Appendix C

Methods, Supporting Tables, and Maps for National-Level Analysis of Supermarket Access

I. Methods

Store Directory Development. The directory of authorized SNAP foodstores was merged with the TDLinx listing of supermarkets to take advantage of the strengths found in each source. SNAP-authorized stores that did not match with the TDLinx directory were examined in detail. After reviewing SNAP stores classified as either SM (supermarkets) or SS (superstores) in the 2006 listing, it was concluded that many of the SNAP superstores did not meet the same criteria as the industry supermarkets.[51] To further sort through which of the SNAP stores met the definition of a supermarket, researchers examined the annual total sales and food sales of all SNAP stores classified as SM and SS that did not match with TDLinx supermarkets. Stores that did not meet the annual $2 million in sales requirement or that did not have significant food sales were eliminated. Researchers used the Trade Dimensions *Marketing Guidebook* (2008) to verify where food sales data were unavailable or not current. Online sources, such as company Web sites and other online sources for additional information about a store's annual sales and the kinds of foods sold, were checked. By combining the two store listings and using outside sources for verification, researchers obtained a more comprehensive national list of supermarkets and supercenters.

The combined list of supermarkets was converted into a GIS-useable format by geocoding the street addresses into store point locations. In many instances, the two databases referenced the same store, while in other instances the geocoded location differed. To address near-matches, researchers employed a proximity analysis and an automated matching system to analyze similarities in store names and addresses and location. Some manual data analysis was also performed to identify points in the two data sets that were the same supermarket. The resulting data set included all of the more than 34,000 TDlinx store locations, as well as an additional 6,000 SNAP store locations (all of which had annual sales above $2 million and did not match a TDlinx store). The final combined data set included locations for 40,108 supermarkets and supercenters nationwide.

Walking and Driving Distances Measures of Access. Walking access measures a range of distances for which it is feasible to walk to a supermarket, while drivable distance measures a range of distances for which it is feasible to drive to a supermarket. Researchers developed a time-based distance measure equivalent for both walking and driving. A walking speed of 2 miles per hour was assumed; thus, "high" access would equal a 15-minute walk, or one-half mile in distance. For drivability, researchers assumed a point-to-point driving speed of 40 miles per hour; thus, a walking distance equivalent to "high" access based on driving is 10 miles (40mph x 0.25 hours). These measures were extended to obtain driving time-equivalent distances for "medium" and "low" access.

[51] SNAP-authorized stores were, for the most part, self-classified by the applicant. A large number of SNAP superstores are known to carry some packaged foods but are less likely to include fresh fruit and vegetables or fresh meat products.

The walkability range is categorized as high, if a supermarket is within a half mile; medium, if a supermarket is between one-half and 1 mile; and low, if the nearest supermarket is more than a mile away. Obviously, whether walking to a supermarket is feasible or not depends on more than just distance—it could also depend on whether the individual is capable of walking that distance, whether there are safe sidewalks on which to walk and controlled intersections, and whether there are other barriers, such as crime that may make walking to a store dangerous. Furthermore, the measures of distance are all Euclidian, or straight-line distance, which may not represent the actually distance that must be walked to access a supermarket. Despite these limitations, this definition of walkability is grounded in the literature. Algert, Agrawal, and Lewis (2006) defined access to stores selling a variety of produce as a walkable distance of 0.8 km., or about a 15-minute walk. Apparicio et al. (2007) measured supermarket access proximity by the number of stores within 1,000 meters, or about 0.6 miles. In a study by the California Center for Public Health Advocacy and the UCLA Center for Health Policy Research (2008), a store radius of 0.5 miles was used to determine adequate access in urban areas.

In rural areas, a drivability measure of access is also used. Drivability is categorized as either high, if a supermarket is within 10 miles; medium, if a supermarket is between 10 and 20 miles; and low, if a supermarket is greater than 20 miles away. The drivability range is not as well grounded in the literature as is the walkability range. Sharkey and Horel (2009) used a walking distance of 1 mile, and driving distances of 3, 5, and 10 miles to measure foodstore proximity in six rural counties in Texas. Morton and Blanchard used any distance outside of 10 miles to describe areas with limited access (2007). Kaufman (1999) found access to a supermarket involved a trip of more than 30 miles for 70 percent of low-income households in a 36-county area of Mississippi, Louisiana, and Arkansas. Researchers know of no other studies that have systematically characterized rural areas as having access to stores or not. The categorization is admittedly somewhat arbitrary, but is not thought to be unreasonable for illustrating national-level trends.

Use of a Kernel Density Function to Define Low-Income Areas. Two criteria were used to identify low-income neighborhoods and communities: a household income had to equal 200 percent or less of the Federal poverty threshold, and, for a given geographic area, at least 40 percent of the population had to meet that criterion. Because the geographic areas consisted of uniform 1-sq.-km. grids, a systematic search criterion was used, where for each grid, the population of adjacent grids extending 3 km. in all directions was used to test whether at least 40 percent of the total population within the search area met the poverty threshold requirement. This procedure is a type of the kernel density function. Its primary purpose here is to test each grid within the context of adjacent grids for meeting the low-income area criteria. The kernel density function also serves to smooth the observed variation in population income, resulting in contiguous (less fragmented) low-income and non-low-income geographic areas, in which the population within these areas is more similar than dissimilar with respect to the income thresholds applied.

While the choice of search area to use is not empirically derived, researchers experimented with different search areas and observed the resulting plots of low-income areas. An effort was made to avoid applying search criterion less than 3 km., in which results gave highly fragmented (pixilated) areas, interspersed with many higher income areas. Conversely, the use of a large search criterion, such as 5 km., resulted in very large low-income areas that included many higher income populations within it. In determining the kernel density search area, these factors and outcomes were considered.

II. Tables.

Table C.1

National levels of access households living in higher income areas: walking and driving distances[1]

Access level	Number of square-kilometer grids	Share of total square kilometer grids	Total higher-income area persons, by access level[2]	Share of total higher-income area persons, by access level	Number of low-income persons in higher-income areas, by access level	Share of total low-income persons in higher-income areas, by access level
		Percent	*Millions*	*Percent*		*Percent*
Walking:						
High	57,209	1.0	47.6	22.8	10.6	24.4
Medium	148,359	2.6	68.0	32.6	13.9	32.2
Low	5,433,305	96.4	92.8	44.5	18.8	43.3
Subtotal	5,638,873	100.0	208.3	100.0	43.3	100.0
Driving:						
High	3,776,567	67.0	205.2	98.5	42.4	97.9
Medium	1,292,894	22.9	2.8	1.3	0.8	1.9
Low	569,412	10.1	0.3	0.1	0.1	0.2
Subtotal	5,638,873	100.0	208.3	100.0	43.3	100.0

[1]Includes low-income households living in higher-income areas.
[2]Areas defined according to ERS criteria. See text for details.
Sources: Census of Population, 2000 and ERS-compiled 2006 supermarket directory.

Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences
United States Department of Agriculture

III. Figures.

Measuring Access Using Cumulative Density Functions (CDFs). In the area-based analysis, researchers first separated the U.S. population into one of three urbanicity types (urban, urban clusters, and rural) to compare access to supermarkets among areas having similar levels of built environment. For each urbanicity, low-income areas were compared with non-low income areas, as well as for three additional vulnerable subpopulations. Within each urbanicity-subpopulation class, distance was measured to the nearest supermarket for each grid. All grids were ranked according to distance, from closest to farthest. The corresponding populations of the ranked grids were tallied and accumulated. The resulting distances and corresponding cumulative share of total population were plotted in a cumulative density function (CDF) chart, in which for any point on the curve(s), distance (measured on the horizontal axis) and its corresponding cumulative share of population (measured on the vertical axis) can be determined.

Figure C.1. plots separate CDFs for each of the three urbanicities: urban areas, urban clusters, and rural areas. The vertical axis shows the cumulative share of the total U.S. population, with the horizontal axis indicating distance to the nearest supermarket. Using CDFs, one can examine the CDF curves of two or more subpopulations for potential access inequalities. Large differences in access appear as gaps between the individual CDF curves. The overall shape of individual curves also reflects the range of access experienced for a given share of the population or subpopulation. Access curves initially having very steep slopes and gradually flattening at the top indicate superior access relative to access curves that have a more gradual slope extending to the upper right corner of the chart. These differences can be seen in figure C.1., where dotted vertical lines demarcate distances of 0.5 and 1.0 miles, indicating high, medium, and low access for each of the three urban categories. The intersections of the horizontal lines on each curve indicate the corresponding share of the total population for each urban category. Accordingly, 74.4 percent of the urban population was within 1 mile of the nearest supermarket. Access in urban clusters was only slightly lower, where 64.9 percent of the population had a supermarket within 1 mile. In contrast, rural populations had considerably lower levels of access. Only 12.4 percent of the rural population was within 1 mile of a supermarket. This result is not unexpected, due to the much greater spatial dispersion of the rural population and its greater dependence on owned-vehicle transportation.[52]

Similar CDF charts are presented below for selected urbanicity-subpopulation combinations.

[52]In the analysis of rural area access to supermarkets, we employ both walking and driving distance criteria. We do not show the drivable distance on figures that include urban and urban cluster areas because very few people in these areas are outside of 10 miles of a super-market (in fact, almost all are within 4 miles, as shown in figure 3.1.

Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences
United States Department of Agriculture

Figure C.1
Supermarket access by urbanicity for the total population

Cumulative share of population (%)

Figure C.2
Access to supermarkets by urbanization level for non low income areas

Cumulative share of low income population (%)

Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences
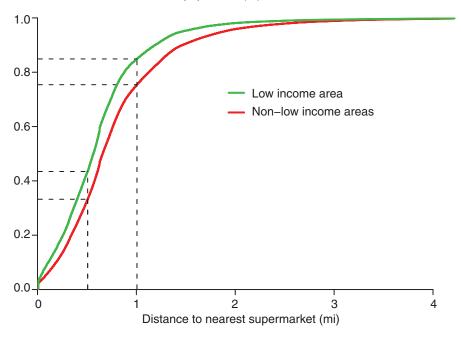United States Department of Agriculture

Figure C.3
Supermarket access for low-income individuals in low-income and non-low-income areas within urban areas
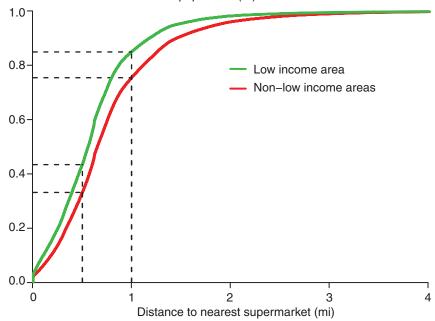
Cumulative share of low income population (%)

Distance to nearest supermarket (mi)

Figure C.4
Supermarket access for households without vehicle by low-income and non-low-income areas within rural areas
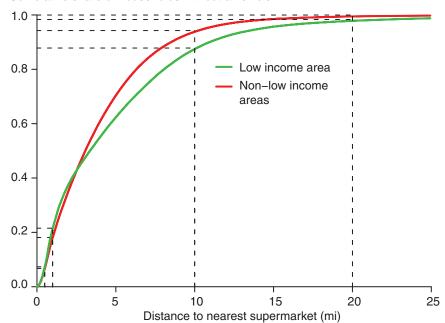
Cumulative share of households without a vehicle

Distance to nearest supermarket (mi)

Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences
United States Department of Agriculture

Figure C.5
Supermarket access for non-White individuals by low-income and non-low-income areas within rural areas

Cumulative share of non–White population

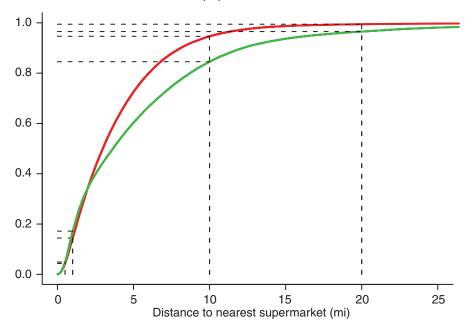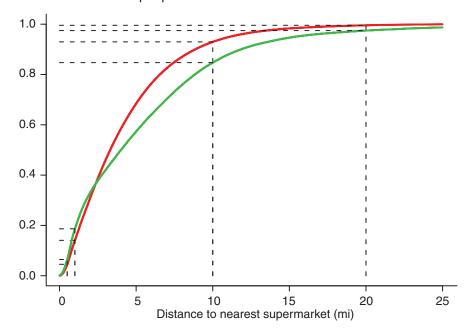Distance to nearest supermarket (mi)

Figure C.6
Supermarket access for elderly households by low-income and non-low-income areas within rural areas

Cumulative share of people 65 and older

Distance to nearest supermarket (mi)

Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences
United States Department of Agriculture

Methodology for Average Time Spent in Travel to Grocery Shopping

The estimates of average time spent in travel to grocery shopping were made using the pooled 2003-2007 American Time Use Survey (ATUS) data. The U.S. Bureau of Labor Statistics' ATUS is a continuous survey that began in 2003. One individual age 15 or older from each sampled household is interviewed about his or her use of time for the 24-hour period from 4 a.m. the day before the interview to 4 a.m. on the interview day. The time diary information includes where the respondent was and whom the respondent was with for each activity. The ATUS also collects information on the respondent's household, labor force participation, and demographic characteristics. (For more information on the ATUS, see www.bls.gov/tus).

The pooled 2003-2007 ATUS microdata files contain 72,922 completed interviews. Of those, 11,726 observations are of respondents who grocery shopped on their diary day.

Measuring travel time can be difficult and complex. Individuals string together activities with travel in between ("trip chaining"), making it difficult to separate out travel specific for a single activity. For example, an individual may leave from home and travel some distance, stop to buy coffee, then continue traveling to the work location, work a full day, then travel to the dry cleaners, pick up dry cleaning, travel to the grocery store, grocery shop, travel to a restaurant, eat at the restaurant, and, finally, travel home. In this case, neither commuting to or from work or traveling to or from the grocery store is a single travel occurrence.

The ATUS time diary data contain extensive information about Americans' travel. The data specify travel as an activity, and record mode of transportation and whom the respondent was with when traveling. If the travel was by vehicle, the data include whether the respondent was the driver or passenger. Travel is defined as moving from one location (or address) to another.

To deal with the complexities of trip chaining, ATUS codes the travel activities as to their purpose, looking ahead to the next activity and location. For example, the time a respondent travels from home to work is coded as 180501, "Travel related to working." If the respondent went grocery shopping after work, the time spent traveling to the store is coded as 180701, "Travel related to grocery shopping." The exception to the "looking ahead" rule is when the respondent is traveling from one location to home, in which case the purpose of the travel is coded as a previous activity. As a result, calculating travel time to the grocery store is complicated by the fact that some diaries will have only one "side" of travel related to grocery shopping coded as travel related to grocery shopping and others will have both sides—the going and coming home—of the trip coded as travel related to grocery shopping.

To best deal with trip chaining and the ATUS coding, researchers decided to estimate average time to the grocery store as follows. For each time diary with grocery shopping as an activity in the respondent's time diary, researchers added all legs of travel from home to the grocery store and then

Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences
United States Department of Agriculture

added all legs of travel from the grocery store to home. The total travel time home-to-shopping was compared with the total travel time shopping-to-home, and the shorter total time was chosen as the "time distance" to grocery shopping. In doing this, researchers did not have to consider the coded purpose of the travel, and they also did not have to consider the "dwell time," that is, the time spent on an activity between two travel occurrences. All the characteristics of travel to grocery shopping, such as mode of transportation, were characteristics of the shorter travel side. In cases where the respondent did not start the day at home or did not end the day at home, researchers only had information for one side (home to grocery shopping or grocery shopping to home). In these cases, the total travel time for that side was used as the time distance to the grocery shopping.

For 6.4 percent of the grocery shoppers in the ATUS data, the shortest time distance is actually from work and not from home. It was decided that the work location is a relevant means of access to grocery shopping, so for these respondents the travel time is work-to-store or store-to-work. Consequently, the average times presented use two "anchors," home and work.

Grocery shopping was defined as the ATUS activity 070101 with the location of grocery store, restaurant or bar, other store/mall, outdoors away from home, or other place. Grocery shopping with the locations home, work, church, and library were not included so as to exclude online grocery shopping. If the respondent was not at home for any activity in the 24-hour time diary, that observation was not included. If the mode of travel was by airplane, the observation was excluded as the grocery shopping was likely in an airport during out-of-town travel. Thirty-seven observations with grocery shopping were excluded for the following reasons: it was determined that the respondent was out of town when the grocery shopping took place; the time diary was not clear as to what the respondent did on the diary day, usually due to "can't remember" or "none of your business" responses; or the respondent had a large number of errands or other activities on the diary day such that the total travel time to grocery shopping would likely be an overestimate of the time distance to the grocery store. The resulting dataset used to calculate the estimates contains 11,569 observations. A small number of extreme cases were included—respondents who had zero minutes travel to grocery shopping, which is possible if the store is in the same building as the respondent's previous activity, and respondents who had 120 minutes or longer travel to grocery shopping. Additional details about variable coding and definitions are available on request.

The Bureau of Labor Statistics (BLS) and the Census Bureau calculated the estimates to ERS specifications to have estimates by whether or not the respondent was in a low-access area. Without BLS and Census cooperation, this analysis could not have been done as detailed geographical information is not publicly available in the ATUS data files.

The analysis used 90 percent confidence intervals to determine whether two average time estimates were statistically different.