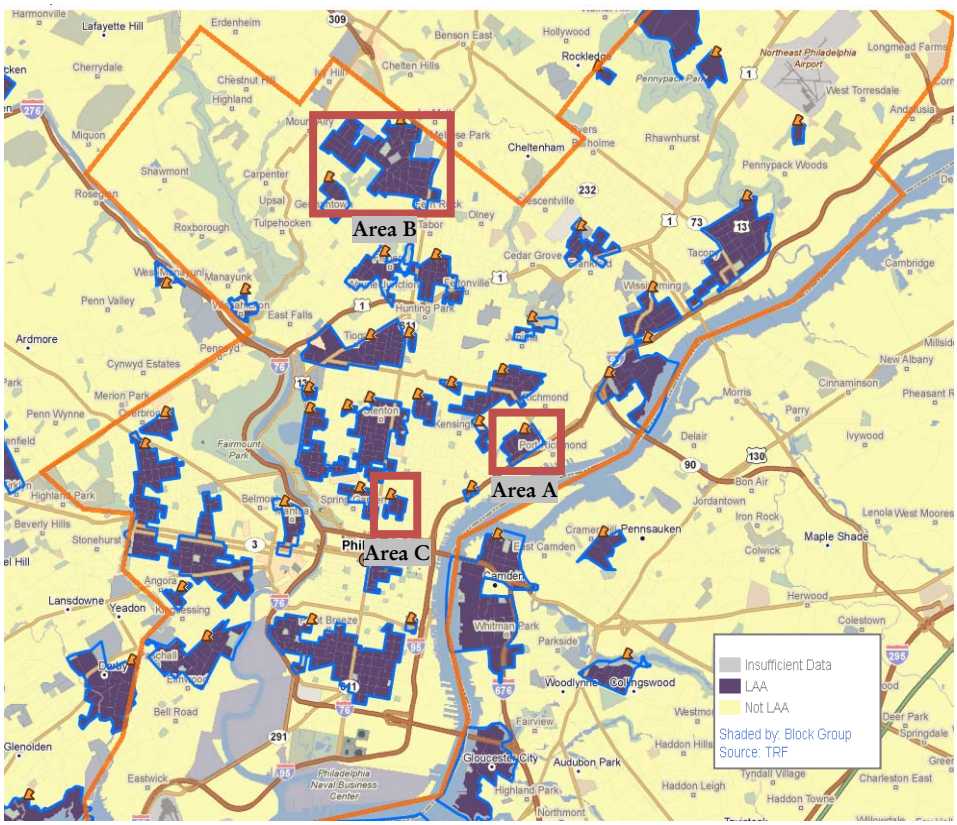

*Table 6: Data for individual Low Access Areas located in Philadelphia, PA*

| Low Access Area (LAA) ID | Average LAA Score (population weighted) | # Block Groups in LAA | Grocery Retail Leakage Amount | Grocery Retail Leakage Rate | Total Grocery Retail Demand | # Grocery Retail Sq Ft Leaked | # Limited Service Stores in LAA |
|---|---|---|---|---|---|---|---|
| 421010203004 | 44 | 1 | $1,265,515 | 100 | $1,265,515 | 1,951 | 0 |
| 421010262004 | 50 | 48 | $65,362,318 | 86.5 | $75,583,438 | 100,790 | 7 |
| 421010280001 | 42 | 11 | $12,446,303 | 79.2 | $15,711,383 | 19,192 | 2 |
| 421010045005 | 57 | 66 | $112,670,174 | 85.6 | $131,550,854 | 173,740 | 13 |
| 421010062005 | 54 | 20 | $23,182,244 | 88.6 | $26,163,404 | 35,747 | 2 |
| 421010101002 | 55 | 75 | $95,553,719 | 74.2 | $128,788,999 | 147,346 | 18 |
| 421010167002 | 62 | 36 | $22,444,343 | 66.8 | $33,611,863 | 34,610 | 7 |
| 421010169022 | 68 | 22 | $18,447,427 | 74.3 | $24,835,627 | 28,446 | 4 |
| 421010171004 | 59 | 21 | $27,419,494 | 94.8 | $28,933,734 | 42,281 | 1 |
| 421010176013 | 57 | 21 | $26,634,714 | 94.8 | $28,101,634 | 41,071 | 1 |
| 421010184001 | 51 | 8 | $10,745,403 | 87.6 | $12,259,643 | 16,570 | 1 |
| 421010191001 | 22 | 2 | ($3,328,151) | -132 | $2,521,849 | -5,132 | 1 |
| 421010209004 | 33 | 4 | $7,273,805 | 100 | $7,273,805 | 11,216 | 0 |
| 421010285002 | 55 | 18 | $17,681,983 | 79.1 | $22,366,663 | 27,266 | 3 |
| 421010168006 | 42 | 1 | $779,068 | 100 | $779,068 | 1,201 | 0 |
| 421010176024 | 61 | 1 | $645,494 | 100 | $645,494 | 995 | 0 |
| 421010169024 | 33 | 1 | ($3,613,032) | -128 | $2,822,488 | -5,571 | 4 |
| 421010009004 | 53 | 9 | $21,729,856 | 92.5 | $23,480,696 | 33,508 | 1 |
| 421010025005 | 45 | 1 | $1,897,094 | 100 | $1,897,094 | 2,925 | 0 |
| 421010071005 | 55 | 5 | $5,844,219 | 100 | $5,844,219 | 9,012 | 0 |
| 421010109002 | 62 | 12 | $13,420,522 | 100 | $13,420,522 | 20,695 | 0 |
| 421010118004 | 58 | 6 | $9,837,285 | 100 | $9,837,285 | 15,169 | 0 |
| 421010132001 | 65 | 3 | $1,782,213 | 100 | $1,782,213 | 2,748 | 0 |
| 421010135004 | 53 | 11 | $10,696,017 | 100 | $10,696,017 | 16,493 | 0 |
| 421010143004 | 37 | 1 | ($85,092) | -5.6 | $1,523,788 | -131 | 1 |
| 421010160008 | 48 | 3 | $4,133,516 | 100 | $4,133,516 | 6,374 | 0 |
| 421010164008 | 53 | 6 | $5,337,119 | 100 | $5,337,119 | 8,230 | 0 |
| 421010180008 | 54 | 8 | $17,002,245 | 100 | $17,002,245 | 26,218 | 0 |
| 421010252007 | 48 | 6 | $10,083,494 | 86.9 | $11,597,734 | 15,549 | 1 |
| 421010297003 | 77 | 1 | $830,213 | 100 | $830,213 | 1,280 | 0 |
| 421010303002 | 34 | 11 | $17,792,731 | 100 | $17,792,731 | 27,437 | 1 |
| 421010319003 | 50 | 14 | $20,397,626 | 100 | $20,397,626 | 31,454 | 0 |
| 421010331007 | 48 | 29 | $44,741,679 | 82.3 | $54,351,279 | 68,993 | 2 |
| 421010344004 | 48 | 3 | $9,526,360 | 100 | $9,526,360 | 14,690 | 0 |
| 421010345003 | 31 | 1 | $7,284,039 | 100 | $7,284,039 | 11,232 | 0 |
| 421010348024 | 27 | 1 | $1,374,615 | 100 | $1,374,615 | 2,120 | 0 |

## Discussion

Access scores and grocery leakage are two metrics which can help policymakers sort and compare LAAs. The access score describes the severity of the access limitation in a Low Access Area, while the leakage measure represents the volume of need. These two measures can help policymakers match appropriate policy interventions to areas that need the interventions most. For instance, an area with a high access score but relatively low amount of leakage might benefit from an intervention designed to bring fresh produce into corner stores, while an area with a moderate access score but high leakage might be able to support the opening of a new supermarket. Access scores and estimates of grocery leakage help policymakers classify LAAs according to the two characteristics which are most relevant from a policy perspective: the degree to which these LAAs are underserved, and their capacity to support a retail intervention.

The literature review is intended to present the reader with a general view of the methods that have been employed and problems researchers have encountered in defining areas with low access to healthy, affordable food. We hope that a deeper understanding of the methodological challenges associated with identifying low access areas will justify some of our method's assumptions and lead to a better understanding of our methodology for identifying these LAAs. However, we understand that our methodology is not the final word on food access, nor should it be. Specifically, we conclude with four caveats, highlighting aspects of the food access problem which our method does not address.

First, supermarkets are not the only possible source of affordable, nutritious food. Our method does not account for the possibility of affordable and nutritious food found in small grocery stores, corner stores, superettes, farmer's markets, produce stands or trucks, community gardens, food pantries, and other sources. As discussed above, we do not consider these possible sources of food because corner stores and superettes do not reliably provide an adequate selection of affordable and

nutritious food; for farmer's markets and the other sources there is no reliable nationwide data. However, we do recognize that these non-supermarket food sources may well provide healthy, affordable food in some areas. Accordingly, we describe our methodology as a first step toward identifying low access areas: this methodology directs attention toward areas that *may* have low access to affordable healthy food because they *are* underserved by supermarkets. Whenever possible, an area that is designated as low access should be investigated further, using local data sources and on-the-ground observation, before initiating an intervention in the neighborhood food environment.

Second, this methodology describes the relative accessibility or inaccessibility of *spatial areas* to supermarkets, not the accessibility of specific individuals. Some individuals living in well-served areas may nonetheless have limited access to food, such as the elderly or disabled (who often have diminished mobility), or individuals who live in an area with high car ownership but do not themselves have access to a vehicle. Our method cannot account for such low-access individuals, who live in an area that is adequately served but are not well-served themselves. For these individuals, some form of targeted assistance, such as a transportation stipend, might be the most appropriate solution. Again, this is an area in which local actors could supplement the LAA analysis with additional information more readily available at a local level.

Third, our method locates areas with relatively low access to food, but is silent on the question of how much additional distance to a supermarket constitutes an unacceptable degree of inaccessibility. Policymakers and local stakeholders have to determine for themselves what percentage of additional distance to food is unacceptably far. This is not merely an empirical question: it depends not only on market conditions and the viability of policy interventions, but also on a *subjective* determination of how much farther is too much farther to expect someone to go to a grocery store. The method does not make this determination, but rather provides access scores for Low Access Areas so that

policymakers and community stakeholders can determine – for themselves – which areas warrant an intervention in their food environment.

Finally, this research does not answer the question of why low access areas exist. It does, however, identify small areas across the United States where the residents must travel significantly farther than residents of similarly situated neighborhoods in order to reach a full service supermarket. We define these neighborhoods as "low access." Data on these low access areas can be studied individually in order to determine what might have led to the low access status, and what is the market potential for a new store located in one of these areas?

In presenting this brief literature review and description of our method for identifying areas with low access to food, this contribution hopes to inform a growing debate on food access in America. Before we can remedy the problem of low access to food, we have to be able to identify, through a realistic and evidence-based approach, those areas in the United States where there is a real barrier to obtaining healthy and affordable food. We at TRF hope that our method for identifying these low-access areas is realistic and convincing, and that, by controlling for key features of the built environment, we have been able to tease out real disparities in access to food. In offering our methodology and the low-access areas it has identified, we hope to move the pressing questions in the food access debate forward from the realm of preliminary research (i.e., if and where low-access areas exist) into the realm of planning and implementing appropriate policy responses to this urgent national priority.

We are currently in the process of running a new analysis for the United State using current an update Trade Dimensions store location dataset (previous analysis used a December 2009 dataset) and updated car ownership, population density and income data available from the 2005-2009 American

Community Survey data and the 2010 Census (previous analysis used Census 2000 data). Accompanying this updated data will be the following changes to the methodology.

Previous analysis used straight line distance to between block group weighted centroid and store location. Two changes will be made to this step. First, distance calculation will start at the Census block rather than block group weighted centroid. This will allow different populations within the block group to have a different "nearest store" rather than assigning one store per block group. The block level distance data will be aggregated back to the block group level by calculating a population weighted (based on the block population) average distance for each block. Second, the distance to nearest store will be calculated based on network distance rather than straight line distance. This will improve the analysis by providing more accurate measures of the distance needed to travel over the road network in order to reach the nearest store.

Another planned change involves taking steps to ensure that block groups included in the references group are normally distributed with a limited amount of variance. Since our use of reference group distance, calculated from block groups greater than 120% AMI, is based on our assumption that high income areas are in adequately served by the full service supermarket industry, it is important that all block groups included in the reference group calculation reflect high income markets. For all block groups with AMI greater than 120%, we plan on calculating an entropy index based on population by income categories. Block groups with high entropy will be removed from the reference group calculation as these would block groups would not be reflect the high income areas implicit in our assumption.

A final change relates to the spatial clustering method used to identify LAAs. When using LISA analysis to identify spatial clusters of high low access score in the previous analysis we defined neighboring block groups as those that share a boundary. A closer examination of some of these

neighboring relationships reveal that while proximate the two block groups are not accessible to each other and should not be considered neighbors in the market sense. As a result, the revised analysis will identify neighbors as block groups that share a boundary and have a road network connecting the two.

**Endnotes:**

[1] Martinez Hernandez M. "Impact of Commercial Development on Inner City Revitalization: An Analysis of Projects in Boston." August 2001. Unpublished thesis for Massachusetts Institute of Technology Department of Urban Studies and Planning. Available to view at http://dspace.mit.edu/handle/1721.1/32220. Accessed 7/27/2010. pp. 114-16.

[2] The Reinvestment Fund. "Reinvestment Brief: The Economic Impacts of Supermarkets on their Surrounding Communities." Issue 4. Available at http://www.trfund.com/resource/downloads/policypubs/supermarkets.pdf. Accessed 7/27/2010.

[3] Goldstein I., Loethen L., Kako E., Califano C. "CDFI Financing of Supermarkets in Underserved Communities: A Case Study." The Reinvestment Fund, August 2008. Available at http://www.trfund.com/resource/downloads/policypubs/TRF_CDFI_SupermarketStudy.pdf. Accessed 7/27/2010. Pp. 35-36.

[4] *Ibid*, pp. 49-51.

[5] Beaulac J., Kristjansson E., Cummins S. "A Systematic Review of Food Deserts, 1966-2007." *Preventing Chronic Disease* 2009; Vol. 6 Num. 3. http://www.cdc.gov/pcd/issues/2009/jul/08_0163.htm. Accessed 7/29/2010.

[6] Larson N., Story M., Nelson M. "Neighborhood Environments: Disparities in Access to Healthy Foods in the U.S." *American Journal of Preventive Medicine* 2009; Vol. 36 No. 1, pp. 74-81. See p. 76.

[7] Farley T., Rice J., Bodor J.N., Cohen D., Bluthenthal R., Rose D. "Measuring the Food Environment: Shelf Space of Fruits, Vegetables, and Snack Foods in Stores." *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 2009; Vol. 86, No. 5, pp. 672-682.

[8] Franco M., Diez Roux A., Glass T., Caballero B., Brancati F. "Neighborhood Characteristics and Availability of Healthy Foods in Baltimore." *American Journal of Preventive Medicine* 2008; Vol. 35, No. 6, pp. 561–567.

[9] The HFAI assesses the availability of eight food groups, assigning points for higher availability of items within each group (up to a total of 27 possible points). The eight food groups assessed are: nonfat/low-fat milk, fruits, vegetables, low-fat meat, frozen foods, low-sodium foods, 100% whole wheat bread, and low-sugar cereals. More information on the HFAI may be found in the "Methods" section of Franco et al, pp. 562-63, and Table 1.

[10] See, for example: Glanz K., Sallis J., Saelens B., Frank L. "Nurtition Environment Measures Survey in Stores (NEMS-S): Development and Evaluation." *American Journal of Preventive Medicine* 2007; Vol. 32 Num. 4, pp. 282–289. Also see: Zenk S., Schulz A., Israel B., James S., Bao S., Wilson M. "Fruit and Vegetable Access Differs by Community Racial Composition and Socioeconomic Position in Detroit, Michigan." *Journal of Ethnicity and Disease* 2006; Volume 16, Winter, pp. 275-280.

[11] For instance, Franco et al. found that "In fact, several stores coded as grocery stores in predominantly white higher-income neighborhoods had a higher availability of healthy foods than did supermarkets in predominantly black neighborhoods." (565).

[12] Short A., Guthman J., Raskin S. "Food Deserts, Oases, or Mirages?: Small Markets and Community Food Security in the San Francisco Bay Area." *Journal of Planning Education and Research* 2007; No. 26, pp. 352-364, especially p. 359.

[13] Hee-Jung S., Gittelsohn J., Kim M., Suratkar S., Sharma S., Anliker J. "A corner store intervention in a low-income urban community is associated with increased availability and sales of some healthy foods." *Public Health Nutrition* 2009; Vol. 12, No. 11, pp 2060-2067.

[14] Good examples of this type of local research can be found in Franco et al., described and cited above, in Short et al., also cited above, which combined a market basket study, in-store observations, and store manager interviews, and in Bodor et al., which measured shelf space devoted to healthy foods in small stores and its relationship to consumption in four New Orleans neighborhoods. (Bodor J. N., Rose D., Farley T., Swalm C., Scott S. "Neighborhood Fruit and Vegetable Availability and Consumption: The Role of Small Food Stores in an Urban Environment." *Public Health Nutrition* 2007; Vol. 11, No. 4, pp 413-420.)

The USDA has published an example study covering Washington, D.C. that makes use of its Thrifty Food Plan, "a national standard for a nutritious diet at a minimal cost," to assess food access and affordability. (See: Andrews M., Kantor L. S., Lino M., Ripplinger D. "Using USDA's Thrifty Food Plan to Assess Food Availability and Affordability." *Food Review* 2001; Vol. 24, No. 2, pp. 45-53.)

[15] Kowaleski-Jones, Fan, J.X., Yamada, I., Zick, C.D., Smilth K.R., and Brown, B.B. (2009). Alternative Measures of Food Deserts: Fruitful Options or Empty Cupboards? National Poverty Center Working Paper.

[16] Rose D.J., Bodor N., Swalm C.M., Rice, J.C., Farley T.A. and Hutchison (2009). Deserts in New Orleans? Illistrations of Urban Food Access and Implications for Policy. National Poverty Center Working Paper.

[17] The most geographically extensive study to use direct observation covered six rural counties in Texas. While impressive, this approach is probably too resource-intensive to be a viable approach for mapping food sources nationwide. For a description of the methods these researchers employed in collecting store data directly, see Sharkey J. and Horel S. "Characteristics of Potential Spatial Access to a Variety of Fruits and Vegetables in a Large Rural Area." Paper presented at the joint USDA and National Poverty Center Conference on "Understanding the Economic Concepts and Characteristics of Food Access," January 23rd, 2009. Available at: http://www.npc.umich.edu/news/events/food-access/sharkey.pdf. Accessed 7/29/2010. p. 8.

[18] http://www.naics.com/Censusfiles/ND445110.HTM

[19] E.g., Apparicio P., Cloutier M., Shearmur R. "The Case of Montréal's Missing Food Deserts: Evaluation of Accessibility to Food Supermarkets." *International Journal of Health Geographics* 2007; Vol. 6, No. 4. Available at http://www.ij-healthgeographics.com/content/6/1/4. Accessed 7/27/2010.

[20] Kowalski-Jones et al. p. 14.

[21] Sparks A., Bania N., Leete L. "Finding Food Deserts: Methodology and Measurement of Food Access in Portland, Oregon." Paper presented at the joint USDA and National Poverty Center Conference on "Understanding the Economic Concepts and Characteristics of Food Access," January 23rd, 2009. Available at http://www.npc.umich.edu/news/events/food-access/sparks_et_al.pdf. Accessed 7/29/2010. Pp. 4-5.

[22] A heat map presents information on both proximity and variety by describing a continuous surface, in which "the value of any point on the surface is determined by the number of food outlets within a given search radius (with greater weight given to outlets that are closer to a given point)." (Definition found in Neckerman K., Bader M., Purciel M., Yousefzadeh P. "Measuring Food Access in Urban Areas." Paper presented at the joint USDA and National Poverty Center Conference on "Understanding the Economic Concepts and Characteristics of Food Access," January 23rd, 2009. Available at http://www.npc.umich.edu/news/events/food-access/neckerman_et_al.pdf. Accessed 7/29/2010. P. 6)

Heat maps, therefore, are able to depict finer gradations of access, because they are not tied either to pre-designated administrative divisions or to a fixed-distance buffer. However, other data, such as detailed demographic information, are often not available at such a detailed level, mitigating the advantages of the heat-map approach.

[23] Sparks et al., pp. 4, 6.

[24] E.g., Neckerman et al., pp. 9, 18.

[25] Kowalski-Jones et al. p. 11; Neckerman et al. p. 20.

[26] United States Department of Agriculture. "Report to Congress: Access to Affordable and Nutritious Food Measuring and Understanding Food Deserts and Their Consequences," June 2009. Available at: http://www.ers.usda.gov/Publications/AP/AP036/AP036.pdf

[27] Strother S., Strother B., Martin B. "Retail Market Estimation for Strategic Economic Development." *Journal of Retail & Leisure Property* 2007; Vol. 8, No. 2, pp. 139-152.

[28] http://www.Census.gov/econ/

[29] Williams C. "Rethinking the Role of the Retail Sector in Economic Development." *The Service Industries Journal* 1997; Vol. 17, No. 2, pp. 205-220.

[30] United States Department of Agriculture (USDA), Economic Research Service (ERS). "Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences." Report to Congress. June 2009. http://www.ers.usda.gov/Publications/AP/AP036/ Accessed 7/29/2010.

In order to locate supermarkets nationwide, the USDA combined the Trade Dimensions database with a national directory of SNAP-authorized food stores that met the definition of a supermarket. (See p. 131.)

[31] For more information on the Trade Dimensions retail site database, see Nielsen's websites at http://ca.nielsen.com/content/nielsen/en_ca/product_families/nielsen_trade_dimensions.html and http://www.plainvanillashell.com/rsdb.asp.

[32] Described in detail above: see endnote 13.

[33] For example, residents with access to cars might find it easier to shop at supermarkets less frequently, but purchase more food per trip, so that a longer travel time to food might be more acceptable for these populations that shop for food less frequently. One study found these shopping patterns in a case study conducted among low-income residents of Austin, TX. Clifton, Kelly. "Mobility Strategies and Food Shopping for Low-Income Families: A Case Study." *Journal of Planning Education and Research* 2004; No. 23, pp. 402-413.

[34] Not all combinations of density and car ownership were used to describe neighborhoods; for instance, TRF's criteria do not find any Extreme Low Density, Low Car Ownership block groups in the U.S.

[35] The population-weighted block group centroid is the point that represents the average location of all people in an area. Because the locations of individual people are not available, population-weighted centroids are calculated by using the population count at the Census block level. Based on how the block population is distributed within a block group, the location of the population-weighted centroid will vary.

The formula to find a population-weighted centroid is as follows:

The Weighted Mean Center extends to the following:

$$\bar{X}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \quad , \quad \bar{Y}_w = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}$$

where $w_i$ is the weight at feature $i$.

[36] Percentage of income spent on "food at home" for each household income category is obtained from the 2009 Consumer Expenditure Survey, a quarterly survey conducted by the Bureau of Labor Statistics (BLS) of the US Department of Labor that provides information on American buying habits.

[37] Metropolitan and micropolitan statistical areas (metro and micro areas) are geographic entities defined by the U.S. Office of Management and Budget (OMB) for use by Federal statistical agencies in collecting, tabulating, and publishing Federal statistics.. A metro area contains a core urban area of 50,000 or more population, and a micro area contains an urban core of at least 10,000 (but less than 50,000) population. Each metro or micro area consists of one or more counties and includes the counties containing the core urban area, as well as any adjacent counties that have a high degree of social and economic integration (as measured by commuting to work) with the urban core.

[38] Steps for creating the State ranking table:

1. Sum block group level LAA figures for LAA population, leakage, low income population (those living in block groups with median household income at or below 80% of AMI), and population living below 200% of poverty level
2. Rank municipalities based on their values for *% Total Pop in LAA*, *% LAA Pop in low income,* and *% LAA Pop below 200% poverty*
    a. For sub-state geographies, we used *LAA Population* to apply a weighting factor that is not normalized (i.e. relative percentage) to account for sheer scale – however, at the state level this doesn't make sense, as each state is likely to manage its own HFFI program (or equivalent)
3. Calculate a composite ranking that blends all four rankings mentioned I point 3, above
    a. *% Total Pop in LAA*, *% LAA Pop in low income,* and *% LAA Pop below 200% poverty* were each weighted at 33.3% of the composite
    b. This effectively makes income-related variables 67% of the composite rank, which avoids higher rankings for states with less income segregation and a significant number of affluent areas with residential development that is intentionally located far away from commercial retail establishments.

[39] Low income block groups are those where the block group median household income is $<$80% of the area median income.