USDA Economic Research Service
U.S. DEPARTMENT OF AGRICULTURE

Economic Research Service

Technical Bulletin Number 1964

March 2024

Utilizing the USDA's National Household Food Acquisition and Purchase Survey to Calculate a Household-Level Food Environment Measure

Michele Ver Ploeg, Benjamin Scharadin, Lauren Miller, Jeffrey Gonzalez, and Rodney Odom

Economic Research Service
U.S. DEPARTMENT OF AGRICULTURE

Economic Research Service
www.ers.usda.gov

Recommended citation format for this publication:

Ver Ploeg, M., Scharadin, B., Miller, L., Gonzalez, J., & Odom, R. (2024). *Utilizing the USDA's National Household Food Acquisition and Purchase Survey to Calculate a Household-Level Food Environment Measure* (Report No. TB-1964). U.S. Department of Agriculture, Economic Research Service.

Cover photo from Getty Images.

USDA Economic Research Service
U.S. DEPARTMENT OF AGRICULTURE

Economic Research Service

Technical Bulletin Number 1964

March 2024

Utilizing the USDA's National Household Food Acquisition and Purchase Survey to Calculate a Household-Level Food Environment Measure

Michele Ver Ploeg, Benjamin Scharadin, Lauren Miller, Jeffrey Gonzalez, and Rodney Odom

Abstract

Utilizing the National Household Food Acquisition and Purchase Survey (FoodAPS) in conjunction with Circana's OmniMarket Core Outlets and USDA's Purchase to Plate Suite (PP-Suite), the authors calculated a household-level Food Retail Environment Healthfulness Quality (FREHQ) measure. The measure approximates a household's exposure to a healthy food retail space, weighting each store within a 20-mile radius of a household by its distance from the household. This process is used to calculate the FREHQ measure for FoodAPS respondents. The FREHQ measure allows for greater household heterogeneity and more nuanced analyses of the influence of the food retail environment on household food spending, food security, and other diet-related health conditions. This FREHQ can be used in conjunction with the rich granularity provided by the 2012–13 FoodAPS data but is also intended to be calculated with updated data and used in conjunction with future FoodAPS data.

Keywords: National Household Food Acquisition and Purchase Survey (FoodAPS), Circana, IRI Retail Panel, Purchase to Plate Tool, food retail environment, household-level healthy food access, inverse distance weighting functions

Acknowledgments

The authors thank the following for their contributions: Chris Lowe of USDA, Economic Research Service (ERS) for technical support using the Purchase to Plate Suite to calculate store-level healthfulness measures; Jessica Todd of USDA, ERS for constructive suggestions of metrics for comparisons; Debbie Rubas, Jay Variyam, and Xiao Dong of USDA, ERS, and Caitlin Caspi of the University of Connecticut, Allison Karpyn of the University of Delaware, TusaRebecca Pannucci of USDA's Center for Nutrition Policy and Promotion, and Rebecca Nemec of USDA's Office of the Chief Economist for their reviews of the draft report; Katherine Li, Yajuan Si, and Brady West of the University of Michigan for providing a county-level food insecurity measures for comparison; Jill Clark of The Ohio State University and Alannah Glickman of the Gephardt Institute for Civic and Community Engagement at Washington University in St. Louis for providing expertise on developing food access measures; Elaine Symanski, Jeff Chaltas, and Chris Sanguinett of USDA, ERS for editorial and design assistance.

About the Authors

Michele Ver Ploeg is food assistance branch chief at USDA, Economic Research Service (ERS) and is the corresponding author. Benjamin Scharadin, assistant professor in the Economics Department at Colby College, is the lead author. Lauren Miller is an agricultural economist with USDA, ERS. Jeffrey Gonzalez is a mathematical statistician with U.S. Department of Labor, Bureau of Labor Statistics. Rodney Odom is a geographer with USDA, ERS.

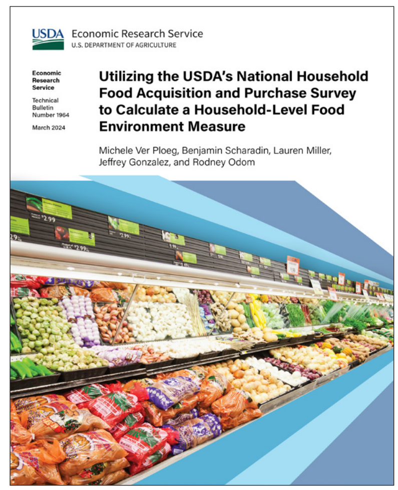
Contents

Economic Research Service
U.S. DEPARTMENT OF AGRICULTURE

A report summary from the Economic Research Service — *March 2024*

Utilizing the USDA's National Household Food Acquisition and Purchase Survey to Calculate a Household-Level Food Environment Measure

Michele Ver Ploeg, Benjamin Scharadin, Lauren Miller, Jeffrey Gonzalez, and Rodney Odom



What Is the Issue?

Many commonly used food retail environment (FRE) measures are based on the presence or number of food stores within a specified area. These measures often treat stores equally and do not account for factors such as varying distances and the healthfulness of available choices, which limit their accuracy and usefulness. Utilizing the National Household Food Acquisition and Purchase Survey (FoodAPS), in conjunction with Circana's (formerly IRI), OmniMarket Core Outlets (formerly IRI InfoScan), and USDA's Purchase to Plate Suite (PP-Suite), the authors developed a household-level Food Retail Environment Healthfulness Quality (FREHQ) measure to address these limitations. The measure approximates a household's exposure to a healthy food retail space by first proxying the healthfulness of the food inventory in stores near a household and then considering the location of each household relative to each store and discounting the weight of a store by its distance from each household. When compared with a homogenous geographic measure—a county-level, low-income, low-access measure—the new measure allows for greater household heterogeneity and provides opportunities to extend analysis on the impact of the food retail environment. The purpose of developing the FREHQ is to provide current and future FoodAPS users a household-specific measure of the FRE that may be used to consider the impact of the FRE on food spending, food security, and other outcomes in a more nuanced way in future studies. This study describes how USDA, Economic Research Service conceptualizes and calculates the FREHQ measure. The measure has advantages over other measures of the FRE in that it allows researchers to investigate the influence of the FRE for households that reside in the lowest quality food retail environments. Further, because it is a household and not a geographic measure, it allows researchers who use this measure to partially address endogeneity concerns by applying econometric techniques, such as geographic-fixed effects.

What Did the Study Find?

The authors constructed four versions of the household-level FREHQ, first using the straight-line versus road network distance between the household and store and then separately considering whether a household had access

ERS is a primary source of economic research and analysis from the U.S. Department of Agriculture, providing timely information on economic and policy issues related to agriculture, food, the environment, and rural America.

to a car. The average FREHQ for households that lived in an urban area is statistically significantly higher than the average for households that lived in a rural area in 2012–13 (the most recent available data). Differences in the average FREHQ were also observed across race, ethnicity, and income:

- Households with incomes above 185 percent of the poverty threshold had a higher average FREHQ measure across all FREHQ constructions.
- No statistically significant differences were found between the average FREHQ scores for White and Black households, with average scores close to 40. However, Asian and Hispanic households both have statistically significant higher FREHQ scores than White and Black households, with scores ranging between 41.3 and 42.4.

A component of the calculation of the FREHQ is to estimate the healthfulness of the products purchased in stores in each household's choice set—the collection of stores from which households can choose to shop. To do so, the authors calculated the FREHQ as a store-level 2020 Healthy Eating Index (HEI-2020) score, which ranges from 0 to 100, to reflect how closely a basket of food meets the 2020 Dietary Guidelines for Americans for each Circana retailer in 2012 and 2013 near FoodAPS households. These store-level estimations provide novel information about the healthfulness of stores by type:

- The mean HEI-2020 score of purchased products for grocery stores was 51.6. This score was close to the national HEI-2020 average scores estimated at the individual and household level based on food intake data, which follows intuition because a large portion of an average household's food at home is purchased at grocery stores.
- Dollar stores had an average HEI-2020 score of about 41.9. This suggests that in 2012 and 2013, the types of foods sold and purchased at dollar stores were dissimilar to the types of foods sold and purchased at grocery stores.

How Was the Study Conducted?

The FREHQ was constructed for each household in the FoodAPS, a nationally representative survey fielded in 2012 and 2013. Store locations from the OmniMarket Core Outlets for the corresponding period were used to calculate straight-line and network distances between stores and households. All stores within a 20-mile radius of each household were included in the sample. USDA, Economic Research Service's PP-Suite was used to link each food item in the OmniMarket Core Outlets with detailed nutrition information to calculate a store-level HEI-2020 score, which is used to measure store healthfulness in each household's FREHQ measure. Descriptive statistics for the four versions of the FREHQ were compared with expected trends based on past literature. While the data used to calculate the FREHQ for this study are from 2012–13 FoodAPS data, which are the only available source of comprehensive national-level data on household food purchase and acquisitions behaviors that allow linking to rich geographic information, future FoodAPS users can easily update the FREHQ as new data become available.

Utilizing the USDA's National Household Food Acquisition and Purchase Survey to Calculate a Household-Level Food Environment Measure

Introduction and Motivation

The food retail environment (FRE) continues to be implicated as a potential cause of differences in dietary outcomes and diet-related health (Black et al., 2014). A robust body of literature on this topic generally shows that residents of areas with easier access to food stores that offer healthy food options have healthier diets and lower rates of obesity (Larson et al., 2009; Caspi et al., 2012). Low-income and communities of color have more limited access to healthy food retail outlets, such as supermarkets, compared with higher income and predominantly White neighborhoods (Walker et al., 2010). In response, policymakers at all levels of government have created funding mechanisms to address lower healthfulness quality FREs (Holzman, 2010; Giang et al., 2008). Attention to the quality of the FRE has been reinvigorated with a new emphasis on nutrition security[1] and inequitable access to food (USDA, 2022).

In contrast, some studies that investigated the impact of the FRE found that limited access to healthy food explained only a small part of nutritional inequities (Allcott et al., 2019) and that the introduction of healthy retail outlets was not strongly associated with an increase in healthy food purchases (Cummins et al., 2014; Dubowitz et al., 2015). One possible reason for these contrasting results is that common methods used to measure exposure to the FRE have multiple methodological limitations. These include how the boundary of the FRE is defined, whether the measure accounts for how specific households interact with the FRE, how more proximate stores are weighted, and how the healthfulness of the food offered in stores is defined (Wilkins et al., 2019; Cummins et al., 2017; Ver Ploeg et al., 2015a).

For example, many studies define the FRE measure based on geographic and/or administrative boundaries that do not take into account a household's location in the boundary. This type of homogenous geographic measure artificially limits the collection of stores to which households are actually exposed in their local FRE because households do not travel only within their census tract, city, or county. These types of measures can mischaracterize a household's FRE exposure due to the Modifiable Areal Unit Problem (MAUP) (Scharadin et al., 2022). The MAUP occurs when the measure or the estimated impact of the measure changes with the boundary used to calculate the measure (Kwan, 2018). As a result, reviews of the literature have called for a move away from using administrative units to measure the FRE (Ver Ploeg et al., 2015a; Perchoux et al., 2013).

These homogenous geographic-based measures are also limited due to the Uncertain Geographic Context Problem (UGCP) (Scharadin et al., 2021), which occurs when a measure does not accurately capture how a household interacts with its FRE. Given that administrative boundary measures are calculated for a particular geographic area (e.g., county), all households in the geographic boundary have the same value for each FRE measure, which implies they have the same FRE exposure. However, due to household heterogeneity in location, ease of travel, etc., each household does not have the same exposure (Ver Ploeg et al., 2015b). As a result, FRE literature has called for household-specific measures that take into account both the location of the household and the locations of food retail stores. This echoes findings from geography studies that

[1] The USDA has defined nutrition security as having consistent and equitable access to healthy, safe, affordable foods essential to optimal health and well-being. Actions being taken to address nutrition insecurity can be found on USDA's Food and Nutrition Security web page.

call for decreased use of absolute measures, measures that consider only a single location, and increased use of measures that consider the relative position of at least two locations (Glickman et al., 2021). Considering the relative position between the household and store locations can better account for heterogeneity among households within a particular geography (Cummins et al., 2017).

Household-specific measures that consider the relative position between residence and store location attempt to capture the "activity space" of a household, the local area that people travel within during the course of their daily activities (Sherman et al., 2005). A household's "activity space" has been shown to better explain the relationship between exposure to the FRE and food consumption behavior but is often limited in use (Clary et al., 2015; Clary et al., 2017). Common examples of FRE measures that consider relative position are the household's distance to the nearest grocery store (Ware et al., 2021), the distance to the store used for a majority of shopping (Ver Ploeg et al., 2015b), and the distance between nearest fast-food and school locations (Davis & Carpenter, 2009). Although addressing the limitations with administrative boundary measures, these common measures still have limitations because they assess only exposure to one environmental "bad" (e.g., fast-food outlet) or "good" (e.g., grocery store). These measures would be improved by considering exposure to numerous types of retail food stores and how that exposure varies for households by proximity.

Finally, many studies use store type (e.g., supermarket versus convenience store) as proxies for the healthfulness of the food available in these stores. Glanz et al. (2016) summarized the large body of literature capturing the variety of in-store healthfulness measures. In-store measures, such as the Nutrition Environment Measures Survey (NEMS) and Market Basket Assessment Tool (MBAT), are often limited to a small locality because data collection that relies on store audits is impracticably labor-intensive on larger scales (Shaver et al., 2018; Hedrick et al., 2022). In addition, store willingness to participate in audits and share that level of information is not universal. Instead, store-type proxies that attempt to score a store type rather than an individual store are often used for national or larger-scale studies (Black et al., 2014).

The authors describe a new measure of the FRE that builds upon and improves existing measures in four ways:

(1) The measure captures the ease of access by considering the distance between a household and all store types in the FRE. This is an improvement from previous measures because it considers exposure to multiple "bads" and "goods" in an environment (Clary et al., 2015; Cummins et al., 2017).

(2) The measure is household-specific rather than geographic unit-specific in order to mitigate the impact of the UGCP and MAUP (Scharadin et al., 2021; Perchoux et al., 2013). The household-specific measure the authors developed has the additional benefit of being continuous (as opposed to the discreet "in a bad food environment or not," upon which many previous studies relied). This allowed analysis across the distribution of FRE quality, rather than analyzing the effect on the average household, and enabled more nuance in the analysis of policy levers targeted directly at households rather than areas.

(3) The measure uses store-specific information to create a proxy measure for store healthfulness that allows for greater variation within store type. Despite clear trends by store type, significant within-type variation exists if store-specific data are not used (Black et al., 2014; Leon et al., 2011).

(4) The measure is comparable across a variety of geographies and robustly measures the FRE for the location of the household.

One approach to address the limitations of other FRE measures is to consider inverse distance weighting functions (IDWFs), which are frequently used in environmental exposure literature to calculate the weighted impact of multiple sources of exposure on a single point (Paramasivam & Venkatramanan, 2019). When calculating exposure to a "bad," such as pollution, the increased distance to the "bad" decreases the impact of that pollution particle source (Liu et al., 2019; Blanco et al., 2018). Similarly, when considering access to a "good," such as educational services and public open space, increased distance limits the benefit of a particular source (Hu et al., 2013; Giles-Corti et al., 2005). A household's FRE, the food retailers available in an area (Glanz et al., 2005), resembles other environmental exposure applications because it contains numerous healthy food retail locations (e.g., supermarkets and grocery stores) and less healthy retail locations (e.g., gas stations and dollar stores). Calculating a household's exposure to the FRE using an IDWF addresses the four measure improvements listed. By discounting the weight of each store in a household's local FRE by the distance to that store, the authors calculated a Food Retail Environment Healthfulness Quality (FREHQ) measure that is household-specific and applicable across the United States.[2]

This USDA, Economic Research Service (ERS) Technical Bulletin discusses the construction of the FREHQ measure that leverages unique datasets—Circana OmniMarket Core Outlets (formerly IRI InfoScan); USDA, ERS's Purchase to Plate Suite (PP-Suite); and USDA, ERS's National Household Food Acquisition and Purchase Survey (FoodAPS).

Constructing the Food Retail Environment Healthfulness Quality Measure

The authors constructed a new FREHQ measure following an IDWF approach. This approach is used frequently to calculate an environmental exposure measure (Liu et al., 2019; Blanco et al., 2018) and investigate food retail and food store access topics. Botkins and Roe (2018) created an IDWF measure to show that increased access to farmers' markets increased the likelihood that a school participates in the USDA's Farm to School program. IDWF measures have also been used to demonstrate negative behaviors related to the FRE. For example, an increase in access to places that serve alcohol is strongly related to an increase in crime (Groff, 2014).

Most directly related to the authors' approach, Glickman et al. (2021) developed a proximate food retail quality (PFRQ) score to measure the quality of a household's FRE. The PFRQ score addresses multiple limitations in that it incorporates store audit data to capture the in-store environment, discounts the exposure of a household to a store by the distance between store and household, and includes all stores within a household's local FRE. However, the PFRQ was developed to assess two relatively small neighborhoods in Cleveland, Ohio, so the measure cannot be compared across locations with dissimilar population and store densities, nor are the findings generalizable.

The authors extended the work of Glickman et al. (2021) by creating a measure that does not rely on in-store audits and can be applied across a variety of geographies. Eq(1) presents the FREHQ measure.

$$FREHQ_i = \frac{\sum_{j=1}^{n} \frac{healthfulness_j}{distance_{ij}^{p}{}_{i}}}{\sum_{j=1}^{n} \frac{1}{distance_{ij}^{p}{}_{i}}} \tag{1}$$

[2] To further clarify, many FRE measures used by researchers classify all households within a geographic boundary (e.g., a census tract) as having the same FRE and the same exposure to each store within that geography. This measure, instead, is specific to each household in that it considers the FRE as all stores of different types within a radius of each specific household and weighs the exposure of each store by the distance of the store from the household.

In Eq(1), $healthfulness_j$ represents the healthfulness of store j; $distance_{ij}$ represents the distance between household i and store j; and p_i is a household-specific parameter representing the i^{th} household's barrier to travel. Intuitively, the FREHQ can be considered a weighted average of the healthfulness of food retailers within 20 miles of the household, where the weight of each retailer is the inverse distance between the household and store. As a result, the FREHQ measures the collective healthfulness quality of a household's community FRE, which is the food retailers available in an area (Glanz et al., 2005). This measure does not claim households shop at every store used to calculate the FREHQ, but rather that each store is in a household's available choice set.

Calculating the measure in Eq(1) requires detailed data on stores and households. The authors calculated the new FREHQ measure by using the Circana OmniMarket Core Outlets and the FoodAPS dataset. They used FoodAPS data for household location, which in conjunction with store location is used to calculate distances, and household-level characteristics to account for a household's barrier to travel. FoodAPS is a nationally representative survey of U.S. household food purchases and acquisitions collected between April 2012 and January 2013. Information on foods purchased and otherwise acquired for consumption at home and away from home was collected through barcode scanning, receipt validation, and food diaries.[3] Food spending estimates in FoodAPS are comparable with other food purchasing datasets, such as the Consumer Expenditure Survey and the Circana Consumer Network™ data (Clay et al., 2016). However, these two datasets cannot match the information on household economic and resource conditions, shopping behaviors, nutritional knowledge, and location and FRE, nor are they representative of low-income populations and those people who participate in the Supplemental Nutrition Assistance Program (SNAP). FoodAPS has been used in a growing body of literature on diet quality and food choice (Christensen & Bronchetti, 2020; Taylor & Villas-Boas, 2016; Dorfman et al., 2019; Scharadin & Jaenicke, 2020).

The authors used the Circana OmniMarket Core Outlets for store location and purchase information. Ideally, store inventory data would be used to measure what is offered in stores and then to characterize the healthfulness of those offerings. Using store offerings, rather than store sales, is essential to reduce endogeneity concerns in future analyses. Unfortunately, the authors did not have inventory data. Instead, they used sales data for 2012 and 2013 to create a list of products sold in each store that matches the collection timeframe of the FoodAPS sample. Although sales data capture only items bought rather than offered, using sales from a 2-year period should provide a relatively complete list of store offerings and account for seasonality trends. The healthfulness of the store offerings list was then measured using the PP-Suite[4] to calculate the 2020 Healthy Eating Index (HEI-2020) for each store.

The Circana OmniMarket Core Outlets provides weekly retail sales data about revenue and quantity information for products with Universal Product Codes (UPCs) and random weight products. This information is collected from a variety of store types, including grocery, drug, convenience, mass merchandiser, club, dollar, and defense commissary stores (although commissaries are not used in this study).[5] Although much of this information is provided at the store level, certain retailers provide information for a retail market area (RMA), a large geographic area containing numerous stores. Stores that provide information on the RMA level are usually large chains that have consistent offerings across stores in an RMA. As a result, the RMA-level data used for these cases can be considered a close proxy for store-level data for stores in

[3] For more information on FoodAPS see the USDA, ERS FoodAPS National Household Food Acquisition and Purchase Survey, available online.

[4] The Purchase to Plate Suite uses grocery store scanner data to estimate the retail cost of foods and beverages that participants reported eating in a national dietary intake survey (Carlson et al., 2020).

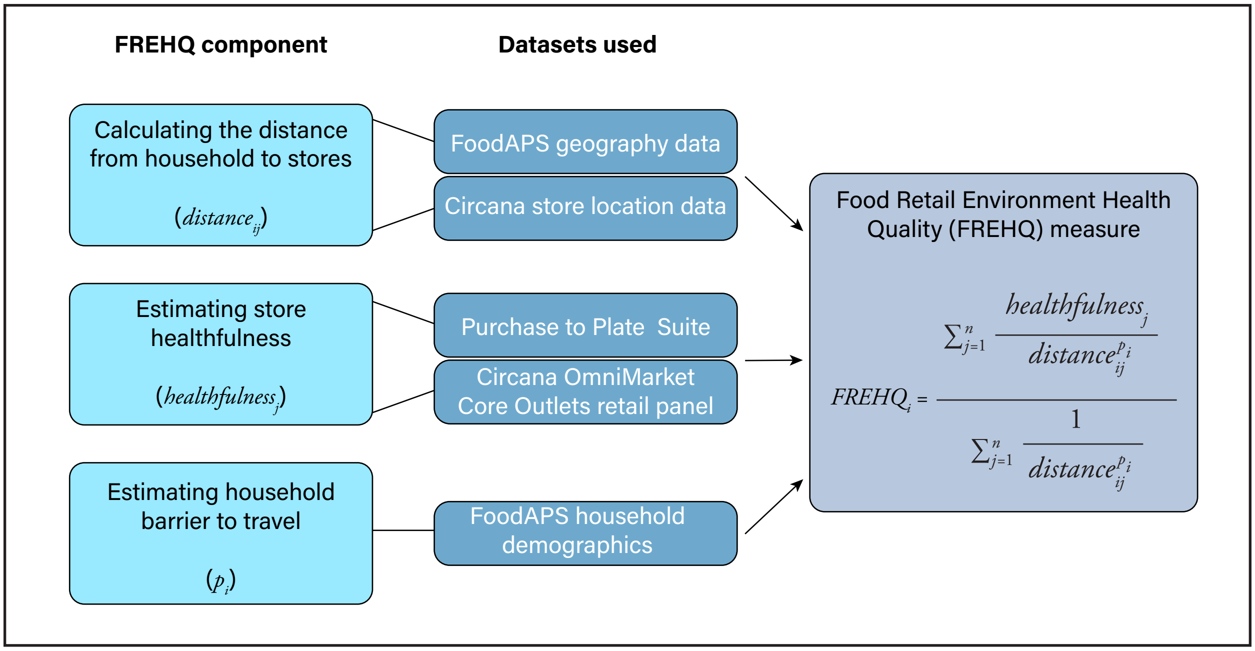
[5] Defense commissaries are important components of the food retail environment for military personnel but are not available to the public. Therefore, they are not included in the study because they serve a very limited role in the healthfulness quality of the average individual or household.

the RMA. Circana OmniMarket Core Outlets store locations were previously used in conjunction with FoodAPS household data for applications such as calculating food basket prices (Gunderson, 2016) and estimating the purchasing power of SNAP benefits (Christensen & Bronchetti, 2020). Muth et al. (2016) provided a detailed description of retail and household scanner data collected by Circana.

Figure 1 provides a general roadmap of how each dataset was used in the FREHQ calculation. (In the following sections, the calculation of each of the three components is described in detail.)

Figure 1

A summary of the datasets used and how values are combined to calculate a household Food Retail Environment Health Quality (FREHQ) measure



FREHQ = Food Retail Environment Health Quality; FoodAPS = National Household Food Acquisition and Purchase Survey.

Source: USDA, Economic Research Service.

Distance From Household to Stores in the Local Food Retail Environment

The first step in calculating the distance between each store in a household's local FRE is to establish which stores are located in the local FRE. There is no clear consensus in the literature on what boundary distance is appropriate for defining the local FRE (Lytle & Sokol, 2017; Caspi et al., 2012). Previous systematic reviews used boundary distances ranging from 100 meters to multiple miles (Charreire et al., 2010). For this 2024 report, the authors included all Circana OmniMarket Core Outlets stores located within a 20-mile straight-line radius of each FoodAPS household for two reasons: First, a 20-mile radius is the farthest distance considered by USDA Food Access measures (Rhone, 2019).[6] Second, only 1.26 percent of FoodAPS households reported a straight-line distance from household residence to primary store farther than 20 miles.[7] This boundary definition provides a consistent boundary for all households, is an inclusive approach, and matches the Circana OmniMarket Core Outlets sample included in the FoodAPS Geographic Component (GC) (Gunderson et al., 2016). Note that these stores are not necessarily stores that FoodAPS households have

[6] Estimates showed that approximately 0.1 percent of the U.S. population is more than 20 miles from the nearest large store (Rhone, 2019).

[7] Note, these households did not have a Circana store within 20 miles, although there may be other stores, as shown in the TDLinx comparison. Therefore, the authors were unable to calculate a FREHQ score for this small subsample, which are likely in remote, rural areas. See appendix A for a TDLinx comparison.

utilized. Understanding store utilization is not needed to create the FREHQ because the set represents stores from which FoodAPS households can choose to shop.

The second step to define each household's local FRE is to calculate the distance between the household and each store within a 20-mile straight-line radius of each FoodAPS household. Recent studies that measured distance between two locations consistently used road network distance because this provides a more realistic representation of the household's travel experience (Havewala, 2021; Athens et al., 2016). However, food access measures used by the USDA, including the Food Access Research Atlas, use straight-line distances. Therefore, the authors calculated the distance using both the straight-line and road network methods because results may differ, particularly in rural areas (Scharadin et al., 2021). The Circana OmniMarket Core Outlets data provided the exact address of each store location in the sample. These addresses were then geocoded to obtain a precise latitude and longitude for each of the store locations. The store street addresses were processed using Environmental Systems Research Institute's (ESRI) StreetMap Premium Geographic Information System package. The shortest route between each household and store location was used when calculating network distances to each store.

Figure 2 provides a description of the distance calculation process. The figure is for illustrative purposes only and does not represent actual observations from either dataset. First, the local FRE is defined by including all Circana stores within 20 miles of the FoodAPS household. The black dashed line represents the 20-mile straight-line radius and is used to form the circle boundary. Store locations within the boundary, colored blue, are included in the household's FREHQ calculation. Store locations farther than 20 miles from the household, colored red, are not included in the household's FREHQ calculation. The road network was used to calculate the shortest distance between the household and store locations. This is represented by the gold road lines between the household and each blue store location.

Figure 2
Representation of distance calculation process

Source: USDA, Economic Research Service.

Store Healthfulness

Measuring the healthfulness of the in-store environment is a consistent challenge because of the type and level of detail of the data needed to make these assessments. In the United States, the Nutrition Environment Measures Survey (NEMS) and Market Basket Assessment Tool (MBAT) are commonly used to measure the healthfulness of products in grocery stores, corner stores, restaurants, vending machines, and the perceived food environment (Glanz et al., 2016; Misyak et al., 2018). The NEMS and MBAT score stores on three dimensions—availability, price, and quality—across a variety of healthy and unhealthy food items ranging from fruits, vegetables, and frozen dinners to items like hot dogs and baked chips. Although these two commonly used in-store assessment tools have been shown to accurately approximate components of the in-store environment and directly observe foods in the stores, they are difficult to implement at the national level because they require labor-intensive store audits.

As an alternative to the in-store assessment tools, the authors calculated the HEI-2020 score using store-level retail sales data. Broadly, the HEI is usually estimated for individuals or households and measures alignment with the Dietary Guidelines for Americans (DGA) (U.S. Department of Agriculture and U.S. Department of Health and Human Services, 2020), with a score of 100 perfectly aligned with the DGA. The HEI-2020 measure is composed of nine adequacy components (Total Fruits, Whole Fruits, Total Vegetables, Greens and Beans, Whole Grains, Dairy, Total Protein Foods, Seafood and Plant Proteins, and Fatty Acids), and four moderation components (Refined Grains, Sodium, Added Sugars, and Saturated Fats). Broadly, the HEI uses a density approach (based on 1,000 kilocalories consumed) that allows the index to assess healthfulness in a wide variety of applications. It has been used extensively to assess the diet quality of the U.S. population and subpopulations using consumption and purchase data (for examples, see Guenther, 2013; Volpe & Okrent, 2012; Rehm et al., 2015). The HEI has also been used to assess the healthfulness of the U.S. food supply (Krebs-Smith et al., 2010; Miller et al., 2015) and healthfulness of food stores (Reedy et al., 2010; Jahns et al., 2016).

In order to proxy the healthfulness of store offerings and calculate a store-level HEI-2020, the authors used USDA, ERS's PP-Suite with the Circana OmniMarket Core Outlets (Carlson et al., 2019; Carlson et al., 2020; Carlson et al., 2022). The Circana OmniMarket Core Outlets contains information on weekly sales data at the UPC level for many stores across the United States. Although it includes some general nutrition information, it misses details necessary to calculate an HEI-2020. The PP-Suite links UPCs in the Circana OmniMarket Core Outlets with the Food and Nutrient Database for Dietary Studies food codes, Food Patterns Equivalents Database codes, and Food Patterns Equivalents Ingredient Database codes (Carlson et al. 2022). This matching process provides nutritional values (e.g., vitamin A, calcium) per 100 grams, and food group equivalents (e.g., cups of fruit, ounces of protein) for each UPC. These values are multiplied by the edible weight of the UPC item to get the nutritional profile of a product in density format and passed through the HEI-2020 formula. (See Carlson et al. (2022) for details on the matching process.) Following the process described in Carlson et al. (2022), the authors calculated the store-level HEI-2020 score for each unique product sold in each store during 2012 and 2013 from the weekly sales data. The list of products is unique by UPC, such that it can be interpreted as a complete list of products sold in each store at least once during the 2 years. Each product appears only once regardless of sales volume because the goal is to more closely proxy which products are offered in the store. Alcott (2019) used a similar approach when investigating the impact of the FRE to reduce endogeneity concerns based on how customers utilized specific store channel types. The authors used data from all of 2012 and 2013 to match the collection period of FoodAPS, between April 2012 and early January 2013, to avoid seasonality issues.

Estimating Household Barrier To Travel

A large body of literature showed that households are more likely to shop at a retail outlet that is closer to their homes (Blut et al., 2018). All else equal, consumers prefer retail locations that are closer rather than farther because of monetary and temporal travel costs (Huff, 1966). Grocery store utilization is no exception. Feather (2003) found that the average household would be willing to pay up to $2 per week more for a supermarket to be 1 mile closer to home. However, the preference for proximity to supermarkets and supercenters is not homogeneous across income, food assistance participation, and other household characteristics (Taylor & Villas-Boas, 2016; Hillier et al., 2017). In particular, SNAP participants and nonparticipants with income between 101 and 185 percent of the poverty threshold were willing to pay about $6 more per week. The greater willingness to pay more for a closer grocery store implies that the barrier to travel is higher for some groups.

Despite these differences, many studies utilizing inverse distance weighting functions (IDWFs) implied that the barrier to travel for all households was the same (Li & Heap, 2011). Operationally, this means that 1 additional mile from the household discounts all households' preference or ability to choose a store in the same way (i.e., the parameter p_i in Eq(1) is the same for all households). However, in reality, the barrier to travel 1 mile more to a food retailer depends on household characteristics such as car access, income, and time constraints (Hamrick & Hopkins, 2012; Ver Ploeg et al., 2015b). Approximately 90 percent of FoodAPS households had access to a car when needed for grocery shopping (Ver Ploeg et al., 2015b). Although this implies that car access is not a major barrier for the average U.S. household, it is still an important aspect to consider for households that do not have access.

Therefore, the authors constructed the FREHQ with two theoretical assumptions about the barrier to travel. First, they followed past research (Glickman et al., 2021) and assume p=1 for all households. This provided a baseline for comparison. Second, they assigned p_i=1 if the household i had access to a car and p_i=2 if the household i did not have access to a car when needed for grocery shopping. These values were supported by Taylor and Villas-Boas (2016), who found that the distance discount to choosing a store was more than double for households without access to a car, regardless of their level of food availability. Operationally, this means that the exposure a household experiences from a store (i.e., the ability to benefit from the store's healthfulness) decreases more rapidly by distance for households that do not have access to a car compared with those that do have access to a car.

Descriptive Statistics for Components of the FREHQ

Multiple novel and complex components are involved in calculating the full FREHQ. Therefore, understanding how each piece quantitatively contributes to the measure allows for a more indepth understanding of its interpretation.

Distance Between FoodAPS Household and Circana OmniMarket Core Outlets Locations

Table 1 presents the household average distance to all Circana OmniMarket Core Outlets stores within each FoodAPS household's local FRE (within 20 miles of the household) and the average of the closest store to the household. These important values present a full picture of the calculation but are too broad to glean any nuanced interpretations. The values represent distances to only the Circana stores in the FoodAPS sample. The mean straight-line distance to all stores is 10.8 miles for all FoodAPS households in the sample, 10.42

miles for households that reside in an urban area, and 11.7 miles for households that reside in a rural area. The actual value of these means is not particularly important because it is largely a function of defining a household's FRE using a 20-mile straight-line radius. However, the relative values follow expected patterns. Each mean network distance is larger than its straight-line counterpart, which is to be expected given that roads often meander. The mean network distance for all households is 13.67 miles. The distance is 13.27 miles for urban households and 14.61 miles for rural households.

Similar trends are present for the average difference to the closest Circana OmniMarket Core Outlets store with the mean network distances larger than their straight-line counterparts and rural distances larger than their urban counterparts. The difference between the mean straight-line and network distance to the closest store is statistically significant only for urban households; however, the difference between the two means is only 0.3 miles. This is because the standard errors for urban households are smaller than rural households. In other words, the variation in the distance to the closest store is more consistent in urban settings, but the variation fluctuates more in rural settings.

Table 1

Descriptive statistics for the average distance to all stores and closest store

	Average distance in miles to all stores in a food retail environment				
Household type	**Mean**	**Standard error**	**95-percent confidence interval**		**Observations**
All straight-line	10.80	0.05	10.71	10.89	4,799
Network	13.67[a]	0.06	13.54	13.79	4,799
Urban straight-line	10.42	0.05	10.32	10.52	4,389
Network	13.27[a]	0.07	13.14	13.41	4,389
Rural straight-line	11.70*	0.11	11.49	11.90	410
Network	14.61*[a]	0.14	14.33	14.89	410
	Distance in miles to closest store in a food retail environment				
Household type	**Mean**	**Standard error**	**95-percent confidence interval**		**Observations**
All straight-line	2.16	0.33	1.48	2.84	4,799
Network	2.82	0.39	2.01	3.64	4,799
Urban straight-line	0.94	0.07	0.80	1.08	4,389
Network	1.29[a]	0.08	1.12	1.47	4,389
Rural straight-line	4.49*	0.81	2.81	6.16	410
Network	5.74*	0.99	3.70	7.77	410

Note: Straight-line is the Euclidean distance between the FoodAPS household member and the nearest or all stores in the food retail environment (FRE). Network is the distance between the FoodAPS household member and the nearest or all stores in the FRE as measured along the network of roads and highways.

[a] Represents a statistical difference from straight-line measure at least at a 95-percent confidence level.

* Represents a statistical difference from urban measure at least at a 95-percent confidence level.

Source: USDA, Economic Research Service, based on National Household Food Acquisition and Purchase Survey (FoodAPS) data on households and Circana OmniMarket Core Outlets (formerly IRI InfoScan) data.

Table 2 presents the average distance to the closest Circana OmniMarket Core Outlets store by store channel type for households that reside in urban and rural areas. On average, drug stores are the closest food retail type for urban households and grocery stores are the farthest. In contrast, dollar stores are on average the closest store type for rural households. This may be a result of the growth of dollar stores in the food retail space, particularly in rural areas (Chenarides et al., 2021). In rural areas, mass merchandisers[8] have the largest

[8] Mass merchandisers are retailers that sell a large number of goods belonging to multiple different categories, including, but not limited to food, clothing, and toys.

mean distance to the closest store of that type at more than 10 miles. In urban areas, the mean distance to the closest mass merchandiser is 3.67 miles, less than the mean distance to the closest grocery store. FoodAPS household members provided information on the store where they do their primary food shopping. The mean values that the authors calculated are in line with the weighted mean road network distance to the stated primary store for urban FoodAPS households (3.97 miles) but are slightly below the mean distance to the stated primary store for rural FoodAPS households (12.31 miles)[9].

Table 2

Descriptive statistics for the closest store by store type for urban and rural households

Average network distance to closest store, by store type, in a food retail environment					
Households in urban areas	**Mean**	**Standard error**	**95-percent confidence interval**		**Observations**
Grocery	4.84	0.09	4.65	5.03	3,716
Drug store	3.15[a]	0.06	3.01	3.27	4,305
Mass merchandiser	3.67[a b]	0.16	3.57	3.77	4,388
Convenience	3.91[a b]	0.07	3.76	4.06	4,051
Dollar store	4.25[a b c d]	0.07	4.11	4.39	3,863
Households in rural areas	**Mean**	**Standard error**	**95-percent confidence interval**		**Observations**
Grocery	9.37	0.42	8.54	10.21	285
Drug store	9.73	0.43	8.88	10.59	290
Mass merchandiser	10.62	1.04	9.82	11.41	369
Convenience	10.48	0.47	9.55	11.42	237
Dollar store	7.91[a b c d]	0.31	7.25	8.52	395

Note: Statistical significance accounts for multiple comparisons with a Bonferroni correction.

[a] Represents a statistical difference from grocery stores at least at a 95-percent confidence level.

[b] Represents a statistical difference from drug stores at least at a 95-percent confidence level.

[c] Represents a statistical difference from mass merchandiser at least at a 95-percent confidence level.

[d] Represents a statistical difference from convenience stores at least at a 95-percent confidence level.

Source: USDA, Economic Research Service, based on National Household Food Acquisition and Purchase Survey (FoodAPS) data on households and Circana OmniMarket Core Outlets (formerly IRI InfoScan) data.

The Circana OmniMarket Core Outlets does not have perfect coverage of all retail stores in FoodAPS-sampled areas. Levin et al. (2019) found that Circana OmniMarket Core Outlets coverage was more sparse than other retail scanner datasets, such as Nielsen's TDLinx, the National Establishment Time Series, County Business Patterns (CBP), and Economic Census. In addition, coverage of stores varied geographically, with a relatively higher coverage in the Eastern United States. Despite these known coverage issues, the Circana OmniMarket Core Outlets is a valuable source of data for analysis involving UPC transaction data. UPC-level information is necessary to estimate store healthfulness values because it allows for detailed nutrient calculations that can be attributed to a specific store location.

The authors used Nielsen TDLinx as a comparison for Circana OmniMarket Core Outlets coverage in FoodAPS geographic areas. Appendix A presents a table showing how Circana OmniMarket Core Outlets coverage in FoodAPS sampled areas compared with Nielsen's TDLinx coverage in FoodAPS sampled areas. In general, the Circana OmniMarket Core Outlets has consistent coverage of FoodAPS sampled areas across the four USDA regions (Midwest, West, Northeast, South), but has better coverage in urban areas compared

[9] The weighted means for network distance to stated primary store, 3.97 miles for urban households and 12.31 miles for rural households, are calculated using the study sample of 4,799 FoodAPS and household replicate weights. Further information on the distance to the stated primary store variable in FoodAPS is available in the User's Guide (2016).

with rural FoodAPS sampled areas.[10] About 31 percent of rural FoodAPS households fell into the lowest coverage quadrant, as opposed to 23 percent of urban households. In terms of store counts, urban households had, on average, 213 Circana OmniMarket Core Outlets store locations within 20 miles, while the average rural household had about 13 locations. Comparatively, urban households had, on average, 1,555 TDLinx stores within 20 miles, and rural households had 271 stores in radius.

The third and fourth quadrants of coverage included 48 percent of all rural households and 51 percent of urban households. Additionally, households with income less than 185 percent of the Federal poverty line (FPL) made up a disproportionate amount of the lowest coverage areas. White and non-Hispanic populations tended to be in better coverage areas than other races and ethnicities. It is important to acknowledge this difference in coverage. A full analysis of how these coverage issues may reflect differences in the FREHQ across these groups would require significant further research.

Store Healthfulness of Circana OmniMarket Core Outlets Retail Locations

The authors estimated store healthfulness for each Circana OmniMarket Core Outlets location within a 20-mile straight-line distance of each FoodAPS household by calculating a store-level HEI-2020 score. Table 3 presents the mean store-level HEI-2020 for all 41,999 stores within FoodAPS household FRE and separately by store type. The average HEI-2020 score for all stores in the sample was 48.15. This is slightly lower than the 2011–2012 U.S. national average HEI-2020 score for individuals (50.9), using food intake data from the National Health and Nutrition Examination Survey (Gu & Tucker, 2017) and the average HEI-2020 score for FoodAPS households (53.0) (Mancino et al., 2018). Although natural comparisons, there are behavioral and technical reasons that the overall average HEI-2020 score for food retailers would not match an overall average for individual or household consumption. Households often shop at multiple stores and therefore optimize their overall basket of goods across various store types. In contrast, store types specialize in a particular market space, meaning their measure of overall healthfulness may be lower. Muth et al. (2016) discussed numerous technical aspects of the Circana OmniMarket Core Outlets, including the unavailability of sample weights for USDA's ERS sample of the Circana OmniMarket Core Outlets and underreporting of random weight perishable products. These aspects combined with using store offerings, rather than sales, may explain lower retail HEI-2020 scores compared with consumption HEI scores.

It is more important for the store-level HEI-2020 scores to follow expected trends by store channel type. The mean HEI-2020 score for grocery stores was 51.59, the second highest of the five store types considered in the sample. The average HEI-2020 score is between the national individual and household consumption averages, which is expected because a large portion of an average household's food at home is purchased at grocery stores. Dollar stores had an average HEI-2020 score of about 41.91 points, which is the lowest average HEI-2020 score of the five store types. This suggests that in 2012 and 2013, the types of foods sold at dollar stores were dissimilar to the types of foods sold at grocery stores. Counterintuitively, convenience stores have the highest average HEI-2020 score at 53.79. This high value is likely a result of the HEI-2020 score being a calorie density measure. Although often utilized for less healthful snack foods, convenience stores often also sell canned beans and vegetables, milk, and eggs. Under the store offerings approach, each of these items is equally weighted. Therefore, as a proportion of calories, convenience stores may look comparably healthy. The difference between this result and intuition may also highlight the premium prices for convenience (easy access) that these types of stores charge. This premium may make healthy foods relatively more expensive. In contrast, candy and chips may be similarly priced with other retail channel types. As a result, convenience stores are not often utilized for more healthful food items.

[10] See table A.1 in the appendix for more coverage information.

Table 3

Mean 2020 Healthy Eating Index (HEI-2020) score for aggregated 2012 and 2013 purchases by store type

Store type	Mean	Standard error	Minimum	Maximum	Store count
Grocery	51.59	0.03	41.29	68.26	7,396
Drug store	46.73[a]	0.02	30.42	74.65	12,687
Mass merchandiser	45.79[a b]	0.03	38.41	57.93	3,324
Convenience	53.79[a b c]	0.04	31.05	62.31	9,817
Dollar store	41.91[a b c d]	0.01	30.66	63.02	8,777
All	48.15	0.02	30.42	74.65	42,001

HEI-2020 = 2020 Healthy Eating Index, calculated at the store level.

Note: Statistical significance accounts for multiple comparisons with a Bonferroni correction.

[a] Represents a statistical difference from grocery stores.

[b] Represents a statistical difference from drug stores.

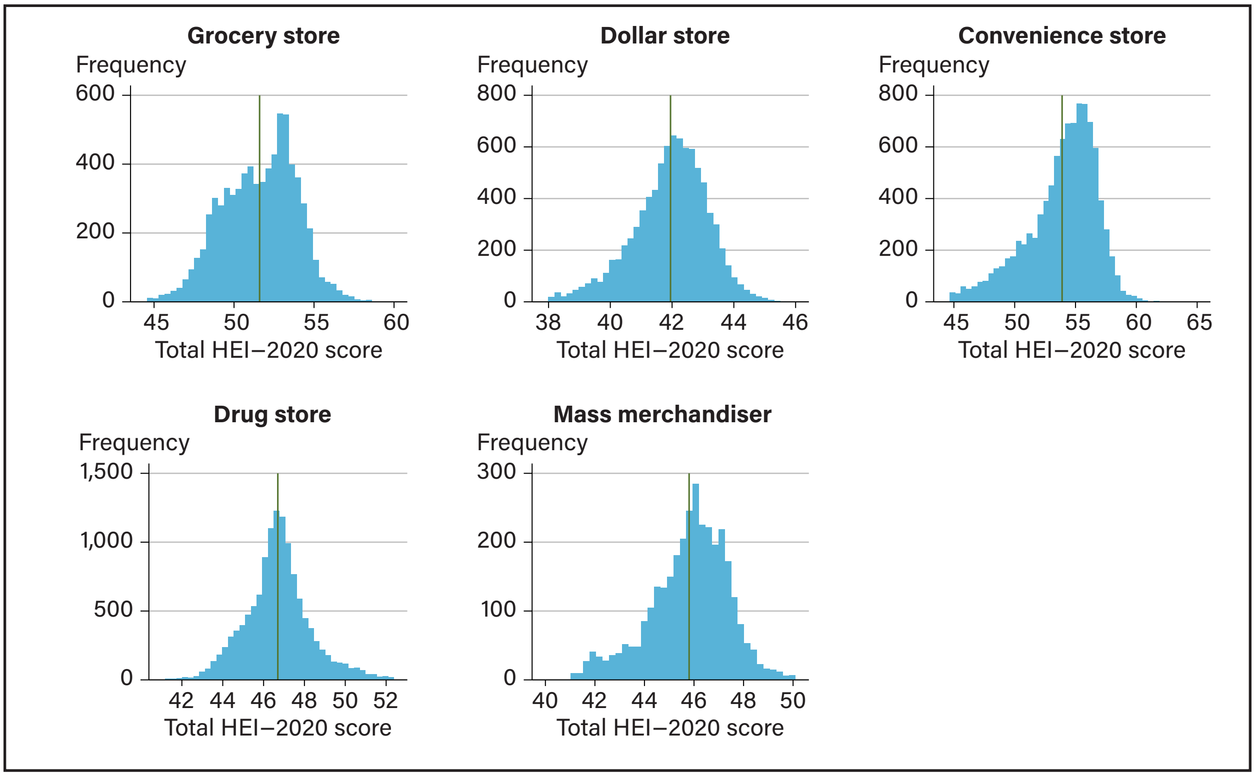
[c] Represents a statistical difference from mass merchandisers.

[d] Represents a statistical difference from convenience stores.

Source: USDA, Economic Research Service, based on data from National Household Food Acquisition and Purchase Survey (FoodAPS) data on households and Circana OmniMarket Core Outlets data.

Considering the shape of the full HEI-2020 score distribution rather than just one point can provide further insights on the value of store-specific measures. Figure 3 provides histograms for the HEI-2020 scores by store channel type. The frequency histograms show that while on average, grocery stores have one of the highest average HEI-2020 scores, they also have a relatively wide distribution. In addition, when considering convenience stores and mass merchandisers, there is a distinct left tail. This suggests that although most of these channel types may have similar mean healthfulness, a significant number of stores have less healthful offerings. The spread in each of these distributions illustrates the value of using a store-specific healthfulness measure, rather than a healthfulness measure that is specific to a store type.

Figure 3
Frequency histograms for store 2020 Healthy Eating Index (HEI-2020) by store type



HEI-2020 = 2020 Healthy Eating Index, calculated at the store level.

Source: USDA, Economic Research Service based on data from National Household Food Acquisition and Purchase Survey (FoodAPS) data on households and Circana OmniMarket Core Outlets data.

Descriptive Statistics and Comparisons for the FREHQ

Each of the components described in the previous section are used to calculate the FREHQ measure for individual households in the FoodAPS dataset. FoodAPS has been used to investigate the impact of numerous topics on food spending and diet quality, including food sales taxes (Dong et al., 2020), school meal programs (Kuhn, 2018), household time constraints (Scharadin, 2022), and the SNAP benefit cycle (Smith et al., 2016). Data from the FoodAPS Geography Component (GC), which collects data on the local food environment in the 50 primary sampling units (PSUs) of the FoodAPS survey, have been used to conduct research on topics related to the local FRE (Taylor & Villas-Boas, 2016; Fan et al., 2018). However, the studies often must rely on homogenous geographic FRE measures. The FREHQ measure was developed to provide current and future FoodAPS users a household-specific measure of the FRE and improve the ability of future studies to consider the impact of the FRE on food spending, food security, and other outcomes, such as diet-related health conditions, with more nuance.

The following sections present descriptive statistics for the estimated FREHQ. Although 4,826 households are in the FoodAPS respondent sample, 26 of the households had no Circana OmniMarket Core Outlets stores within 20 miles of their residence. Therefore, the FREHQ is calculated for 4,799 households. The authors compared differences in household demographics between each quintile of the FREHQ to demographic differences in previously used FRE measures to show how the novel measurement method may differ.

FREHQ Descriptive Statistics: Full Sample and Groups of Policy Interest

Table 4 presents descriptive statistics for the FREHQ measure for all FoodAPS households in the sample and subsamples of particular policy interest (metro/non-metro status, race and ethnicity, and whether family income is at or below 185 percent of the poverty threshold for family size). The rows show four forms of the FREHQ measure to demonstrate the impact of accounting for the road network and household car access on a household's FRE.

The first column presents the average FREHQ measure for all households in the sample. When measuring the distance to store in a straight line, the average FREHQ measure is 40.23 without considering car access and 40.29 when considering car access. When measuring the distance to the store by the road network, the average FREHQ is 40.21 without considering car access and 40.28 when considering car access. The average FREHQ is lower when car access is considered. Overall, little variation exists between the four measures, despite the significant methodological differences. One explanation is that a majority of FoodAPS households have access to a car for food shopping (Ver Ploeg et al., 2015b) and therefore the FREHQ differs only for a minority of households. As a result, the average may not change, but the individual values for households without car access more closely represent their exposure.

The values of the FREHQ can be loosely interpreted like an HEI-2020 score because the measure is a weighted average of stores in a household's local FRE. Given this interpretation, it is clear that the average food environment, regardless of methodology, does not score well using the HEI-2020 rubric with an approximate score of 40 points out of 100. Even with this interpretation, it is difficult to determine what is a "poor" or "good" FREHQ value for an FRE. Therefore, the measure is best used to compare across subgroups to understand if a difference in healthfulness exists between two FREs. Average FREHQ values by subgroups are also presented in table 4. On average, households that live in rural areas have a statistically significant lower FREHQ score compared with households that live in urban areas. Part of this can be attributed to a greater distance between households and food retailers for rural households. However, the composition of food retailers may also play a role. Dollar stores are increasing in number in rural food retail spaces (Chenarides et al., 2021) but also have the lowest HEI-2020 score of all channel types included in the sample.

Differences exist in the average FREHQ by income, race, and ethnicity. Households below 185 percent of the poverty threshold had a lower average FREHQ. Although the difference is somewhat small, about 0.4 points, it is statistically significant. This highlights previous studies that have found that households with more rigid income constraints, i.e. lower incomes, may also have lower access to healthy foods. Both barriers can contribute to lower diet quality outcomes because it costs lower income households more to access healthy food, whether through travel costs or time costs, but the households are less able to absorb that additional cost. Differences also exist in the FREHQ across race and ethnicity groups. Asian and Hispanic households have the highest FREHQ score. Many other factors contribute to household consumption, but this follows a pattern in HEI-2020 scores, as Asian and Hispanic households also have higher diet quality (Tao et al., 2022). Despite this relatively greater access to foods as measured in relation to the DGA, Asian and Hispanic households may still have difficulty finding culturally appropriate food in the local FRE. No statistical difference exists in the average FREHQ between White and Black households. Both kinds of households have an average FREHQ score of approximately 40. The lack of statistical difference between these values may suggest the importance of considering other determinants of diet quality and health outcomes, such as income and time constraints, to explain differences between groups, in addition to the FRE.

Table 4
Mean of Food Retail Environment Healthfulness Quality (FREHQ) for full FoodAPS sample of households and subgroups of policy interest

FREHQ type	Full sample	Metro	Nonmetro	White	Black	Asian	Hispanic	Above 185-percent poverty threshold	Below 185-percent poverty threshold
Straight-line simple	40.23	40.55	38.18[a]	40.04	40.39	41.94[b c]	41.32[b c]	40.55	40.17[e]
	(0.48)	(0.06)	(0.36)	(0.08)	(0.14)	(0.28)	(0.14)	(0.10)	(0.10)
Straight-line car	40.21	40.56	38.13[a]	40.03	40.34	42.26[b c]	41.35[b c]	40.56	40.17[e]
	(0.49)	(0.07)	(0.38)	(0.09)	(0.15)	(0.27)	(0.14)	(0.09)	(0.10)
Network-simple	40.29	40.59	38.36[a]	40.09	40.34	42.07[b c]	41.34[b c]	40.60	40.22[e]
	(0.49)	(0.07)	(0.36)	(0.09)	(0.13)	(0.27)	(0.14)	(0.10)	(0.10)
Network-car	40.28	40.54	38.31[a]	40.10	40.40	42.44[b c]	41.37[b c d]	40.61	40.24[e]
	(0.50)	(0.07)	(0.37)	(0.10)	(0.14)	(0.26)	(0.15)	(0.09)	(0.10)
Observations	4,799	4,389	410	3,345	704	195	921	2,281	2,518

FREHQ = Food Retail Environment Healthfulness Quality; FoodAPS = National Household Food Acquisition and Purchase Survey.

Note: Straight-line simple is calculated as Euclidean distance without considering whether the household has access to a vehicle. Straight-line car is calculated using Euclidean distance but separately considers whether the household has a vehicle. Network simple is calculated using road network distance but does not consider whether the household has a vehicle. Network car is calculate using the road network and separately considers whether the household has a vehicle. When applicable, statistical significance accounts for multiple comparisons with a Bonferroni correction.

[a] Represents a statistical difference between households living in a nonmetropolitan area and households living in a metropolitan area at least at the 95-percent level.

[b] Represents a statistical difference between White and other households at least at the 95-percent level.

[c] Represents a statistical difference between Black and other households at least at the 95-percent level.

[d] Represents a statistical difference between Asian and other households at least at the 95-percent level.

[e] Represents a statistical difference between households with incomes above and below 185 percent of the poverty threshold for family size at least at the 95-percent level.

Source: USDA, Economic Research Service based on data from National Household Food Acquisition and Purchase Survey (FoodAPS) data on households and Circana OmniMarket Core Outlets data.

Comparisons of the FREHQ to Other County-Level Measures of the Food Retail Environment Overall, and by Household Characteristics

It is important to compare the new FREHQ measures with the FRE measures used in past research to understand how they are similar and how they differ. A direct comparison between the FREHQ and existing measures of food retail access is difficult because the FREHQ is measured on the household level and common food access measures are usually measured at a geographic level (e.g., census tract or county). The authors first divided all counties into quintiles based on the percentage of households in a county that were considered low-income and had low access to food retail based on data from the USDA's Food Environment Atlas (FEA). Individuals are considered to be in the low-income category if their annual family income is at or below 200 percent of the Federal poverty threshold for family size. Households are considered to be in the low-access category if they live farther than 1 mile from the nearest supermarket, supercenter, or large grocery store for an urban area or farther than 10 miles for a rural area.

Once the quintiles were constructed, the authors used them to compare county-level household characteristics of those quintiles to the household characteristics of FoodAPS households based on their quintile of the FREHQ measure. They compared the quintiles from the two FRE measures by race/ethnicity, poverty rates, obesity rates, and food insecurity rates. The first three of these four county-level characteristics came from the FEA, and the county-level estimates of food security are model-based estimates from Li et al. (2023). The FEA data are widely cited in FRE literature (Cooksey-Stowers et al., 2017; Ahern et al., 2011; Chi et al., 2013) and provided numerous county-level characteristics for years similar to the FoodAPS collection period. The three indicators from the FEA are: (1) proportion of households that identify as non-Hispanic White, non-Hispanic Black, and Hispanic; (2) the county poverty rate; and (3) the county obesity rate. However, the FEA provides food insecurity rates only at the State level. Therefore, the authors used model-based small area estimates calculated by Li et al. (2023). The estimates were calculated by linking the FoodAPS and the American Community Survey from 2010 and 2012. Following a weighted finite population Bayesian bootstrap method, Li et al. (2023) were able to estimate food insecurity rates more efficiently at the county level. Appendix B contains a short summary of the methods used to make the county-level food security estimates. For each of the four county-level indicators, the authors ranked counties by each respective indicator's value and then divided the counties into quintiles. Each of these proportions is weighted by the population in each county in a quintile according to the methodology described in appendix B. The authors then compared each quintile with the quintiles of the FREHQ using households in the FoodAPS sample and their characteristics. Following a similar process, they calculated the proportion of households that identify as non-Hispanic White, non-Hispanic Black, and Hispanic in each quintile using the race of the primary respondent. They also calculated the proportion of FoodAPS households with household annual family income at or below 100 percent of the Federal poverty threshold for family size; low or very low food security based on USDA's standard, 10-item U.S. Adult Food Security Survey Module to assess household food security status in the last 30 days; and the proportion of households with an average household body mass index (BMI) above 30, based on the average of the BMI of all respondents in the household.

Finally, a higher FREHQ is associated with better access to healthy food, while a higher county-level, low-income, low-access rate is associated with lower access to food. Therefore, the authors inverted the county-level measure from the FEA to make interpreting the quintile comparisons more intuitive. As a result, a higher quintile is associated with better FRE in both cases.

Comparing Race and Ethnicity Across Food Access Quintiles

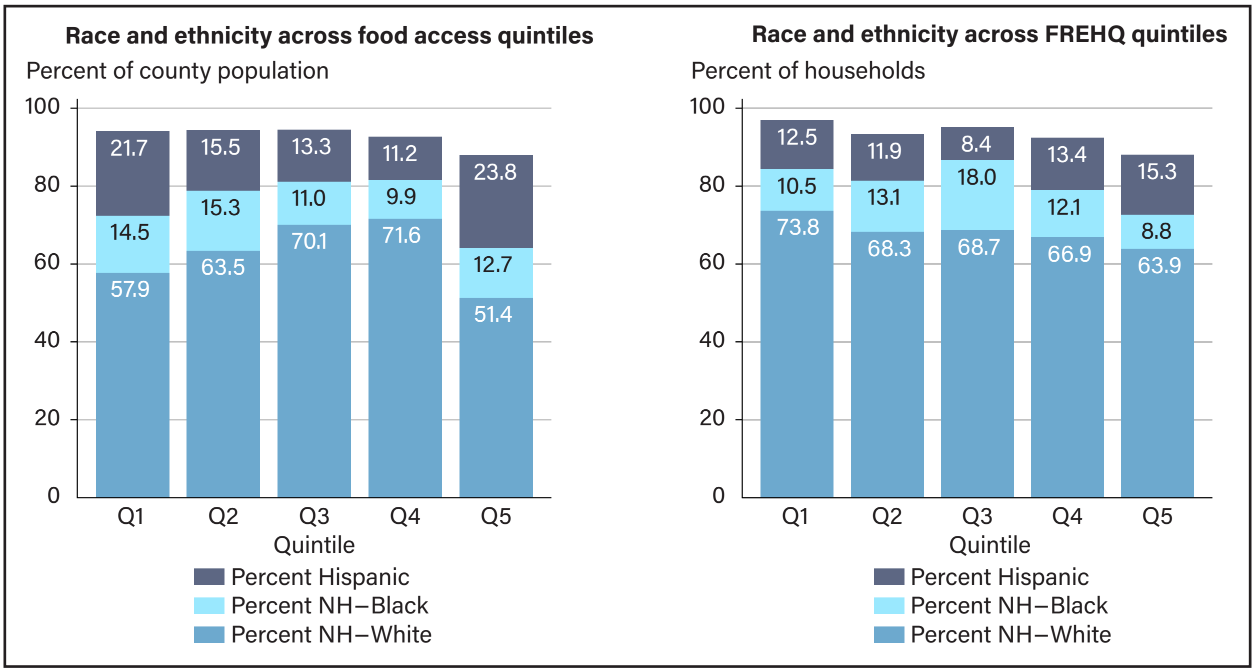
Figure 4 presents two stacked bar graphs. The average proportion of individuals in a county who identify as non-Hispanic White, non-Hispanic Black, and Hispanic for each quintile of the FEA county-level, low-income, and low-access measure are on the left. The proportion of households with a primary respondent who identifies as non-Hispanic White, non-Hispanic Black, and Hispanic in each quintile of the FREHQ are on the right. The FEA chart follows patterns found in previous research. The proportion of individuals identifying as White ranged from 57.9 percent in the first quintile, which is comprised of counties with the highest proportion of individuals who are low income and low access, to 71.6 percent in the fourth quintile, which is comprised of counties with the lowest proportion of individuals who are low income and low access. The proportion of non-White individuals decreased between the first and fourth quintiles. This overall trend matches results that used a geographic-based measure of the FRE (Walker et al., 2010). The fifth quintile does not follow the pattern and has a much larger proportion of non-White individuals. Previous work found that income is negatively correlated with population density, while the distance to a store is positively correlated with population density (Wilde et al., 2014). Together, geographic-based measures like the low-income, low-access FEA measure show a somewhat counterintuitive result that Black and Hispanic populations have better access to stores (Wilde et al., 2014).

The FoodAPS chart presenting the race and ethnicity of the primary respondent does not follow a similar clear pattern. All FREHQ proportions were calculated using FoodAPS household weights. The proportion of White households generally decreased from the first (73.8 percent) to the fifth quintile (63.9 percent). Interestingly, the proportion of non-Hispanic Black households in a quintile increased from the first quintile (10.5 percent) to the third quintile (18.0 percent), and then quickly decreased through the fifth quintile (8.8 percent). In contrast, the proportion of Hispanic households decreased from the first quintile (12.5 percent) to the third quintile (8.4 percent), and then increased through the fifth quintile (15.3 percent).

The difference between these two charts suggests that individual heterogeneity is lost when considering geographic boundary measures such as food store access on a county level. Race and ethnicity trends across the distribution of a county-level access measures are linear and pronounced. In contrast, the trends across the distribution of the household-level measure are less pronounced for White households and nonlinear for non-Hispanic Black and Hispanic households. As a result, the connection between the FRE and individual- or household-specific decisions, like food purchasing behaviors, may seem weaker when an empirical study uses county-level measures of the FRE.

Figure 4

Comparing race/ethnicity of quintiles of low-income/low-access percentages from the Food Environment Atlas (FEA) and quintiles of the Food Retail Environment Healthfulness Quality (FREHQ)



FEA = Food Environment Atlas; FREHQ = Food Retail Environment Healthfulness Quality; NH = non-Hispanic.

Note: The stacked bars in figure 4 do not sum to 100 because only non-Hispanic White, non-Hispanic Black, and Hispanic groups are included. Figure 4 does not include individuals identifying as Asian American and Pacific Islanders, Native American, Middle Eastern, North African, or multiple races or ethnicities because these groups comprised a small portion of the respondent sample.

Source: USDA, Economic Research Service based on data from National Household Food Acquisition and Purchase Survey (FoodAPS) data on households and Circana OmniMarket Core Outlets data.

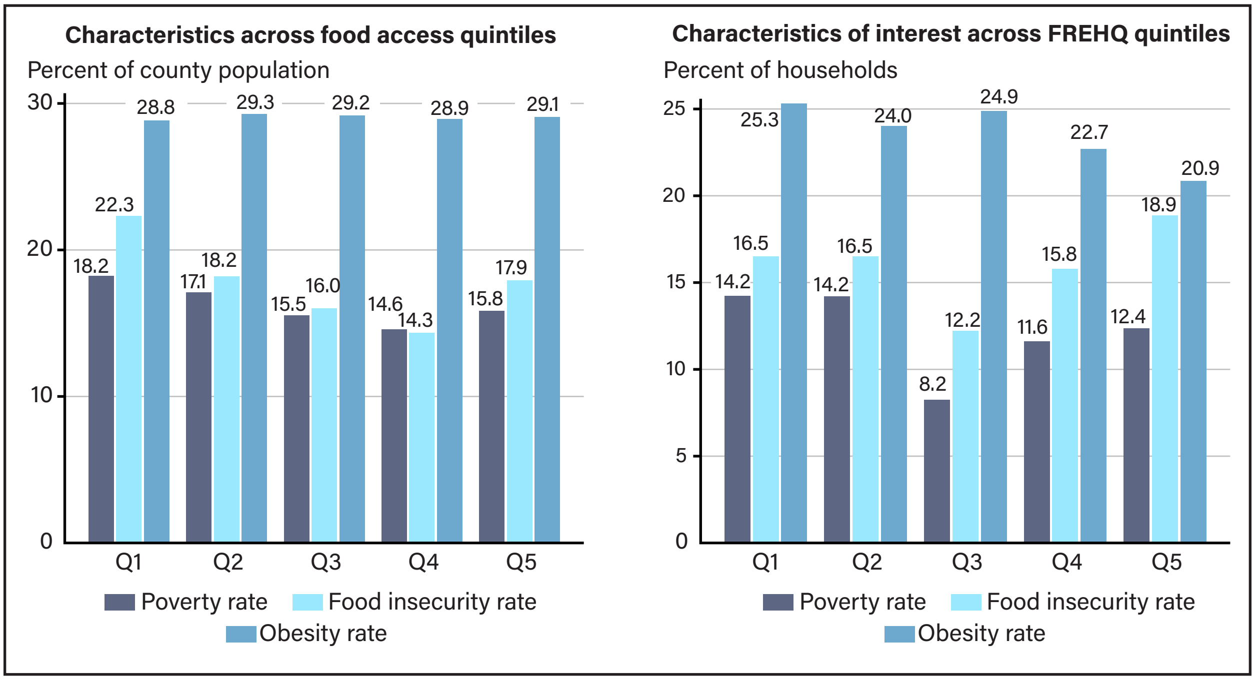
Comparing Poverty, Food Insecurity, and Obesity Across FREHQ Quintiles

Figure 5 compares the county-level poverty rate, food insecurity rate, and obesity rate—three characteristics commonly related to the FRE—across food store access quintiles from the FEA and the FREHQ quintiles. The FEA provides county-level poverty and obesity rates. However, food insecurity rates are only available at the State level. Therefore, the authors used model-based small area estimates calculated by Li et al. (2023). The county-level poverty rate and food insecurity follow expected patterns from the first to the fourth quintile. The first quintile, comprised of counties with the lowest healthy food store access, had an average poverty rate of 18.2 percent and an average food insecurity rate of 22.3 percent. These rates decrease to 14.6 percent and 14.0 percent, respectively, as food store access increased in the fourth quintile. These rates increased slightly in the fifth quintile. Similar to above, this deviation was driven by the high correlation between population density and the distance measure of low store access (Wilde et al., 2014). In contrast, very little change exists in the obesity rate by healthy food store access quintile. The second quintile has the highest average obesity rate, 29.3 percent, and the first quintile has the lowest average obesity rate. This reinforces common knowledge that obesity is a widespread issue.

The bar graphs presenting the FREHQ quintiles provide more nuance because the household-level measure allows for more variability. Rates of poverty and food insecurity are higher in the first and second quintile, 14.2 percent and 16.5 percent, respectively, compared with quintiles with more nutritious food environments. However, the increasing trends across quintiles three, four, and five are opposite of the trends for the county-level measure, with rates of poverty and food insecurity increasing as the healthfulness quality of the FRE increases. In general, the percent of households with an average household BMI above 30 decreases as the FREHQ increases. While the first two quintiles reinforce past research highlighting high rates of poverty, food insecurity, and obesity in areas with lower access to healthy food, the last three quintiles suggest that other household constraints may dominate as access to healthy food increases.

Future analysis can further expand on the trends, but one explanation is that households with relatively easy access to stores selling healthy food (quintiles four and five), may not have the income to support consistently purchasing healthy foods or the time to consistently prepare healthy meals. FRE literature often focuses on the impact on the average household, which combines the impact of the FRE for those constrained by their FRE (households in the first two quintiles), and those households that may not be constrained (households in the last three quintiles). As a result, analysis focusing on the average household may miss the disproportionate impact on households in the lowest quality FRE. The higher prevalence of poverty, food insecurity, and obesity for households in the first quintile illustrates how focusing FRE analysis on individuals and households in the poorest local FRE can provide more nuance in understanding the influence of the FRE.

Figure 5
Comparing characteristics of the Food Environment Atlas (FEA) and Food Retail Environment Healthfulness Quality (FREHQ)



Note: The percent of households in each Food Environment Atlas (FEA) quintile are calculated using data from the FEA, with poverty defined as the county poverty rate, obesity defined as the county obesity rate, and food insecurity defined as the model-based small area estimates calculated by Li et al. (2023) for county-level food insecurity rates. The percent of households in each Food Retail Environment Healthfulness Quality (FREHQ) quintile are calculated using FoodAPS data, with poverty defined as household annual family income at or below 100 percent of the Federal poverty threshold for family size; food insecurity defined as low or very low food security based on USDA's standard 10-item, U.S. Adult Food Security Survey Module; and obesity defined as an average household body mass index (BMI) above 30 based on the average of the BMI of all respondents in the household.

Source: USDA, Economic Research Service based on data from National Household Food Acquisition and Purchase Survey (FoodAPS) data on households and Circana OmniMarket Core Outlets data.

Poverty and food insecurity rates increase as the healthfulness quality of the FRE increases from the second to fifth quintile. This suggests that households face constraints beyond the local FRE. Although the FRE is still a factor in household diet quality and food choice, other constraints such as income or time may be more binding for these households.

Conclusion, Applications, and Extensions

The authors developed the household-level FREHQ to address limitations in current homogenous geography-based FRE measures. The new FREHQ measure was calculated for each household in FoodAPS based on a store-level HEI-2020 score for each store within 20 miles of the household, discounting the weight of a store by its distance from each household. The store-level HEI-2020 score was based on the Circana OmniMarket Core Outlets for store location and UPC-level weekly sales data. The FREHQ improves upon previous FRE measures in four key ways: It incorporates exposure to all food retailers in an environment, is household-specific, uses store-specific information to create a healthfulness measure for individual stores, and is comparable across a variety of geographies. The FREHQ is calculated for FoodAPS households, and despite the age of the data used for this study, the process can be repeated to calculate the FREHQ measure for future FoodAPS iterations as more updated data become available. This study also provides a new way to proxy store inventory data using Circana OmniMarket Core Outlets scanner sales data since food retail inventory data are not available.

Comparing the FREHQ with a commonly used FRE access measure, the county-level estimate of the number of low-income and low-access people, shows that the FREHQ retains more household heterogeneity than the county-level measure. The FEA county-level measure is largely driven by population density because it considers the distance to the closest supermarket or grocery store. In addition to considering only one store location, the measure does not consider the healthfulness of the in-store environment. In contrast, the FREHQ distribution shows that households face different constraints to making healthy diet decisions depending on the quality of their FRE. For example, the proportion of households under the poverty line in the first and second quintile of the FREHQ, those with the least healthful FRE, is two times higher than the proportion of households under the poverty line in the third quintile. Previous literature using homogeneous geographic-based FRE measures focused on the average effect of the FRE and do not account for these nuances across the FRE distribution. One reason that the estimated impact of the FRE on household diet quality and food choice varies across studies is that other household constraints are more binding than the FRE when healthful food is available. Because the FREHQ is calculated on the household level, researchers can use it to focus on how the FRE influences food choices for households with the least access to healthful food.

Because the FREHQ is a household-level measure, it can increase the number of econometric techniques that can be used to investigate the relationship between diet quality and health outcomes and the local FRE. Specifically for FoodAPS users where sampled households are clustered in secondary sampling units of block groups or groups of block groups, the household-level FREHQ measure allows for the use of geographic-fixed effects to control for the local infrastructure, such as walkability of an area or access to public transportation. This is more difficult for homogenous geographic-based measures of the FRE because there is not variation in the geographic unit. Some research attempts to address this limitation by including fixed effects on a higher level (e.g., State-fixed effect with county-level FRE measure); however, a high-level fixed effect is a rough proxy. In contrast, the FREHQ has ample variation within geographies where the FoodAPS sample was collected, allowing for relatively granular fixed effects.

Besides being able to update the FREHQ as new data become available, future research may also consider how to build upon the foundation of the FREHQ to incorporate additional ways that individuals and households interact with their FRE. One future component could consider grocery delivery services. Participation in such services has been increasing, with a particular increase during the height of the COVID-19 pandemic (Jones, 2021). Federal food assistance programs, such as SNAP and the USDA Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), have also recognized the ability for these services to increase healthy food accessibility (USDA, Food and Nutrition Service (FNS), 2022). A second adaptation could consider time to travel to a store rather than distance. Although these two metrics are highly correlated, previous research has shown that the importance of travel time to a location, rather than distance to a location, can vary depending on the locality (Scharadin et al., 2021). Considering time costs could be particularly useful for household members without vehicles who may rely on public transportation or rides from friends or family members.

One final adaptation could consider barriers to accessing a store beyond physical distance. Recent calls to account for how individuals and households move through and are influenced by their built environments require incorporating household attitudes, such as how comfortable a household is visiting a particular store. A Black or Hispanic household's probability of exposure to discrimination when making food purchases increases as the distance and the dissimilarity between their race or ethnicity of household members and the predominant race or ethnicity of the store's neighborhood increases (Zenk et al., 2014; Lee, 2000). This may be one reason that White households are more likely to travel farther than non-White households for grocery purchases (Hillier et al., 2017). Incorporating these components into a FRE measure could be used to examine the potential impact of exposure to discrimination in food purchasing behaviors.

References

Ahern, M., Brown, C., & Dukas, S. (2011). A national study of the association between food environments and county-level health outcomes. *The Journal of Rural Health,* 27(4), 367–379.

Allcott, H., Diamond, R., Dubé, J. P., Handbury, J., Rahkovsky, I., & Schnell, M. (2019). Food deserts and the causes of nutritional inequality. *The Quarterly Journal of Economics,* 134(4), 1,793–1,844.

Athens, J. K., Duncan, D. T., & Elbel, B. (2016). Proximity to fast-food outlets and supermarkets as predictors of fast-food dining frequency. *Journal of the Academy of Nutrition and Dietetics,* 116(8), 1,266–1,275.

Black, C., Ntani, G., Inskip, H., Cooper, C., Cummins, S., Moon, G., & Baird, J. (2014). Measuring the healthfulness of food retail stores: variations by store type and neighbourhood deprivation. *International Journal of Behavioral Nutrition and Physical Activity,* 11(1), 69.

Blanco, I., Diego, I., Bueno, P., Fernández, E., Casas-Maldonado, F., Esquinas, C., Soriano, J. B., & Miravitlles, M. (2017, December). Geographical distribution of COPD prevalence in Europe, estimated by an inverse distance weighting interpolation technique. *International Journal of Chronic Obstructive Pulmonary Disease*, 13, 57–67.

Botkins, E. R., & Roe, B. E. (2018, January). Understanding participation in farm to school programs: Results integrating school and supply-side factors. *Food Policy,* 74, 126–137.

Carlson, A. C., Tornow, C. E., Page E. T., McFadden, A. B., & Zimmerman, T. P. (2022). Development of the Purchase to Plate Crosswalk and Price Tool: Estimating prices for the National Health And Nutrition Examination Survey (NHANES) foods and measuring the healthfulness of retail food purchases. *Journal of Food Composition and Analysis,* 106, 104, 344.

Carlson, A., Lino, M., Juan, W. Y., Hanson, K., & Basiotis, P. P. (2007). Thrifty Food Plan, 2006. (2007). *AgEcon* Search.

Carlson, A., Page, E. T., Zimmerman, T. P., Tornow, C. E., & Hermansen, S. (2019). *Linking USDA nutrition databases to IRI household-based and store-based scanner data* (Report No. TB-1952). U.S. Department of Agriculture, Economic Research Service.

Carlson, A., Kuczynski, K., Pannucci, T. R., Koegel, K., Page, E. T., Tornow, C. E., & Zimmerman, T. P. (2020). *Estimating prices for foods in the National Health and Nutrition Examination Survey: The Purchase to Plate Price Tool* (Report No. TB-1955). U.S. Department of Agriculture, Economic Research Service.

Caspi, C. E., Sorensen, G., Subramanian, S. V., & Kawachi, I. (2012). The local food environment and diet: A systematic review. *Health & Place,* 18(5), 1,172–1,187.

Charreire, H., Casey, R., Salze, P., Simon, C., Chaix, B., Banos, A., Badariotti, D., Weber, C., & Oppert, J. M. (2010). Measuring the food environment using geographical information systems: A methodological review. *Public Health Nutrition,* 13(11), 1,773–1,785.

Chenarides, L., Cho, C., Nayga, R. M., & Thomsen, M. R.. (2021, September). Dollar stores and food deserts. *Applied Geography,* 134, 102,497.

Chi, S. H., Grigsby-Toussaint, D. S., Bradford, N., & Choi, J. (2013, October). Can geographically weighted regression improve our contextual understanding of obesity in the U.S.? Findings from the USDA Food Atlas. *Applied Geography,* 44, 134–142.

Christensen, G., & Bronchetti, E. T. (2020, August). Local food prices and the purchasing power of SNAP benefits. *Food Policy,* 95, 101,937.

Clay, M., Ver Ploeg, M., Coleman-Jensen, A., Elitzak, H,, Gregory, C., Levin, D., Newman, C., & Rabbitt, M. (2016). *Comparing National Household Food Acquisition and Purchase Survey (FoodAPS) data with other national food surveys' data* (Report No. EIB-157). U.S. Department of Agriculture, Economic Research Service.

Clary, C. M., Ramos, Y., Shareck, M., & Kestens, Y. (2015, February). Should we use absolute or relative measures when assessing foodscape exposure in relation to fruit and vegetable intake? Evidence from a wide-scale Canadian study. *Preventive Medicine*, 71, 83–87.

Clary, C., Matthews, S. M., & Kestens, Y. (2017, March). Between exposure, access and use: Reconsidering foodscape influences on dietary behaviours. *Health & Place,* 44, 1–7.

Cooksey-Stowers, K., Schwartz, M. B., & Brownell, K. D. (2017). Food swamps predict obesity rates better than food deserts in the United States. *International Journal of Environmental Research and Public Health,* 14(11), 1,366.

Cummins, S., Clary, C., & Shareck, M. (2017). Enduring challenges in estimating the effect of the food environment on obesity. *The American Journal of Clinical Nutrition*, 106(2), 445–446.

Cummins, S., Flint, E., & Matthews, S. A. (2014). "New neighborhood grocery store increased awareness of food access but did not alter dietary habits or obesity." *Health Affairs* 33(2), 283–291.

Davis, B., & Carpenter, C. (2009). Proximity of fast-food restaurants to schools and adolescent obesity. *American Journal of Public Health,* 99(3), 505–510.

Dong, D., Yuqing, Y., & Stewart, H. (2020). The effects of food sales taxes on household food spending: An application of a censored cluster model. *Agricultural Economics,* 51(5), 669–684.

Dorfman, J. H., Gregory, C., Liu, Z., & Huo, R. (2019). Re-examining the SNAP benefit cycle allowing for heterogeneity. *Applied Economic Perspectives and Policy*, 41(3), 404–433.

Dubowitz, T., Ghosh-Dastidar, M., Cohen, D. A., Beckman, R., Steiner, E. D., Hunter, G. P., Flórez, K. R.. Huang, C., Vaughan, C. A., Sloan, J. S., Zenk, S. A., Cummins, S., & Collins, R. L. (2015). Diet and perceptions change with supermarket introduction in a food desert, but not because of supermarket use. *Health Affairs,* 34(11), 1,858–1,868.

Fan, L., Baylis, K., Gundersen, C., & Ver Ploeg, M. (2018). Does a nutritious diet cost more in food deserts? *Agricultural Economics* 49(5), 587–97.

Feather, P. M. 2003. "Valuing Food Store Access: Policy Implications for the Food Stamp Program." Valuing food store access: Policy implications for the food stamp program. *American Journal of Agricultural Economics,* 85(1), 162–172.

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

Giles-Corti, B., Broomhall, M. H., Knuiman, M., Collins, C., Douglas, K., Ng, K., Lange, A., & Donovan R. J. (2005). Increasing walking: How important is distance to, attractiveness, and size of public open space? *American Journal of Preventive Medicine*, Active Living Research, 28 (2, Supplement 2), 169–176.

Glanz, K., Johnson, L., Yaroch, A. L., Phillips, M., Ayala, G. X., & Davis, E. L. (2016). Measures of retail food store environments and sales: Review and implications for healthy eating initiatives. *Journal of Nutrition Education and Behavior*, 48(4), 280-288.e1.

Glickman, A. R., Clark, J. K., & Freedman, D. A. (2021, May). A relational approach to evaluate food environments finds that the proximate food environment matters for those who use it. *Health & Place,* 69, 102,564.

Groff, E. R. (2014). Quantifying the exposure of street segments to drinking places nearby. *Journal of Quantitative Criminology*, 30(3), 527–548.

Gu, X., & Tucker, K. L. (2017). Dietary quality of the U.S. child and adolescent population: Trends from 1999 to 2012 and associations with the use of Federal nutrition assistance programs. *The American Journal of Clinical Nutrition,* 105(1), 194–202.

Gunderson, C., Baylis, K., Fan, L., House, L., & Dutko, P. (2016). *Construction of weekly store-level food basket costs.* U.S. Department of Agriculture, Economic Research Service documentation.

Guenther, P. M., Kellie O., Casavale, K. O., Reedy, J., Kirkpatrick, S. I., Hiza, H. A. B., Kuczynski, K. J., Kahle, L. L., & Krebs-Smith, S. M. (2013). Update of the Healthy Eating Index: HEI-2010. *Journal of the Academy of Nutrition and Dietetics,* 113(4), 569–580.

Hamrick, K. S., & Hopkins, D. (2012). The time cost of access to food–distance to the grocery store as measured in minutes. *International Journal of Time Use Research*, 9(1), 28–58.

Havewala, F. (2021). The dynamics between the food environment and residential segregation: An analysis of metropolitan areas. *Food Policy*, Urban food policies for a sustainable and just future, 103, 102,015.

Hedrick, V. E., Farris, A. R., Houghtaling, B., Mann, G., & Misyak, S. A. (2022). Validity of a market basket assessment tool for use in Supplemental Nutrition Assistance Program education healthy retail initiatives. *Journal of Nutrition Education and Behavior*, 54(8), 776–783.

Hillier, A., Smith, T. W., Whiteman, E. D., & Chrisinger, B. W. (2017). Discrete choice model of food store trips using National Household Food Acquisition and Purchase Survey (FoodAPS). *International Journal of Environmental Research and Public Health,* 14(10), 1,133.

Holzman, D. C. (2010). Diet and Nutrition: White House proposes healthy food financing initiative. *Environmental Health Perspectives,* 118(4), A156.

Hu, R., Dong, S., Zhao, Y., Hu, H., & Li, Z. (2013). Assessing potential spatial accessibility of health services in rural China: A case study of Donghai Count. *International Journal for Equity in Health*, 12(1), 35.

Huff, D. L. (1966). A programmed solution for approximating an optimum retail locations. *Land Economics,* 42(3), 293–303.

Jahns, L., Scheett, A. J., Johnson, L. K., Krebs-Smith, S. M., Payne, C. R., Whigham, L. D., Hoverson, B. S., & Kranz, S. (2016). Diet quality of items advertised in supermarket sales circulars compared to diets of the U.S. population, as assessed by the Healthy Eating Index–2010. *Journal of the Academy of Nutrition and Dietetics*, 116(1), 115-122.e1.

Jones, J. W. (2021). *COVID-19 working paper: Supplemental Nutrition Assistance Program and Pandemic Electronic Benefit Transfer redemptions during the Coronavirus pandemic* (Report No. AP-089). U.S. Department of Agriculture, Economic Research Service.

Krebs-Smith, S. M., Reedy, J., & Bosire, C. 2010. Healthfulness of the U.S. food supply: Little improvement despite decades of dietary guidance. *American Journal of Preventive Medicine*, 38(5), 472–477.

Kuhn, M. A. (2018, October). Who feels the calorie crunch and when? The impact of school meals on cyclical food insecurity. *Journal of Public Economics,* 166, 27–38.

Kwan, M. P. (2018). The limits of the neighborhood effect: Contextual uncertainties in geographic, environmental health, and social science research. *Annals of the American Association of Geographers,* 108(6), 1,482–1,490.

Larson, N., & Story, M. (2009). A review of environmental influences on food choices. *Annals of Behavioral Medicine,* 38 (S1), 56–73.

Lee, J. (2000). The salience of race in everyday life: Black customers' shopping experiences in Black and White neighborhoods. *Work and Occupations,* 27(3), 353–376.

Li, J., & Heap, A. D. (2011). A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecological Informatics*, 6(3), 228–241.

Li, K., Si, Y., West, B. T., Zhang, X., Page, E. T., Gonzalez, J. M., & Miller, L., (2023). Estimating county-level food insecurity prevalence: A comparison of model-based procedures. Technical report, University of Michigan, Ann Arbor.

Liu, Y., Pan, J., Zhang, H., Shi, C., Li, G., Peng, Z., Ma, J., Zhou, Y., & Zhang, L. (2019). Short-term exposure to ambient air pollution and asthma mortality. *American Journal of Respiratory and Critical Care Medicine,* 200(1), 24–32.

Lytle, L. A., & Sokol, R. L. (2017, March). Measures of the food environment: A systematic review of the field, 2007–2015. *Health & Place,* 44, 18–34.

Mancino, L., Todd, J. E., & Scharadin, B. (2018). USDA's National Household Food Acquisition and Purchase Survey: Methodology for imputing missing quantities to calculate Healthy Eating Index—2010 scores and sort foods into ERS food groups (Report No. TB-1947). U.S. Department of Agriculture, Economic Research Service..

Miller, P. E., Reedy, J., Kirkpatrick, S. I., & Krebs-Smith, S. M. 2015. The United States food supply is not consistent with dietary guidance: Evidence from an evaluation using the Healthy Eating Index–2010. *Journal of the Academy of Nutrition and Dietetics,* 115(1), 95–100.

Misyak, S. A., Hedrick, V. E., Pudney, E., Serrano, E. L., & Farris, A. R. (2018). Reliability of a Market Basket Assessment Tool (MBAT) for use in SNAP-Ed healthy retail initiatives. *Journal of Nutrition Education and Behavior,* 50(5), 511–515.

Muth, M. K., Sweitzer, M., Brown, D., Capogrossi, K., Karns, S. A., Levin, D., Okrent, A., Siegel, P., & Zhen, C., eds. (2016). *Understanding IRI household-based and store-based scanner data* (TB-1942). U.S. Department of Agriculture, Economic Research Service.

National Household Food Acquisition and Purchase Survey (FoodAPS): User's guide to survey design, data collection, and overview of datasets. (2016). U.S. Department of Agriculture, Economic Research Service.

Perchoux, C., Chaix, B., Cummins, S., & Kestens, Y. (2013, May). Conceptualization and measurement of environmental exposure in epidemiology: Accounting for activity space related to daily mobility. *Health & Place,* 21, 86–93.

Reedy, J., Krebs-Smith, S. M., & Bosire, C. (2010). Evaluating the food environment: Application of the Healthy Eating Index–2005. *American Journal of Preventive Medicine,* 38(5), 465–471.

Rehm, C. D., Monsivais, P., & Drewnowski, A. (2015, April). Relation between diet cost and Healthy Eating Index 2010 scores among adults in the United States 2007–2010. *Preventive Medicine,* 73, 70–75.

Rhone, A., Williams, R., & Dicken, C. (2019). *Low-income and low-foodstore-access census tracts, 2015–19* (Report No. EIB-236). U.S. Department of Agriculture, Economic Research Service.

Scharadin, B. (2022, February). The efficacy of the dependent care deduction at maintaining diet quality. *Food Policy,* 107, 102,205.

Scharadin, B., Ver Ploeg, M., & Dicken, C. (2022). Geographic boundary definitions and the robustness of common food retail environment measures. *Annals of the American Association of Geographers,* 112(5), 1,403–1,423.

Scharadin, B., & Jaenicke, E. C. (2020). Time spent on childcare and the Household Healthy Eating Index. *Review of Economics of the Household,* 18(2), 357–86.

Shaver, E. R., Sadler, R. C., Hill, A. B., Bell, K., Ray, M., Choy-Shin, J., Lerner, J., Soldner, T., & Jones, A. D. (2018). The Flint store survey: Combining spatial analysis with a modified Nutrition Environment Measures Survey in Stores (NEMS-S) to measure the community and consumer nutrition environments. *Public Health Nutrition,* 21(8), 1,474–1,485.

Sherman, J. E., Spencer, J., Preisser, J. S., Gesler, W. M., & Arcury, T. A. (2005). A suite of methods for representing activity space in a healthcare accessibility study. *International Journal of Health Geographics,* 4(1), 24.

Smith, T. A., Berning, J. P., Yang, X., Colson, G., & Dorfman, J. H. (2016). The effects of benefit timing and income fungibility on food purchasing decisions among Supplemental Nutrition Assistance Program households. *American Journal of Agricultural Economics,* 98(2), 564–580.

Taylor, R., & Villas-Boas, S. B. (2016). Food store choices of poor households: A discrete choice analysis of the National Household Food Acquisition and Purchase Survey (FoodAPS). *American Journal of Agricultural Economics,* 98(2), 513–532.

U.S. Department of Agriculture. (2022). Food and Nutrition Security webpage.

U.S. Department of Agriculture, Food and Nutrition Service. Thrifty Food Plan. (2021, August) FNS-916.

U.S. Department of Agriculture and U.S. Department of Health and Human Services. *Dietary Guidelines for Americans, 2020-2025* (9th ed.) (2020).

Venkatramanan, S., & Paramasivam, C. R. (2019). An introduction to various spatial analysis techniques. *In GIS and Geostatistical Techniques for Groundwater Science*. Elsevier.

Ver Ploeg, M., Dutko, P., & Breneman, V. (2015a.) Measuring food access and food deserts for policy purposes. *Applied Economic Perspectives and Policy,* 37(2), 205–225.

Ver Ploeg, M., Mancino, L., Todd, J. E., Clay, D. M., & Scharadin, B. (2015b). Where do Americans usually shop for food and how do they travel to get there? *Initial findings from the National Household Food Acquisition and Purchase Survey* (Report No. EIB-138). U.S. Department of Agriculture, Economic Research Service.

Volpe, R., & Okrent, A. (2012). *Assessing the Healthfulness of Consumers' Grocery Purchases* (Report No. EIB-102) U.S. Department of Agriculture, Economic Research Service.

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

Ware, B. O., Amin, M. D., Jessup, E. L., & McCluskey, J. J. (2021). Neighborhood racial composition, income, and distance to grocery retailers in Seattle. *Agricultural and Resource Economics Review,* 50(3), 512–532.

Wilde, P., Llobrera, J., & Ver Ploeg, M. 2014. Population density, poverty, and food retail access in the United States: An empirical approach. *International Food and Agribusiness Management Review,* 17(A), 1–16.

Wilkins, E., , Morris, M., Radley, D., & Griffiths, C. (2019, August). Methods of measuring associations between the retail food environment and weight status: Importance of classifications and metrics. *SSM—Population Health,* 8, 100,404.

Zenk, S. N., Schulz, A. J., Israel, B. A., Mentz, M., Miranda, P. Y., Opperman, A., & Odoms-Young, A. M. (2014). Food shopping behaviours and exposure to discrimination. *Public Health Nutrition,* 17(5), 1,167–1,176.

Appendix A: Discussion of Circana OmniMarket Core Outlets Coverage Compared With Nielsen's TDLinx Coverage in FoodAPS Sampled Areas

Circana OmniMarket Core Outlets (formerly IRI Infoscan) data has imperfect coverage. Levin et al. (2019) provided a detailed discussion of these issues, including geographic coverage and coverage by store type. The report found that in regional case studies Circana OmniMarket Core Outlets had lower coverage than other food purchasing datasets but had higher coverage rates in the Eastern United States compared with Texas. Given these regional and outlet-type coverage variations, it is important to consider the Circana OmniMarket Core Outlets coverage patterns that overlap with FoodAPS sampled areas.

Following a similar approach to Levin et al. (2019), the authors compared the Circana store counts to the TDLinx store counts across numerous individual and household characteristics using FoodAPS households. They calculated a crude coverage measure of the TDLinx sample covered in the Circana OmniMarket Core Outlets by dividing the count of Circana stores in the 20-mile radius by the count of TDLinx stores in the 20-mile radius and multiplying by 100 (ignoring whether the Circana store had a match within the TDLinx sample). This calculation is given by the equation below.

$$(100 * \frac{Count\ of\ Circana\ Stores\ in\ Radius\ 20\ miles}{Count\ of\ TDLinx\ Stores\ in\ Radius\ 20\ miles})$$

Table A.1 represents Circana coverage of TDLinx stores across the quartiles of this calculated ratio. The first quartile represents households with the worst coverage and the fourth quartile represents households with the best. For example, the 33 percent for the first quartile under rural status means that 33 percent of households with the lowest Circana-to-TDLinx store ratio live in a rural area.

Overall, the authors did not identify any coverage issues that limit the use of the Circana OmniMarket Core Outlets in the analysis. In general, the composition of rural and nonrural households, composition of FoodAPS target groups, and race and ethnicity compositions remain constant across the quartiles of the Circana TDLinx ratio. Some regional variability exists with a larger proportion of households living in the Midwest and the South in the third and fourth quartile.

Table A.1

Comparison of Circana and TDLinx store counts (demographics split into coverage quartiles)

	Crude coverage of Circana to TDLinx stores in households' FRE, by coverage quartile									
	Quartile 1[1] Lowest coverage		Quartile 2[2]		Quartile 3[3]		Quartile 4[4] Highest coverage		Overall[5] (n = 4,826)	
Race	n	Percentage	n	Percentage	n	Percentage	n	Percentage	**n**	**Percentage**
White	760	63%	851	71%	863	72%	897	74%	**3,371**	**70%**
Black	198	16%	162	14%	191	16%	150	12%	**701**	**15%**
Other	138	11%	97	8%	59	5%	95	8%	**389**	**8%**
Ethnicity	n	Percentage	n	Percentage	n	Percentage	n	Percentage	**n**	**Percentage**
Hispanic	338	28%	261	22%	125	10%	218	18%	**942**	**20%**
Non-Hispanic	878	72%	936	78%	971	81%	989	82%	**3,884**	**80%**
Rural status	n	Percentage	n	Percentage	n	Percentage	n	Percentage	**n**	**Percentage**
Rural	402	33%	281	23%	235	19%	393	33%	**1,311**	**27%**
Non-rural	814	67%	916	77%	971	81%	814	67%	**3,515**	**73%**
Target group (% FPL)	n	Percentage	n	Percentage	n	Percentage	n	Percentage	**n**	**Percentage**
Non-SNAP < 100%	106	9%	90	8%	84	7%	66	5%	**346**	**7%**
Non-SNAP >= 100% and <185%	235	19%	205	17%	207	17%	204	17%	**851**	**18%**
Non-SNAP >= 185%	455	37%	517	43%	551	46%	525	43%	**2,048**	**42%**
SNAP	420	35%	385	32%	364	30%	412	34%	**1,581**	**33%**
Region	n	Percentage	n	Percentage	n	Percentage	n	Percentage	**n**	**Percentage**
Midwest	181	15%	184	15%	601	50%	204	17%	**1,170**	**24%**
West	461	38%	313	26%	68	6%	214	18%	**1,056**	**22%**
Northeast	175	14%	113	9%	345	29%	183	15%	**816**	**17%**
South	399	33%	587	49%	192	16%	606	50%	**1,784**	**37%**

FRE = food retail environment; n = number value; FPL = Federal poverty line; SNAP = Supplemental Nutrition Assistance Program.

Note: Households that identified as American Asian and Pacific Islander, American Indian and Alaskan Native, or multiple races were dropped from this table due to their small sample size.

[1] The households in this quartile have a crude coverage of Circana to TDLinx stores within (0%, 14.29%).

[2] The households in this quartile have a crude coverage of Circana to TDLinx stores within (14.29%, 17.68%).

[3] The households in this quartile have a crude coverage of Circana to TDLinx stores within (17.68%, 23.45%).

[4] The households in this quartile have a crude coverage of Circana to TDLinx stores within (23.45%, 47.29%).

[5] Each category may not add to 4,826 due to missing responses. The % columns represent what percentage each category is of the total households in quartile (or in total, for the last column).

Source: USDA, Economic Research Service, based on data from Circana OmniMarket Core Outlets and TDLinx.

Appendix B: Methodology to Weight County-Level Characteristics by County Population

Estimating the proportion of the population that has a particular characteristic (e.g., food insecurity) using a county-level measure requires weighting the county-level proportions by the population of the county. Otherwise, counties with very large populations or with small populations would carry even greater weight when calculating the quintile average. The following methodology describes the steps taken to weight USDA's ERS Food Environment Atlas (FEA) county-level measures by county population. The description uses food insecurity as the example, but this method of weighting is applied to all comparison characteristics in the section entitled, "Comparisons of the FREHQ to Other County-Level Measures of the Food Retail Environment Overall, and by Household Characteristics."

First, the quintile based on the low-income, low-access measure from the FEA is determined for each county Federal Information Processing Series (FIPS) code. Second, the proportion of the population in counties within that quintile that are food insecure is calculated. This process is described below:

Let FI_q = the proportion of the population in quintile q for q = 1, 2, …, 5 that are food insecure

$$FI_q = \frac{X_q}{N_q}.$$

Define N_q as the total population in counties within quintile q, then

$$N_q = \sum_{i \epsilon q} N_{qi}$$

where N_{qi} is the total population in county i of quintile q.

Define X_q as the total population that are food insecure (or have some other characteristic) in counties within quintile q, then

$$X_q = \sum_{i \epsilon q} x_{qi} = \sum_{i \epsilon q} N_{qi} \hat{p}_{qi}$$

where x_{qi} is the total population that is food insecure in county i of quintile q, and $\hat{p}_{qi}$ is the food insecurity prevalence estimate (e.g., the estimated proportion of food insecure) in county i of quintile q.