

NBER WORKING PAPER SERIES

IS THE FOCUS ON FOOD DESERTS FRUITLESS? RETAIL ACCESS AND FOOD
PURCHASES ACROSS THE SOCIOECONOMIC SPECTRUM

Jessie Handbury
Ilya Rahkovsky
Molly Schnell

Working Paper 21126
<http://www.nber.org/papers/w21126>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2015

Previously circulated as "What Drives Nutritional Disparities? Retail Access and Food Purchases Across the Socioeconomic Spectrum." Prottoy Aman Akbar, Yue Cao, and Hae Nim Lee provided us with outstanding research assistance. Marianne Bitler, Anne Case, David Cuberes, Amanda Chuan, Janet Currie, Jan De Loecker, Gilles Duranton, Joe Gyourko, Jakub Kastl, Ephriam Liebttag, Ilyana Kuziemko, Todd Sinai, Diane Whitmore Schanzenbach, Jesse Shapiro, Tom Vogl, and David Weinstein provided helpful comments. We thank participants in seminars at the 2014 Urban Economics Association Meeting, the 2015 American Economic Association Meeting, the 2015 NBER Summer Institute, the Becker-Friedman Institute at the University of Chicago, Brown, NYU, Princeton, the Robert Wood Johnson Foundation, the University of New South Wales, the University of Sydney, the University of Toronto, the University of Warwick, Wharton, the Federal Reserve Bank of Kansas City, the Federal Trade Commission, and the Economic Research Service at the USDA. Jessie Handbury would like to thank the Wharton Social Impact Initiative, the Research Sponsors' Program of the Wharton Zell-Lurie Real Estate Center, and the Economic Research Service at the USDA for generous financial support. This paper was previously circulated under the title "What drives nutritional disparities? Retail access and food purchases across the socioeconomic spectrum," NBER Working Paper No. 21126, April 2015. The views expressed are those of the authors and should not be attributed to the Economic Research Service, the USDA, Nielsen, or IRI. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Jessie Handbury, Ilya Rahkovsky, and Molly Schnell. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Is the Focus on Food Deserts Fruitless? Retail Access and Food Purchases Across the Socioeconomic Spectrum

Jessie Handbury, Ilya Rahkovsky, and Molly Schnell

NBER Working Paper No. 21126

April 2015, Revised October 2016

JEL No. I12,I24,I3,R2,R3

ABSTRACT

Using novel data describing the healthfulness of household food purchases and the retail landscapes consumers face, we measure the role of access in explaining why wealthier and more educated households purchase healthier foods. We find that spatial differences in access, though significant, are small relative to spatial differences in the nutritional content of sales. Socioeconomic disparities in nutritional consumption exist even among households with equivalent access, and the healthfulness of household consumption responds minimally to improvements in local retail environments. Our results indicate that access-improving policies alone will eliminate less than one third of existing socioeconomic disparities in nutritional consumption.

Jessie Handbury
The Wharton School
University of Pennsylvania
1463 Steinberg-Dietrich Hall
Philadelphia, PA 19104
and NBER
handbury@wharton.upenn.edu

Molly Schnell
Princeton University
336 Wallace Hall
Princeton, NJ 08544
mollyks@princeton.edu

Ilya Rahkovsky
Economic Research Service, USDA
355 E Street SW
Washington, DC 20024
irahkovsky@ers.usda.gov

1 Introduction

It is well known that there are large nutritional disparities across different socioeconomic groups in the United States. Unhealthy diets could be attributed to three factors: healthy calories cost more than unhealthy calories, individuals have a lower willingness-to-pay for healthy calories, and/or individuals have limited access to healthy calories. Under the assumption that differential access plays an important role in explaining nutritional disparities, the Agricultural Act of 2014 appropriated \$125 million in federal funds to be spent annually to promote access to healthy foods in underserved communities (Aussenberg (2014)). Many state and local governments have also introduced programs to improve access by providing loans, grants, and tax credits to stimulate supermarket development and encourage existing retailers to offer healthier foods in food deserts (CDC (2011)).^{1,2,3}

Despite the popularity of such programs, little is known about their potential for narrowing nutritional disparities. Combining panel data on household purchases, store locations, product availability, and product pricing across the US from 2006 to 2011, we find that improving access to healthy foods alone will do little to close the gap in the nutritional quality of grocery purchases across households with different levels of income and education. We estimate that even if spatial disparities in access were entirely resolved, over two thirds of the existing socioeconomic disparities in nutritional consumption would remain.⁴

While there has long been agreement among researchers that both spatial disparities in access and socioeconomic disparities in nutritional consumption exist, the actual effect of access to healthy foods on food purchases has been heavily contested (Bitler and Haider (2011)). A large public health literature has inferred the impact of food environments on consumption from a cross-sectional correlation between local store density and food purchases, and finds mixed results (see

¹Government agencies have historically used the term “food desert” to refer to census tracts that meet specific criteria for both income and access. The USDA currently uses the term to refer to “neighborhoods that lack healthy food sources” (USDA (2016)). We employ the broader definition here.

²Between 2004 and 2010, the Pennsylvania Fresh Food Financing Initiative provided \$73.2 million in loans and \$12.1 million in grants to stimulate supermarket development in the state. In 2013, North Carolina House Bill 957 began granting tax credits to retailers who offer healthful foods in food deserts. In 2014, Maryland House Bill 451 provided \$1 million in assistance to food deserts through loans and grants, and the New Jersey Food Access Initiative, launched in 2009, started a private-public partnership to attract supermarkets to underserved areas. California, Nevada, New Mexico, Texas, Oklahoma, Louisiana, Illinois, New York, and DC have enacted similar legislation.

³First Lady Michelle Obama made improving access to healthy foods a cornerstone of her agenda while in the White House, stating in 2011 that “it’s not that people don’t know or don’t want to do the right thing; they just have to have access to the foods that they know will make their families healthier” (Curtis (2011)).

⁴We use “purchases” and “consumption” interchangeably. Differences in food waste, charitable giving, etc. that lead household purchases to systematically differ from household consumption are beyond the scope of this paper.

Larson et al. (2009) for a review).⁵ However, socioeconomic disparities in nutritional consumption and access could be driven entirely by differences in demand, so the observed correlation between consumption and access is not sufficient to uncover the role that access plays in generating nutritional disparities. To highlight this identification challenge we present a simple model that nests two mechanisms, one driven by access and one driven by demand, that can each independently explain the socioeconomic disparities in purchases that we observe. On the supply side, even if preferences are homothetic and identical, high-socioeconomic status (SES) households will purchase healthier bundles than low-SES households if they are more likely to live in locations where the cost of accessing healthy food is lower. On the other hand, if households sort by SES, demand-side differences that lead high-SES households to purchase healthier bundles—such as non-homothetic preferences or social norms—could themselves generate differences in access via preference externalities.

Our model motivates two complementary analyses that exploit the detailed nature of our data to identify the direction of causality in the relationship between nutritional availability and nutritional consumption. Our first empirical strategy is cross-sectional and compares the disparities in nutritional consumption that exist in the full cross-section of sample households to the disparities that persist across sample households subject to the same retail environment. We first compare the purchases of households living in the same census tract. Here, nutritional disparities between households that are above versus below the national medians for both income and education are 34% lower than in the full cross-section. It is possible, though, that households living in the same neighborhood still have differential access, either because they live in different locations within the neighborhood or because of differences in mobility. To eliminate differences in access entirely, we then look at purchases made within the same store. Here, the socioeconomic gap in the healthfulness of food purchases is reduced by only 17%.

If tastes only vary with income and education, these estimates will be exact. If tastes also vary with unobservable household characteristics, and households sort into residential and retail locations according to these tastes, then these observed within-location disparities will underestimate the disparities that would persist if retail access were equalized nationwide. Our cross-sectional results therefore indicate that eradicating spatial disparities in retail access alone would resolve at *most* one third of the observed disparities in nutritional consumption.

Our second empirical strategy is a time-series analysis that directly measures how households

⁵Despite a large policy literature on the topic, the relationship between disparities in access and nutritional consumption has been largely ignored by economists, who have instead focused on the role of differential price elasticities in generating disparities in consumption (Jones (1997); Bertail and Caillavet (2008); Park et al. (1996)).

in our data responded to changes in access over our sample period. In addition to using an alternative source of identifying variation, this method further allows us to compare the effectiveness of two common policies: incentivizing the entry of new stores versus encouraging existing retailers to offer more healthful products. Recent studies measuring the effects of changes in retail landscapes on food purchases are local in scope, looking at either the entry of a few supermarkets or an intervention to increase the availability of nutritious food products in a single urban food desert, and find modest effects (see, for example, Song et al. (2009), Weatherspoon et al. (2013), and Elbel et al. (2015)).⁶ Our time-series results provide conclusive evidence as to the generalizability of these local estimates.

We find that the elasticity of the healthfulness of household food purchases with respect to the density and nutritional quality of retailers in the household’s vicinity is positive but close to zero. Providing the typical low-SES household with the retail environment of the average high-SES neighborhood would decrease the gap in nutritional consumption across these groups by less than 5%. Using an event study specification to then look at changes in access driven by a single store entry within 2km, we find no nutritional response: households change the mix of stores in which they shop when a new store is introduced, but there is no measurable impact on the nutritional quality of household purchases. This result holds even when the healthfulness of available products increases as a result of store entry.

While comparing the purchases of the same household over time removes any correlation between changes in access and time-invariant components of household demand, changes in access may be correlated with unobserved changes in household tastes. As a result of this endogeneity, the observed response of households “treated” with changes in access provides an upper bound for the expected response of households more generally.

Our paper is related to the literature in economics that uses wide-spread changes in built environments to examine the relationship between retail environments and obesity (Currie et al. (2010); Anderson and Matsa (2011); Courtemanche and Carden (2011); Eid et al. (2008)), but departs from these previous studies in three important dimensions. First, we are concerned not just with the relationship between access and nutritional consumption, but rather the interaction between access, nutritional consumption, and household SES.⁷ This is important for evaluating the effectiveness of

⁶A notable exception is Freedman and Kuhns (2016), who examine whether the federal government’s New Markets Tax Credit (NMTC) influenced supermarket entry using a regression discontinuity design based on the program’s median family income threshold. Consistent with our results, they find no systematic difference in household purchases across tracts on either side of this threshold; however, without data on household purchases prior to the introduction of the program, they are unable to look at the response of household purchases to supermarket entry (or associated changes in retail environments) induced by the program.

⁷Currie et al. (2010) examine differences by race and education and find that the impact of fast food entry on weight

current policies, as recent efforts to improve access do so with the intent of reducing disparities in consumption across different socioeconomic groups. Second, we look directly at food purchases, the primary mechanism by which we expect changes in retail environments to impact obesity, rather than obesity itself.⁸ Finally, in addition to leveraging time-series variation in retail environments, we further use a novel cross-sectional approach to bound the maximal long-run impact of access-improving policies on socioeconomic disparities in nutritional consumption.

Dubois et al. (2014) take a structural approach to study the role that varying food environments play in generating international gaps in food purchases between the US, UK, and France. In their context, a structural demand system is necessary to measure how the purchases of households of different nationalities would adjust to being placed in the same retail environment. Our setting and data, however, allow us to use more direct, reduced-form approaches that exploit the fact that we (i) observe different household types of interest already subject to the same retail environment and (ii) can directly measure how household purchases adjust to measurable time-series variation in retail environments. In fact, in recent work, Alcott et al. (2015) estimate the demand system from Dubois et al. (2014) in the context of consumption disparities across income groups in the US and confirm our results.⁹

More broadly, our work contributes to a growing literature that studies the causes and consequences of inequality across different socioeconomic groups in the US. Recent work highlights the role that differences in environmental toxins, school quality, and neighborhoods play in generating socioeconomic disparities in health, education, and labor market outcomes (Aizer et al. (2016); Currie and Walker (2011); Currie et al. (2015); Ludwig et al. (2011); Chetty et al. (2011); Chetty and Hendren (2016)). While these studies show what neighborhoods do, our study instead provides complementary evidence on what neighborhoods do not: contrary to the popular narrative surrounding food deserts, we find that inequality in access to healthy foods is not driving the large socioeconomic disparities in nutritional consumption that we observe.

In the context of products provided in private markets characterized by increasing returns, our results indicate that the direction of causality in the relationship between neighborhood characteristics and outcomes of interest may indeed be the reverse. That is, disparities in retail access are not

gain is greatest among African American mothers and mothers with a high school education or less. In our time-series analysis, we find that more educated households respond slightly more to improvements in access to healthful foods. These differential findings are consistent with the evidence presented by Chen et al. (2010) and Volpe et al. (2013) showing that the impact of store entry depends on neighborhood characteristics and the type of store entering.

⁸While there is a large literature in economics on the relationship between socioeconomic status and various health behaviors that are known to contribute to obesity (e.g., Cutler and Lleras-Muney (2010)), grocery purchases are one health behavior which has received surprisingly little attention.

⁹We look at disparities across both income and education. Consistent with Ogden et al. (2010), we find that disparities across education groups are larger than those across households with different levels of income.

the result of supply-side market failures that in turn cause socioeconomic disparities in outcomes but instead due to efficient supply-side responses to spatial differences in demand. Given the role of fixed costs in the US supermarket industry documented by Ellickson (2006) and Hottman (2014), we expect that home-market effects and preference externalities (Helpman and Krugman (1985); Waldfogel (2003)) explain the existence of the observed disparities in access.¹⁰

If disparities in retail access do not generate the consumption disparities that we observe, then something else is to blame. There are a range of explanations for disparities in purchases, including differences in tastes or social norms, price sensitivities, and budget constraints. For the purposes of this paper, we remain agnostic as to why socioeconomic differences in household consumption exist across households in the same location and persist over time as access improves.

The paper proceeds as follows. In Section 2, we describe the datasets that we use. In Section 3, we document (i) how the nutritional quality of purchases varies across households with different levels of income and education and (ii) how access to nutritious foods varies across markets with different income and education profiles. In Section 4, we present a simple theoretical framework to demonstrate how the detailed nature of our data can be used in two complementary analyses to bound the role that access plays in generating consumption disparities. Section 5.1 implements our cross-sectional approach by looking at whether consumption disparities persist when we control for residential or retail location. Section 5.2 takes an alternative, time-series approach and examines whether we observe the healthfulness of household purchases responding to changes in local access. In Section 6, we provide a discussion of our results and conclude.

2 Data

We combine six datasets that together describe the nutritional quality of grocery purchases that households make and the retail environments that they face. We supplement this data with two additional datasets to demonstrate both the relevance and generalizability of our results. We introduce each dataset and highlight the features most relevant for our analysis below. The interested reader may refer to Appendix A for additional details.

Our identification strategies rely on inter-household (cross-sectional) and intra-household (time-series) variation in household food purchases observed in the Nielsen Homescan data. The Homescan data contains transaction-level purchase records for a representative panel of 114,286 house-

¹⁰Our evidence on the relevance of demand-side factors in explaining product availability relates to Dingel (2014) who shows that home-market demand explains as much of the positive relationship between local income and the export quality of US cities as other supply-side factors.

holds across the US between 2006 and 2011.¹¹ Households in the panel use a scanner to record all of their purchases at a wide variety of stores where food is sold. After scanning the Universal Product Code (UPC) of each item purchased, households record the date, store name, quantity purchased, and price. For items that do not have a standard UPC, households record the purchase in the relevant “random weight” category, such as “bread” or “fish.” As these categories are too broad to infer meaningful nutritional information, we only consider products with standard UPCs in our primary analysis. We demonstrate, however, that our results are robust to the inclusion of random weight items in 2006, when the random weight category definitions were more precise.

The Homescan data has three features that are important for our analysis. First, we observe household demographic data reported on an annual basis, allowing us to measure the SES of each sample household. We measure SES using household income and average years of education across household heads, controlling for household demographics.¹² Second, we observe the census tract in which each household resides. We use this information to measure the degree to which socioeconomic disparities in nutritional consumption persist when we control for each household’s retail environment. Finally, we observe household purchases over a period of up to 72 months, with the typical household appearing in the sample for 20 months. This time-series variation allows us to measure the responsiveness of household consumption to changes in their retail environment.

Since the Homescan data only includes the stores in which panelists shop and the products that they purchase at these stores, it provides a limited picture of local retail environments. Two additional datasets, both maintained by Nielsen, provide a more comprehensive picture of the retail environments that households face. The Nielsen TDLinx data, a geo-coded census of food stores in the US, enables us to calculate concentration indexes that characterize the total number of different types of stores to which households have access. To calculate indexes that depict both the nutritional quality and the relative prices of products offered at a subsample of these stores, we use the Nielsen Scantrack data. The Scantrack data contains weekly sales and quantities by UPC collected by point-of-sale systems located in over 30,000 participating retailers across the US.

¹¹We clean the sample of households provided by Nielsen, as detailed in Appendix A. The final sample used in the majority of our analyses includes 96,849 households.

¹²Much of our analysis examines disparities across households with above versus below median income and education. These cut-offs are based on household income and education that have been residualized from household size fixed effects, average head of household age, a dummy for marital status of household heads, dummies for households with either a female or male household head only, a dummy for the presence of children, and dummies for whether the household reports being white, black, Asian, or Hispanic. We first regress log income and education against these variables and then use the estimates to predict the income and education of each household were it to be a white, married couple with no children and an average age of 23. The resulting medians are \$39,221 for household income and 13.98 for years of education. On average across years, 33% of households are HI/HE, 18% are HI/LE, 17% are LI/HE, and 32% are LI/LE (Table A.1).

While comprehensive, the Nielsen datasets are not without limitations. We address potential limitations of our data in detail in Appendix A, although there are some concerns that are worth mentioning here. First, the Homescan data under-represents households at both extremes of the income distribution. We replicate our analysis non-parametrically to show that our results are not driven by households with near median income.¹³ Second, the Scantrack data over-samples grocery stores and drug stores and does not track sales of random weight products. Some of our results are subject to these concerns, whereas others are not: the TDLinx data contains information on all major food outlets, and the Homescan data includes information on *all* UPC purchases (both standard and random weight items at all purchase outlets). When our results rely on the Scantrack data, we note how these limitations influence their interpretation.

We merge the Nielsen data with three external datasets to obtain UPC-level nutritional information, tract-level driving times to stores within 40km, and tract-level neighborhood demographics. To measure the healthfulness of products purchased by Homescan panelists and offered in Scantrack stores, we use IRI's nutritional database that contains the quantity of macro-nutrients and vitamins per serving, serving size in weight, and the number of servings per container at the UPC-level. To account for the time that it takes households residing in different census tracts to arrive at stores in the TDLinx data, we collect driving and transit times from Google Maps for all stores within 40km of each census tract. Finally, to examine how both the number of stores and the quality and prices of available products vary with local income and education, we use the five-year pooled (2007-2011) American Community Survey (ACS) to measure demographics across census tracts.¹⁴

Two additional sets of data help to establish the validity and relevance of our results more generally. To demonstrate that our measures of nutritional quality correlate with health outcomes, such as BMI and hypertension, we use three waves of the National Health and Nutrition Examination Survey (NHANES; 2005-2006, 2007-2008, 2009-2010). The NHANES data combines objective measures of health outcomes with self-reported recall of two-day food consumption at the individual level. To demonstrate that our results speak to general patterns of nutritional consumption, we use the USDA's National Household Food Acquisition and Purchase Survey (FoodAPS) data. The FoodAPS data contains information on all food purchases made during a single week—that

¹³Nielsen does provide sampling weights that could also be used to address this concern. However, we instead choose to use expenditure weights to deal with measurement error associated with under-reporting.

¹⁴We separate tracts into four categories. Tracts are considered high income (HI) if their median household income falls above the median across all tracts (\$47,500) and low income (LI) otherwise. Tracts are considered high education (HE) if their share of college-educated residents falls above the median share across all tracts (21.4%) and low education (LE) otherwise. 43% of tracts are HI/HE, 16% are HI/LE, 9% are LI/HE, and 32% are LI/LE (Table A.8).

is, products with and without standard UPCs for either consumption at or away from home—for a sample of 4,826 households.

3 Socioeconomic Disparities in Nutritional Consumption and Access

In this section, we use data describing the nutritional quality of the food purchases made by households across the entire US to provide the most thorough depiction of socioeconomic disparities in nutritional consumption to date. Combining data on the spatial distribution of stores, availability of nutritious products, and relative prices of healthy-to-unhealthy foods, we then provide an equally comprehensive depiction of spatial disparities in access.

3.1 Disparities in Nutritional Consumption

We begin by documenting the extent of disparities in nutritional consumption across households with different levels of income and education. We focus on the *quality* rather than the quantity of food a household purchases since the latter is affected by the extent to which a household eats at restaurants, and a propensity for eating out is likely related both to household characteristics and access to restaurants and food stores.¹⁵ We measure the quality of household purchases using two complementary indexes, both of which are calculated at a monthly frequency for each household in our sample. Our first index, the “nutrient score,” measures the extent to which a household’s grocery purchases deviate from the nutrient composition recommended in the federal Dietary Guidelines for Americans (DGA). Our second index, the “expenditure score,” measures the extent to which a household’s grocery purchases deviate from the expenditure shares recommended by the USDA Center for Nutrition Policy and Promotion (CNPP)’s “Thrifty Food Plan.” The expenditure score follows the measure used by Volpe et al. (2013). As results are consistent across indexes, we only present the nutrient score here. The interested reader may refer to Appendix C to view results using the expenditure score.

The nutrient score for the grocery purchases recorded by household h in month t is defined as

¹⁵We use the FoodAPS data to demonstrate that socioeconomic disparities in nutritional quality are similar across (i) food purchases for consumption at home and (ii) all food purchases.

$$\begin{aligned}
Nutrient\ Score_{ht} = & \left[\sum_{j \in J_{Healthful}} \left(\frac{pc_{jht} - pc_j^{DGA}}{pc_j^{DGA}} \right)^2 \Big|_{pc_{jht} < pc_j^{DGA}} \right. \\
& + \left. \sum_{j \in J_{Unhealthful}} \left(\frac{pc_{jht} - pc_j^{DGA}}{pc_j^{DGA}} \right)^2 \Big|_{pc_{jht} > pc_j^{DGA}} \right]^{-1}
\end{aligned}$$

where j indexes nutrients, pc_{jht} denotes the amount of nutrient j per calorie in household h 's grocery purchases in month t , and pc_j^{DGA} is the amount of nutrient j in the DGA recommended diet per calorie consumed.¹⁶ These recommendations are summarized in the FDA's instructions for how to make use of nutritional labels.¹⁷ We assign the nutrients for which the recommendation is an upper bound to the unhealthful category (total fat, saturated fat, sodium, and cholesterol) and the nutrients for which the recommendation is a lower bound to the healthful category (fiber, iron, calcium, Vitamin A, and Vitamin C).¹⁸ The nutrient score penalizes households for purchasing less (more) than the recommended amount of healthful (unhealthful) nutrients per calorie. To account for differences in the units in which nutrients are measured, we normalize the deviations of household nutrient purchases from the DGA's recommendations. We follow Volpe et al. (2013) and summarize the normalized deviations using an inverse squared loss function with equal weighting across nutrients.

While useful for analysis, one drawback of indexes in general is that they can be difficult to interpret. To demonstrate that our nutrient score accords with intuition, Table 1 shows how this measure of nutritional quality varies across three sample bundles. The first bundle consists of only healthy products (broccoli, low-fat yogurt, boneless chicken breast, etc.), the second bundle contains a mix of healthy and unhealthy products, and the third bundle consists of only unhealthy products (potato chips, bacon, Oreo cookies, etc.). We determine the food products included in each bundle by selecting among the most widely purchased UPCs in each of the CNPP's 13 healthful and 10 unhealthful food categories. Full lists of the products in each bundle and the CNPP food categories from which they are drawn are provided in Tables A.14 and A.2, respectively.

¹⁶We exclude household-month nutrient scores that are more than twice the distance between the 50th and 90th percentiles from the median across all household-month observations. When using nutrient scores that include random weight items, we further exclude household-months that violate this restriction for their random weight nutrient score.

¹⁷ Refer to <http://www.fda.gov/Food/IngredientsPackagingLabeling/LabelingNutrition/ucm274593.htm> for these recommendations; last accessed on December 4, 2014.

¹⁸For random weight items, we only have imputed nutritional information for total fat, saturated fat, sodium, cholesterol, and fiber. Our nutrient scores that include random weight items are therefore constructed without information on iron, calcium, Vitamin A, or Vitamin C.

Table 1 yields three takeaways. First, as expected, the healthy bundle has a higher nutrient score than the mixed bundle, which in turn has a higher nutrient score than the unhealthy bundle. Second, the nutrient score correlates well with measures of nutrition commonly used in the literature, including fat per calorie and expenditure share on fruits and vegetables.¹⁹ Finally, lower nutrient scores are associated with higher calorie bundles. Though we prefer our nutrient score as a comprehensive and objective index of the nutritional quality of calories households purchase, it is reassuring that our index is correlated with these recognizable measures of healthfulness.

Table 1: Healthfulness of Sample Bundles

| Sample Bundle: | Nutritional Quality | | |
|--------------------------|---------------------|--------|-----------|
| | Healthy | Mixed | Unhealthy |
| Nutrient score | 0.85 | 0.77 | 0.2 |
| Total calories | 12,160 | 15,343 | 18,525 |
| Total calories per OZ | 25.75 | 32.84 | 40.08 |
| Fat (grams per 100 cal.) | 3.2 | 4.61 | 5.54 |
| Expenditure share on: | | | |
| Candy | 0.00% | 0.45% | 0.80% |
| Soda | 0.00% | 4.26% | 7.59% |
| Fruits & vegetables | 21.66% | 9.49% | 0.00% |

Notes: The above table shows how measures of nutritional quality vary across the three sample bundles defined in Table A.2. In computing expenditure shares, we use the average national price for each item in the bundle.

We are interested in the extent to which the nutritional quality of household purchases varies systematically with household characteristics. The raw averages already suggest that there are significant disparities: the average nutrient score of households with above median income and education is 1.34, 24% of a standard deviation higher than the average nutrient score of households with below median income and education at 1.01.²⁰ To absorb the effects of household composition, seasonality, and nationwide trends, in Table 2 we regress log household-month nutrient scores on log household income, average years of education across household heads, and other demographics with year-month fixed effects.^{21,22} Looking to columns (1) through (3), we see that wealthier and more educated households purchase more healthful foods. Although both effects are statistically significant, the standardized coefficients reported in column (4) reveal that education

¹⁹Refer to Tables A.4 and A.5 for correlations between the household-month and household-month-store nutrient scores used in our analysis and other more common measures of nutritional quality.

²⁰Refer to Table A.3 for summary statistics of the nutritional quality of household purchases, both in aggregate and by SES. Figure A.4 displays average household nutrient scores by deciles of household income and education.

²¹Since income is log-normally distributed whereas education is normally distributed, we include income in logs and education in levels.

²²Refer to Table A.7 for regression results showing disparities across individual nutrients.

explains more of the variation in the quality of household purchases than income. Nutritional disparities across households with different levels of education but the same level of income are over twice as large as disparities across income levels controlling for education.

Columns (1) through (4) of Table 2 only use nutrient scores computed for non-random weight purchases (that is, for products with standard UPCs). If there are systematic differences in the consumption of random weight goods across households with different socioeconomic profiles, the results excluding random weight items could either overstate or understate true disparities in nutritional consumption. In column (5), we replicate the analysis from column (4) using nutrient scores computed using both random weight and non-random weight purchases in 2006.²³ Despite a much smaller sample size, we still see that more educated households purchase food products that accord more closely with recommendations for nutritional intake.

Table 2: Household Characteristics and Nutritional Quality of Purchases

| | Ln(Nutrient Score) | | | | | |
|---------------|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------|
| | Homescan | | | | | FoodAPS |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Ln(Income) | 0.109*** (0.0040) | | 0.0482*** (0.0042) | 0.0303*** (0.0026) | 0.00262 (0.0049) | 0.0661** (0.0225) |
| Education | | 0.0519*** (0.0013) | 0.0456*** (0.0014) | 0.0854*** (0.0026) | 0.0518*** (0.0044) | 0.124*** (0.025) |
| Observations | 2,416,384 | 2,416,384 | 2,416,384 | 2,416,384 | 276,893 | 3,934 |
| R^2 | 0.015 | 0.020 | 0.020 | 0.020 | 0.007 | 0.055 |
| Standardized | No | No | No | Yes | Yes | Yes |
| Random Weight | No | No | No | No | Yes | N/A |

Notes: Standard errors are in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Columns (1)-(5) use the Homescan data; column (6) uses the FoodAPS data. Column (5) uses nutrient scores for 2006 only that include random weight purchases. When the Homescan data is used, observations are at the household-month level, standard errors are clustered by household, and year-month fixed effects are included. Observations in the FoodAPS data are at the household level. All regressions include controls for household demographics, including household size dummies, average head of household age, a dummy for marital status of household heads, dummies for households with either a female or male household head only, a dummy for the presence of children, and dummies for whether the household reports being white, black, Asian, or Hispanic. Refer to Table A.6 for the full regression results. All specifications include expenditure weights.

The Homescan data can only speak to nutritional disparities in food purchased for consumption at home. If nutritional disparities in food for consumption away from home (at restaurants, schools, etc.) follow a different pattern, then the disparities documented in columns (1)-(5) of Table 2 would not speak to food purchases more generally. In column (6) of Table 2, we replicate the analysis

²³Recall that the random weight categories in 2007-2011 are too broad to infer meaningful nutritional information for random weight items.

Table 3: Nutritional Quality of Consumption and Health Outcomes in NHANES

| | Obesity | | Diabetes | | Hypertension | |
|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Nutrient Score | -0.039*** (0.005) | -0.036*** (0.006) | -0.027*** (0.003) | -0.040*** (0.004) | -0.017*** (0.005) | -0.010* (0.005) |
| Healthy Eating Index | | -0.001 (0.000) | | 0.002*** (0.000) | | -0.001** (0.000) |
| Observations | 9,527 | 9,527 | 9,527 | 9,527 | 9,527 | 9,527 |
| R^2 | 0.045 | 0.045 | 0.091 | 0.095 | 0.238 | 0.239 |
| Mean Dep. Var. | 0.364 | 0.364 | 0.108 | 0.108 | 0.331 | 0.331 |

Notes: Standard errors are in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The above table displays the output from a linear probability model of indicators for various health outcomes on individual-level nutrient scores and the healthy eating index. These measures of nutritional quality are calculated using two-day food recalls as reported in NHANES. Explanatory variables are standardized by the variable's standard deviation. All regressions include controls for log income, education, a cubic in age, and indicators for whether the respondent is male, white, black, or Hispanic. We exclude children aged 15 and under. To proxy for expenditure weights, we weight by total calories in all specifications; results are robust to the use of alternative weights.

from columns (4) and (5) using information on the quality of *all* food purchases documented in the FoodAPS data. We see that nutritional disparities are comparable whether we consider food for consumption at home or all food purchases: wealthier and more educated households purchase food products with higher nutrient scores. Furthermore, despite the fact that income is reported as a continuous, rather than a categorical, variable in the FoodAPS data, education still explains over twice as much of the variation in the nutritional quality of food purchases as does income. It is therefore unlikely that the larger disparities by education than by income that we find in the Homescan data are driven by attenuation bias in the income coefficient.

The consumption disparities documented in Table 2 translate into meaningful differences in health outcomes. Using the NHANES data, we calculate nutrient scores for respondents based on their two-day food recall. We then regress indicators for various health conditions on individual-level nutrient scores and controls for income, education, sex, age, and race. Dividing the standardized coefficient estimates shown in Table 3 by the mean levels of each dependent variable, we see that a one standard deviation increase in nutrient scores is associated with a 11% reduction in the probability of being obese, a 25% reduction in the probability of having diabetes, and a 5% reduction in the probability of having hypertension. Furthermore, we see that our nutrient score explains an order of magnitude more of the variation in these health outcomes than the Healthy Eating Index (HEI)—a measure of diet quality commonly used to measure conformance to federal

dietary guidelines.²⁴ While a two-day food recall need not be indicative of an individual’s regular diet, we take these results as evidence that the socioeconomic disparities in nutritional consumption that we observe are important for understanding differences in health outcomes across these groups.

3.2 Spatial Disparities in Access

We now turn to documenting disparities in access to healthy foods across neighborhoods with different income and education profiles. We characterize retail environments using indexes that reflect the number of stores to which consumers have access; the healthfulness of the products available in these stores; and the prices of healthy, relative to unhealthy, products offered by these stores.

3.2.1 Store Concentration

We begin with simple concentration indexes that reflect the spatial distribution of retail food stores in and around each census tract in the US. The concentration indexes are kernel densities based on store locations from the TDLinx data and driving times computed using Google Maps. Let d_{sl} denote the driving time between store s and the centroid of census tract l , and let S_t denote the universe of stores in our sample in year t . We define the concentration index for census tract l in year t as a Gaussian kernel with a bandwidth of 10 minutes of driving time ignoring all stores further than 40km from the tract centroid:²⁵

$$Concentration\ Index_{lt} = \sum_{s \in S_t} w_{sl} \quad \text{where} \quad w_{sl} = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{d_{sl}}{10}\right)^2} & \text{if } distance_{sl} \leq 40\text{km} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

To examine how store concentration varies systematically with neighborhood characteristics, we combine these store concentration indexes with tract demographics from the ACS. Raw averages already suggest significant spatial correlation between income, education, and store concentration:

²⁴In the NHANES data, our nutrient score and the HEI are highly correlated (correlation coefficient of 0.55). While the HEI is commonly used in the literature, we prefer our nutrient score for two reasons: (1) the HEI conflates differences in quantity and quality and (2) our nutrient score explains relatively more of the variation in outcomes of public health concern.

²⁵This bandwidth was selected to match the expenditure-weighted distribution of household trips observed in the Homescan data. It implies, as is the case in the data, that a household is approximately 65% as likely to visit a store that is a 10 minute drive from the centroid of their residential census tract as they are to visit a store at the census tract centroid. Refer to Figure A.5 for a comparison of the implied bandwidth weights and the observed distribution of household shopping trips.

Table 4: Neighborhood Characteristics and Store Concentration

| | Ln(Concentration Index) | | | | | | |
|--------------------|-------------------------|------------------------|------------------------|------------------------|-----------------------|-----------------------|------------------------|
| | (1) All | (2) Grocery | (3) Conven. | (4) Drug | (5) Club | (6) Dollar | (7) Mass Merch. |
| Panel A | | | | | | | |
| Ln(Med. Income) | -0.0353*** (0.0071) | -0.0244*** (0.0071) | -0.0462*** (0.0071) | -0.0262*** (0.0071) | -0.120*** (0.0071) | -0.175*** (0.0094) | -0.0544*** (0.0075) |
| Ln(Share College) | 0.157*** (0.0071) | 0.172*** (0.0071) | 0.146*** (0.0071) | 0.171*** (0.0071) | 0.0820*** (0.0071) | 0.0579*** (0.0097) | 0.187*** (0.0076) |
| Observations | 36,951 | 36,470 | 36,489 | 36,398 | 36,640 | 19,860 | 31,175 |
| R^2 | 0.018 | 0.024 | 0.014 | 0.024 | 0.008 | 0.022 | 0.025 |
| Panel B | | | | | | | |
| Indicator for HIHE | 0.658*** (0.027) | 0.811*** (0.029) | 0.576*** (0.027) | 0.818*** (0.030) | -0.0346 (0.029) | -0.849*** (0.058) | 0.665*** (0.033) |
| Indicator for HILE | -0.239*** (0.036) | -0.187*** (0.038) | -0.247*** (0.036) | -0.238*** (0.040) | -0.463*** (0.038) | -0.713*** (0.079) | -0.482*** (0.044) |
| Indicator for LIHE | 0.555*** (0.043) | 0.638*** (0.046) | 0.541*** (0.043) | 0.605*** (0.048) | 0.257*** (0.047) | 0.150 (0.092) | 0.644*** (0.053) |
| Observations | 36,951 | 36,470 | 36,489 | 36,398 | 36,640 | 19,860 | 31,175 |
| R^2 | 0.027 | 0.032 | 0.022 | 0.031 | 0.006 | 0.015 | 0.031 |

Notes: Standard errors are in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Observations are at the tract-year level. All variables are standardized by the variable's standard deviation. See Footnote 14 for a description of how tracts are separated according to income and education levels. These results are for 2010; they are representative of other years in the TDLinx sample.

households in tracts with above versus below median income and education face concentration indexes that are on average 13% of a standard deviation higher.²⁶

In Table 4, we show how these concentration indexes vary with tract-level characteristics. In Panel A, we regress log concentration indexes on median income and the share of residents with at least a college degree in the tract. We see in column (1) that conditional on income, local education is positively associated with store concentration. The reverse holds for local income conditional on education, although this negative association is much weaker. As local income and education are highly correlated, making these results difficult to interpret, in Panel B we explore how neighborhoods with different socioeconomic profiles compare to neighborhoods with median income and the share of college-educated residents both below their respective medians. Here, we

²⁶Refer to Table A.8 for summary statistics of the variables we use to measure retail environments, both in aggregate and by tract SES. Figure A.6 displays concentration indexes by deciles of median income and college-educated shares.

see that neighborhoods with income and education above their respective medians have a greater concentration of stores than all other neighborhoods types. Columns (2) through (7) display the relationship between tract-level demographics and concentration indexes that each reflect the density of stores of a certain store type. We see that the results in column (1) do not mask significant heterogeneity across most store types. In fact, high-SES neighborhoods have the greatest concentration of all store types other than dollar stores, whose sales make up less than 1% of sales for food for consumption at home over our sample period, and club stores. Since club stores draw from a geographically disperse market, their location decisions are more likely driven by land costs and road accessibility than smaller and more local grocery, convenience, and drug stores.

There are interesting non-linearities in the spatial distribution of stores across neighborhoods with different socioeconomic profiles. While high-SES neighborhoods have more stores between 2 and 32km from their census tract centroids than low-SES neighborhoods, the opposite is the case for stores within 2km. Therefore, if we use a smaller bandwidth when computing the concentration indexes defined in Equation 1 (e.g. a driving time of 5 minutes or a distance of 1km), low-SES neighborhoods appear to have a higher concentration of stores than high-SES neighborhoods. This is consistent with the results of Powell et al. (2007) who find that low-income zip codes have more non-chain supermarkets and grocery stores than higher income zip codes.

Despite these patterns of store concentration within very small geographic bands, the evidence suggests that households primarily travel beyond their nearest store when making food purchases (Rahkovsky and Snyder (2015)). Further, Figure A.5 demonstrates that a bandwidth of 10 minutes of driving time or 3km of distance produces weights that more accurately reflect the distribution of observed shopping trips than 5 minute or 1km bandwidths. We therefore believe that a very small bandwidth overstates access in low-SES relative to high-SES neighborhoods.

3.2.2 Store Inventory and Pricing

Our concentration indexes allow us to examine disparities in the spatial distribution of retailers but fail to account for the fact that not all stores are equal. Importantly, stores may differ in the products and prices they offer, even within store types. Before proceeding to a formal analysis of store inventory and pricing, in Table 5 we explore how cost and availability differ across neighborhoods for the three sample bundles introduced in Section 3.1. The left panel of Table 5 shows the percentage of census tracts in which the entirety of a given bundle can be found across stores, conditional on there being at least one geocoded Scantrack store in the tract. We see that the unhealthy bundle is more likely than the healthy bundle to be available both across all tracts and across tracts

with different income and education profiles. Comparing availability across tracts with different socioeconomic compositions, we further note that each bundle is most likely to be available in tracts with average income and education above their respective medians. In fact, the healthy (unhealthy) bundle is 9 (8) percentage points more likely to be available in tracts with above versus below median income and education.

Table 5: Availability and Cost of Sample Bundles

| Sample Bundle: | Availability (% of Tracts) | | | Avg. Cost Per 100 Cals. (Std. Dev.) | | |
|----------------|----------------------------|-------|-----------|-------------------------------------|-------------|-------------|
| | Healthy | Mixed | Unhealthy | Healthy | Mixed | Unhealthy |
| All tracts | 48.3% | 48.3% | 79.8% | 0.43 (0.05) | 0.35 (0.03) | 0.3 (0.03) |
| HI/HE tracts | 52.1% | 52.0% | 83.2% | 0.45 (0.05) | 0.37 (0.03) | 0.3 (0.03) |
| HI/LE tracts | 47.9% | 47.9% | 78.1% | 0.42 (0.05) | 0.35 (0.03) | 0.3 (0.03) |
| LI/HE tracts | 46.7% | 46.7% | 79.3% | 0.43 (0.04) | 0.35 (0.03) | 0.3 (0.02) |
| LI/LE tracts | 43.1% | 43.0% | 75.3% | 0.4 (0.04) | 0.34 (0.03) | 0.29 (0.02) |

Notes: The above table presents the availability and cost at the tract level of the bundles defined in Table A.2. Bundle availability is calculated as the share of tracts that offer all the items (or similar products) listed in the corresponding bundles, while bundle cost is calculated as the expenditure required to buy the bundle where the price of each component is equal to the average price charged for similar products for each bundle item in a given tract-month. Similar products are defined as products in the same product module whose description contains the same key words as the description of the exact item in the bundle. For example, similar products for “Tuna–Chunk Light” are products in the “SEAFOOD-TUNA-SHELF STABLE” module with a description containing the key words “TUNA WTR CHK LT”. See Footnote 14 for a description of how tracts are separated according to income and education levels.

To measure spatial differences in inventory, we compute healthfulness indexes for each of the stores in the Scantrack sample that we are able to match to location information in the TDLinx data.²⁷ These indexes summarize the nutritional content of the products offered in a given store in a given month using a store-level variant of the nutrient score defined for households in Section 3.1.²⁸ In particular, the store-level nutrient score reflects the per calorie nutrients that a representative household would purchase in store s in month t . The representative household purchases *all* of the products available in a store such that their relative UPC-level expenditure shares for that store reflect the national average. Variation across store-level nutrient scores therefore comes only from differences in the mix of UPCs available across stores, not from differences in the quantities sold.²⁹

Differences in the store-level nutrient indexes are small. The average nutrient score of stores located in census tracts with above median income and education is 0.62 versus 0.59 across stores in census tracts with below median income and education.³⁰ To assess the magnitude of this

²⁷Refer to Appendix A.2 for details on this match.

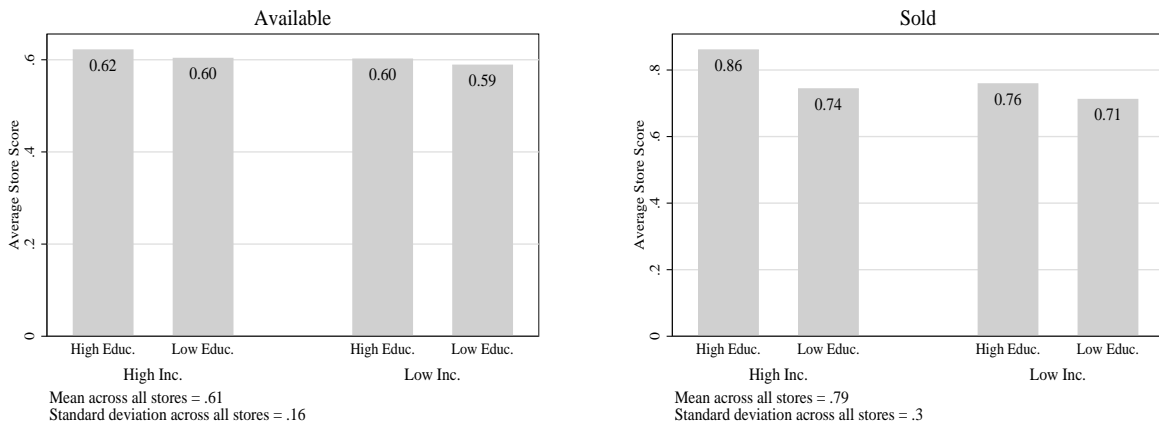
²⁸The formula for store-level nutrient scores is provided in Appendix D.

²⁹Indexes based on store-sales weights will be biased towards the tastes of households visiting that store and, therefore, will mechanically be correlated with the demographics of the store’s local community. By using national weights, we control for the relative importance of UPCs to the typical consumer without introducing this bias.

³⁰Refer to Table A.9 for summary statistics of the variables we use to assess product availability and store sales,

difference, we benchmark the gap in nutritional availability to the gap in sales measured using nutrient scores calculated with store-sales rather than national-sales weights (that is, store-level nutrient scores that reflect differences in local demand as well as product availability). In Figure 1, we see that the differences in the healthfulness of products sold across neighborhoods are much more pronounced than the differences in the healthfulness of the products available.³¹ The disparity in the nutritional quality of products sold across neighborhoods with above versus below median income and education is *five times* as large as the disparity in the nutritional quality of products offered in stores across these neighborhoods.

Figure 1: Nutrient Scores Across Census Tracts: Available vs. Sold



Notes: The above figure presents raw store-level nutrient score averages, computed using either national-sales weights (left) or store-sales weights (right), across census tracts with different socioeconomic compositions. See Footnote 14 for a description of how tracts are separated according to income and education levels. These results are for January 2010; they are representative of other months in the Scantrack sample.

Given the disconnect between the nutritional quality of products available and the nutritional quality of products sold across neighborhoods, it is unlikely that differences in product availability alone drive the observed differences in sales. That being said, there are other store policies, such as pricing and amenities, that may also influence household purchases. Even though a product is on the shelf in a low-SES neighborhood, the product may be prohibitively expensive or offered in an unkept section of the store such that the item is not truly “accessible” to households in that neighborhood.

The Scantrack data includes both revenue and quantity information across products, allowing

both in aggregate and by tract SES. Figure A.7 displays store-level nutrient scores by deciles of median income and college-educated shares.

³¹A meticulous reader may wonder whether it is possible for the nutrient scores of a nationally representative consumer to be lower than the nutrient scores of bundles actually sold across all neighborhoods. This is not an error but rather an artifact of a skewed distribution of store-level nutrient purchases combined with an index that does not reward healthy deviations.

us to infer the prices offered to consumers. In turn, we can use these inferred prices to directly examine the role of differential pricing. One commonly cited hypothesis for why low-income consumers eat less healthy foods is that unhealthy calories are less expensive than healthy calories.³² Since low-income consumers face tighter budget constraints and food is a necessity good, they will allocate more of their expenditure towards cheaper, less healthful foods than high-income consumers. While relative prices of healthy versus unhealthy food products may be a key driver of nutritional disparities in general, they are only relevant for this paper insofar as the pricing practices of the stores in low-SES neighborhoods lead low-SES households to purchase even more unhealthy foods than they would if they had access to the prices offered by stores in high-SES neighborhoods.

If store pricing is to blame for the relative unhealthfulness of sales in low-SES neighborhoods, it must be the case that either (a) these stores charge higher prices for all food products, limiting their customers' consumption possibilities and forcing them to allocate even more of their expenditure towards cheaper products than they would otherwise, or (b) these stores charge relatively more than stores in high-SES neighborhoods for healthful relative to unhealthy food products.³³ We explore these hypotheses by looking at both the spatial distribution of prices for all food products as well as the distribution of prices for healthy relative to unhealthy foods across the US.³⁴ We find that differences in pricing alone cannot be driving consumption disparities: stores in high-SES neighborhoods charge slightly more, on average, than stores in low-SES neighborhoods for all products on average, and healthful foods are actually slightly more expensive relative to unhealthy foods in these neighborhoods. Therefore, if anything, pricing patterns should cause store sales in low-SES neighborhoods to be more, as opposed to less, healthful than store sales in neighborhoods with wealthier and more educated residents.

Recall that the Scantrack sample under-represents convenience stores and other small retailers and only tracks items with standard UPCs. These two key limitations of the Scantrack data could bias our estimates of existing disparities in the availability of nutritious food, although the direction of the bias is ambiguous in both cases. Since high-SES neighborhoods have more convenience

³²In the majority of product groups, we observe that the national average price per calorie of products in healthful CNPP food categories is, on average, higher than the national average price per calorie of products in unhealthy CNPP food categories. For the sample baskets introduced in Section 3.1, Table 5 displays the average cost per 100 calories of each basket across all tracts and across tracts with different income and education profiles. Comparing the cost per calorie across bundles, we see that the healthy bundle is on average more expensive than the mixed bundle which in turn is more expensive than the unhealthy bundle.

³³Neither of these hypotheses are upheld in our samples bundles: Table 5 shows that there are no significant differences in the average prices of any of our sample bundles across tracts with different socioeconomic compositions.

³⁴The reader may refer to Appendix D for this analysis.

stores (Table 4)—small retailers that have lower nutrient scores on average than other store types (Figure A.3)—the limited coverage of convenience stores could lead us to overestimate differences in access across neighborhoods. However, it is possible that differences in the extent of competition with grocery stores allow convenience stores in low-SES neighborhoods to profitably sell healthier foods than convenience stores in high-SES neighborhoods, in which case we will underestimate differences in access. Similarly, the direction of the bias due to the lack of random-weight UPC sales reported in the Scantrack data will depend on whether healthy products without standard UPCs are more likely to be offered in low-SES or high-SES neighborhoods. Note that the majority of our main results below do not rely on these disparity estimates and thus are not subject to these measurement issues. Where appropriate, we note where our results do rely on these estimated disparities, and we provide a variety of robustness checks to help mitigate these concerns.

Finally, the manner in which healthful products are presented, including their shelf space and department cleanliness, may also make these products relatively less attractive in certain stores (see, for example, Zenk et al. (2011)). Unfortunately, we do not have the qualitative data required to assess whether these differences can help explain socioeconomic differences in store sales. In Section 5.1, however, we use fixed effects to control for *all* differences in access across neighborhoods and even across stores to obtain an upper bound on the role that these factors jointly play in explaining socioeconomic differences in household purchases.

4 Theoretical Framework

We have demonstrated that there are large socioeconomic disparities in the nutritional content of household grocery purchases as well as significant, yet more limited, spatial disparities in access to healthy foods. The direction of causality here is undetermined. It is possible that the disparities in nutritional consumption are due entirely to the fact that poorer and less educated households have access to different products than wealthier and more educated households. It is also possible that households sort into locations where they have access to the food products that they prefer to purchase or, more likely, that households sort into locations based on factors correlated with their demand for grocery products (e.g., housing prices, proximity to employment opportunities, schools, etc.) and spatial disparities in product availability arise entirely because stores cater to local demand. In reality, there are likely feedback effects between household demand and retail access.

In this section, we introduce a simple and quite general theoretical framework in which household SES and local supply both influence food purchases. This framework demonstrates the chal-

length that the previous literature has faced in identifying the causal link between access and the nutritional quality of household purchases. It also suggests two ways in which we can use the detailed nature of our data to overcome this challenge. In Sections 5.1 and 5.2, we apply each of these approaches to empirically bound the maximal impact that access-improving policies can have on socioeconomic disparities in nutritional consumption. The interested reader may refer to Appendix E for a more parametric approach to this theory.

4.1 Set-up

Consider a model with M locations indexed by l . Each location l has a population of equal size N composed of heterogeneous households whose SES, indexed by h , can take one of two values: low (L) or high (H). We rank locations by their share of high-SES households, with higher l locations having larger shares of high-SES households. We assume that the share of high-SES households in a neighborhood is exogenously determined.

Each household decides how much to consume of each product in a set of grocery products indexed by nutritional quality $q = 1, \dots, Q$, where a higher q is associated with a more healthful product. The household can also choose to consume an outside good, z . The household selects a consumption bundle to maximize utility subject to their budget constraint, which is determined by the household's income, y_h , and the cost of accessing food products of each quality in their location l , $p_l(q)$. The cost of access reflects the retail price of food products, travel costs, and storage.

The household's problem is therefore given by

$$\max_{\mathbf{x}, z} U_h(\mathbf{x}, z) \text{ subject to } \mathbf{p}'_l \mathbf{x} + p_l(z)z \leq y_h$$

where \mathbf{x} is a $Q \times 1$ vector of quantities of the differentiated grocery varieties and \mathbf{p}_l is a $Q \times 1$ vector of the prices of these products in location l .

The solution to the household's problem yields a Marshallian demand curve for products of each quality q , $x_h(q, \mathbf{P}_l)$, where $\mathbf{P}_l = (\mathbf{p}_l, p_l(z))$ is the vector of access costs in location l . The possibility that both utility and demand are a function of household SES is accounted for by the fact that these functions are indexed by h . This accommodates the possibility that there are either SES-specific tastes and/or non-homothetic preferences.

Denote by $s_h(q)$ the share of total grocery expenditures that households with SES h allocate to products of quality q . We can express this across-location expenditure share as

$$s_h(q) = \sum_{l=1}^M \theta_h(l) s_h(q, \mathbf{P}_l) \quad (2)$$

where $\theta_h(l)$ is the share of SES- h households that reside in location l and $s_h(q, \mathbf{P}_l) = \frac{x_h(q, \mathbf{P}_l)}{y_h - p_l(z) z_h(\mathbf{P}_l)}$ is the within-grocery expenditure share for products of quality q for SES- h households residing in location l . The sales-weighted average quality of food products consumed by SES- h households across all locations is therefore given by

$$Q_h = \sum_{q=1}^Q s_h(q) q \quad (3)$$

Fact 1. *If $s_h(q)$ is supermodular in SES h and product quality q (i.e., $\partial s_H(q)/\partial q > \partial s_L(q)/\partial q$), the average quality of food consumption, Q_h , is increasing in household SES.*

4.2 Mechanisms

Equation 2 highlights the separate roles that access and demand can play in generating the socioeconomic disparities in nutritional consumption that we observed in Section 3.1. We present these distinct mechanisms in the propositions below.

Proposition 1 (Supply-side mechanism). *If (i) demand does not vary with SES (i.e. $s_h(q, \mathbf{P}_l) = s(q, \mathbf{P}_l) \forall h$ in any given market l) and (ii) the spatial distribution of high-SES households is correlated with access to healthful food products (i.e., $\text{Corr}(\theta_H(l), \partial s(q, \mathbf{P}_l)/\partial q) > 0$), then the average quality of food consumption, Q_h , will be increasing in household SES.*

Proof. If $s_h(q, \mathbf{P}_l) = s(q, \mathbf{P}_l) \forall h$ in any given market l , Equation 2 reduces to $s_h(q) = \sum_{l=1}^M \theta_h(l) s(q, \mathbf{P}_l)$. That is, across-location expenditure shares only vary with SES through differences in the spatial distribution of households by SES, $\theta_h(l)$. Since $\theta_H(l)$ and $\partial s(q, \mathbf{P}_l)/\partial q$ are positively correlated across locations (by assumption), $s_h(q)$ is supermodular in SES h and product quality. Therefore, by Fact 1, Q_h is increasing in household SES. \square

By Proposition 1, if high-SES households tend to live in locations where the costs of accessing food products incentivize all households, regardless of SES, to purchase healthier foods, then high-SES households will buy healthier foods than low-SES households on average, even if the two sets of households have the same tastes. This will hold whenever $\partial^2 p_l(q)/\partial q \partial l > 0$, i.e., whenever healthier food products are sold at lower prices or are more available in neighborhoods with a larger share of high-SES residents. In practice, this pricing and availability pattern might arise

because local firms in high-SES neighborhoods cater to local high-SES tastes for these products. However, if tastes do not vary with SES, such cost differences could also arise as the result of differences in both wholesale and retailing costs. In Appendix E, for example, we demonstrate that if healthy foods cost more and rents are higher in high-SES neighborhoods, then firms in high-SES neighborhoods will have a comparative advantage in the distribution of nutritious products.

Proposition 2 (Demand-side mechanism). *If (i) supply does not vary across locations (i.e., $\mathbf{P}_l = \mathbf{P} \forall l$) and (ii) high-SES households purchase relatively more healthy products than low-SES households in all locations regardless of access (i.e., $\partial s_H(q, \mathbf{P})/\partial q > \partial s_L(q, \mathbf{P})/\partial q \forall \mathbf{P}$), then the average quality of food consumption, Q_h , will be increasing in household SES.*

Proof. If $\mathbf{P}_l = \mathbf{P} \forall l$, Equation 2 reduces to $s_h(q) = \sum_{l=1}^M \theta_h(l) s_h(q, \mathbf{P}) = s_h(q, \mathbf{P})$. That is, across-location expenditure shares equal within-location expenditure shares for each SES group. Since high-SES households purchase relatively more healthy products than low-SES households in all locations regardless of access (by assumption), $\partial s_H(q, \mathbf{P})/\partial q = \partial s_H(q)/\partial q > \partial s_L(q, \mathbf{P})/\partial q = \partial s_L(q)/\partial q \forall \mathbf{P}$. Therefore, by Fact 1, Q_h is increasing in household SES. \square

High-SES households may purchase relatively more healthy products than low-SES households for a variety of reasons. For $y_H > y_L$, this could be the result of income effects. That is, households with lower incomes may spend more on low-quality products either because they cost less or because there are complementarities between consumption of the outside good z and the quality of grocery products (as in our parametric model in Appendix E). High-SES households might also spend more on high-quality products because they attain more utility from these products, regardless of their expenditure on the outside good (due to differences in educational attainment, for example). For the purposes of this paper we remain agnostic as to why high-SES households spend more on healthy foods within locations. We simply seek to measure the role that these demand-side factors, relative to supply-side factors, play in generating the differences in purchases that we see across households living in all locations.

4.3 Empirical Approaches

In Section 5, we disentangle these supply-side and demand-side forces with two empirical approaches motivated by the model above. The first, cross-sectional approach turns off the supply-side source of heterogeneity, i.e., differences in access, by measuring the socioeconomic disparities in nutritional quality that persist across households either living or shopping in the same retail environment. The second, time-series approach instead turns off any time-invariant demand-side

sources of heterogeneity by estimating how the purchases of households with constant income and education change over time in response to varying retail environments.

Cross-sectional approach Within a location, the average quality of products purchased by SES- h households is given by

$$Q_h(l) = \sum_{q=1}^Q s_h(q, \mathbf{P}_l) q$$

Comparing the average quality across high-SES and low-SES households, we have that

$$Q_H(l) - Q_L(l) = \sum_{q=1}^Q (s_H(q, \mathbf{P}_l) - s_L(q, \mathbf{P}_l)) q$$

If high-SES households have relatively higher expenditure shares on high-quality products, then $Q_H(l) > Q_L(l)$ on average across locations. That is, we will observe socioeconomic disparities in nutritional consumption even though the households we are comparing have equivalent access. To the extent that these differences in demand yield preference externalities or home-market effects, such that higher quality products are easier to access in locations where there is high demand for them, differences in aggregate local demand will play a role in generating the correlation between $\theta_H(l)$ and $\partial s_h(q, \mathbf{P}_l)/\partial q$. Looking within locations we ignore these effects, whereby potentially underestimating the role of demand-side factors and, in turn, providing an upper-bound for the role of access.

Time-series approach Consider the framework above with locations recast as markets that are separated by time instead of by space. The change in the quality of products purchased by a household with SES h between time t and $t + 1$ is given by

$$Q_h(t + 1) - Q_h(t) = \sum_{q=1}^Q (s_h(q, \mathbf{P}_{t+1}) - s_h(q, \mathbf{P}_t)) q$$

Assuming that a household's income and tastes are constant over time—or at least over the time horizon that we consider empirically—we can estimate the elasticity of healthfulness with respect to access by regressing changes in the healthfulness of household purchases against variables that summarize store concentration and product availability.

It is possible that tastes vary over time, however, and we expect that changes in availability across markets will be correlated with unobserved changes in the prevalent tastes of local residents.

While the tastes of any one household in our panel might not reflect the prevalent local tastes, i.e., a household's tastes may not change or may change in the opposite direction, we expect that the tastes of our sample households will, on average, be correlated and covary with local tastes. As a result, we expect that our estimate of the elasticity of household purchases with respect to changes in their retail environment to be subject to an upward omitted variable bias. We therefore interpret these elasticities as an upper bound for the true elasticity that we expect to govern the response of purchases to improved access that is driven by policy as opposed to endogenous firm responses to changes in market fundamentals. As the correlation between time-variant components of a given household's demand and their retail access is likely more limited than the correlation between unobservable components of the household's tastes and where they choose to live or shop, however, we expect our time-series results to yield a tighter bound than our cross-sectional approach.

5 Role of Access in Explaining Consumption Disparities

We now implement the empirical strategies suggested by the theory above to identify the causal role of access in explaining consumption disparities. We begin by taking a cross-sectional approach and compare the disparities that persist across households living in the same residential location or shopping in the same store to the disparities that exist across the entire US. Leveraging observed changes in households' retail environments over our panel, we then directly measure how nutritional consumption responds to a changing retail environment. This time-series analysis further allows us to explore the relative effectiveness of two common policy types: incentivizing store entry versus incentivizing existing stores to offer more healthful products.

5.1 Looking Within Locations and Stores

In the analysis that follows, we measure the component of existing nutritional disparities that persists across households with equivalent access. We begin by controlling for residential location in Panel A of 6. In column (1), we replicate the analysis from column (4) of Table 2. In column (2), we control for household location using census tract fixed effects.³⁵ Comparing columns (1) and (2), we see that the association between income and the healthfulness of household purchases becomes insignificant when we compare the purchases of households residing in the same cen-

³⁵Census tracts are relatively small areas with approximately 4,000 residents on average. In spite of their small size, we do observe households with different levels of income and education living in the same tract. Over 50% of household-month observations are of households residing in tracts with both high-SES and low-SES sample residents. Refer to Figure A.8 for histograms showing the variation in income and education across households living in the same census tract.

tract. This indicates that differences in nutritional consumption across income groups can be explained by either access or demand-side factors: households may be limited by local access or stores may cater to local tastes. The relationship between education and the nutrient score, however, is more persistent: the coefficient on education remains surprisingly stable when we control for census tract of residence. This within-location analysis indicates that differential access can potentially explain all of the nutritional disparities across different income groups but less than 15% of the disparities across different education groups.

One concern with this within-location analysis is that households living in the same neighborhood may still have differential access. Even within a census tract, distance to retail outlets varies depending on the location of the household and factors such as car ownership or proximity to public transportation may yield differences in the ability of households to travel to stores. To control for these factors, we take a step further and examine how the nutritional quality of purchases varies across households shopping in the *exact* same store.³⁶ To do this, we first calculate household-store-month nutrient scores that reflect the nutritional quality of the purchases that a given household makes in a specific store in a given month.

In Panel B of Table 6, we first regress household-store-month scores against household demographics and time fixed effects with no store controls. Consistent with our previous results, column (1) shows that the healthfulness of household-store purchases is increasing in both income and education. When we control for access by including store fixed effects in column (2), the association between household purchase quality and income is cut in half. As in the within-location analysis, however, we see that access has much less potential to explain the relationship between nutritional quality and household education: comparing columns (1) and (2) in Panel B, we see that the association between household purchase quality and household education falls by less than 10% when we include store fixed effects.

In Section 3.1, we saw that disparities across education groups are larger than those across income groups. The fact that education disparities are also more persistent than income disparities within a given retail environment suggests that even though much of the disparities across income groups are resolved when controlling for access, most of the overall disparities that exist between households with different socioeconomic profiles would remain intact if access were equated. We test whether this is the case by residualizing household nutrient scores from either tract or store fixed effects estimated in regressions that are similar to those depicted in column (2) of Table 6.

³⁶We observe households with different levels of income and education shopping in the same store. Over 80% of household-store-month observations are for store-months with both high-SES and low-SES sample customers. Refer to Figure A.8 for histograms showing the variation in income and education across households shopping in the same store.

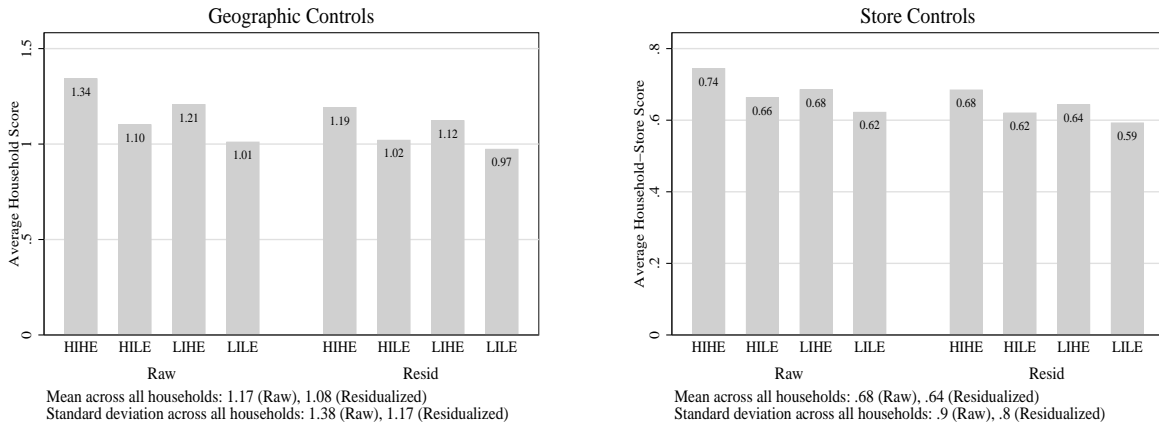
Table 6: Nutritional Quality of Household Purchases: Controlling for Access

| <i>Panel A</i> | Geographic Controls | | | | | |
|----------------------|-----------------------|-----------------------|-----------------------|------------------------|-----------------------|-----------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Ln(Income) | 0.0303*** (0.0026) | 0.00358 (0.0026) | 0.00114 (0.0027) | -0.0148*** (0.0026) | 0.0302*** (0.0026) | 0.00362 (0.0026) |
| Education | 0.0854*** (0.0026) | 0.0741*** (0.0027) | 0.0730*** (0.0026) | 0.0649*** (0.0027) | 0.0855*** (0.0026) | 0.0741*** (0.0027) |
| Ln(\$/ 100 calories) | | | 0.302*** (0.0062) | 0.289*** (0.0050) | | |
| Ln(Trips/ month) | | | | | 0.0193*** (0.0039) | 0.0160*** (0.0030) |
| Observations | 2,416,384 | 2,416,384 | 2,416,379 | 2,416,379 | 2,416,384 | 2,416,384 |
| R^2 | 0.020 | 0.176 | 0.029 | 0.182 | 0.020 | 0.176 |
| Access FEs | None | Tract | None | Tract | None | Tract |

| <i>Panel B</i> | Store Controls | | | | | |
|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Ln(Income) | 0.0239*** (0.0022) | 0.0117*** (0.0020) | 0.00454* (0.0022) | -0.00162 (0.0020) | 0.0228*** (0.0022) | 0.0114*** (0.0020) |
| Education | 0.0467*** (0.0022) | 0.0438*** (0.0019) | 0.0387*** (0.0022) | 0.0382*** (0.0019) | 0.0466*** (0.0022) | 0.0438*** (0.0019) |
| Ln(\$/ 100 calories) | | | 0.226*** (0.0053) | 0.206*** (0.0045) | | |
| Ln(Trips/ month) | | | | | -0.0745*** (0.0031) | -0.0212*** (0.0026) |
| Observations | 5,518,666 | 5,518,666 | 5,518,663 | 5,518,663 | 5,518,666 | 5,518,666 |
| R^2 | 0.014 | 0.120 | 0.018 | 0.123 | 0.015 | 0.120 |
| Access FEs | None | Store | None | Store | None | Store |

Notes: Standard errors are in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. In Panel A (Panel B), observations are at the household-month (household-store-month) level. Standard errors are clustered by household. All variables are standardized by the variable's standard deviation. All regressions include year-month fixed effects and controls for household demographics, including household size dummies, average head of household age, a dummy for marital status of household heads, dummies for households with either a female or male household head only, a dummy for the presence of children, and dummies for whether the household reports being white, black, Asian, or Hispanic. Panel A (B) includes expenditure (expenditure-share) weights.

Figure 2: Residualized Nutrient Scores Across Households



Notes: The subfigure on the left (right) presents average raw and residualized household-level (household-store-level) nutrient scores across households with different socioeconomic profiles. Residualized scores in the subplot on the left (right) are obtained by subtracting census tract (store) fixed effects estimated in regressions of the log scores against demographic controls, including interacted income and education group fixed effects, year-month fixed effects, and census tract (store) dummies. See Footnote 12 for a description of how households are separated according to income and education levels. These results are for January 2010; they are representative of the other months in the Homescan data.

Instead of controlling for continuous values of income and education, however, we control for income and education by including dummies for above-median income, above-median education, and the interaction between the two. Figure 2 depicts the average residuals for households in different income and education groups. Comparing these residualized averages to the averages of the unadjusted nutrient scores, we see that the gap between the nutrient scores of households that are above versus below median income and education are only 34% or 25% lower when we control for household location or exact retail location, respectively.

There are a variety of reasons why disparities in nutritional consumption may persist among households living in the same neighborhood or shopping in the same store. One potential reason is socioeconomic differences in budget constraints. Even though households shopping in the same store face the same product offerings and prices, households with lower levels of income may purchase less healthful bundles because they can only afford a subset of the available products.³⁷ In columns (3) and (4) of Table 6, we explore the relative importance of these economic frictions by replicating our within-location and within-store analyses controlling for monthly food expenditures per 100 calories at the household level. Comparing column (1) with column (3) in each panel, we see, not surprisingly, that the positive association between household income and purchase

³⁷As shown in Appendix D, we find that healthful products are relatively more expensive than unhealthy products in high-SES vs. low-SES neighborhoods. Differential pricing across neighborhoods cannot therefore be driving the consumption disparities that we document. However, since healthful products are relatively more expensive than unhealthy products in all neighborhoods (see, for example, the cost per 100 calories for the healthy and unhealthy sample bundles shown in Table 5), relatively high prices for healthful foods could be to blame.

quality is essentially removed when we control for how much households are spending per calorie. The relationship between household education and the nutritional quality of household purchases, however, is again more persistent: while healthier bundles cost more per calorie on average, more educated households purchase healthier bundles that cost the same on a per calorie basis as the less healthy bundles purchased by less educated households. That is, wealthier households are buying more expensive, healthier foods than poorer households with the same education, whereas more educated households are buying equally-costly, but healthier, foods than households with the same income but less education. Note that this pattern is similar across neighborhoods and stores as well as within them. This indicates that the strong relationship between education and the nutritional quality of household purchases that we find can neither be explained by differences in access nor differences in expenditures.

Of course, food expenditure is an endogenous outcome variable. In order to identify the role of budget constraints and product prices on within-store purchase disparities, one needs exogenous price variation. In ongoing work, we are exploring the relative importance of budget constraints, price sensitivities, and tastes in explaining the socioeconomic disparities in nutritional consumption that persist within a given store.

Robustness We interpret the above results under the assumption that households living in the same neighborhood or shopping in the same store have equivalent access. It is possible, however, that socioeconomic differences in mobility could threaten the credibility of this assumption. For example, differences in car ownership may allow households with higher levels of income to shop more frequently, in which case they may purchase more healthful bundles because they can purchase perishable goods such as fresh produce and dairy. To examine whether differences in shopping behaviors alter the interpretation of our results, we replicate our within-location and within-store analyses controlling for the number of shopping trips a household makes in each month.

Columns (5) and (6) of Table 6 show that both our within-location and within-store results are robust to controlling for shopping frequency. Comparing columns (5) and (6) to columns (1) and (2) in Panel A, we see that the associations between either household income or education and purchase quality are essentially unaffected by the inclusion of shopping trips per month, both across the US and within census tracts. While households that shop more frequently purchase significantly healthier bundles, differences in shopping frequency account for neither within-location nor across-location disparities in consumption. Interestingly, we see in columns (5) and (6) in Panel B that households that make more trips in a month actually purchase less healthful bun-

dles. Comparing columns (5) and (6) to columns (1) and (2) in Panel B, though, we again see that the association between either household income or education and purchase quality is robust to controlling for differences in shopping frequency.

One concern with the Homescan data that has yet to be addressed is measurement error in reported income. To address this concern, we replicate our analysis using household education alone. For completeness, we also replicate our analysis using only income. Our results are robust to both variants, although the coefficient estimates are more sensitive to controlling for access when we look at income alone (Table A.11). As disparities by education have been more persistent throughout, this is not surprising.

In Table 7, we summarize the results of a range of other robustness checks. As results are similar for our within-location and within-store analyses, we only report robustness results for our more conservative, within-location analysis here.³⁸ For each alternative specification, we replicate the analysis presented in Figure 2: In columns (2) and (3), we report average nutrient scores for households with income and education both above or both below their respective medians; the difference between these mean nutrient scores is reported in column (4). In columns (5) through (7), we replicate columns (2) through (4) reporting nutrient scores that have been residualized from census tract fixed effects in place of raw nutrient scores. Recall that the difference between the raw disparity reported in column (4) and the residualized disparity reported in column (7) provides an upper bound for the share of the raw disparity that can potentially be explained by access. For each specification, we report this upper bound in the final column.

The first row of Table 7 presents the results for our base case, where we saw that socioeconomic disparities in nutritional consumption are only reduced by 34% when we compare the purchases of households living in the same census tract. As noted in Section 2, one concern with the Homescan data is that it under-samples households with very high and very low incomes. To ensure that our results are not being driven by households with near-median income, we replicate our analysis by comparing the nutrient scores of households with income and education above the 75th (90th) percentile with the nutrient scores of households with income and education below the 25th (10th) percentile. These results confirm the generality of our results across the entirety of the income and education distributions.

In the second to last panel of Table 7, we verify that our results are robust to the use of alternative samples. First, we verify that our results are not sensitive to our definition of outliers by replicating our main analysis on a sample that excludes household-months with nutrient scores that are either above the 99th or below the 1st percentiles across all household-month observations

³⁸Refer to Table A.12 for robustness results for our within-store analysis.

Table 7: Socioeconomic Disparities in Nutritional Consumption within Tracts: Robustness

| | N (1) | Raw | | | Residualized (Tract FEs) | | | Diff-in-Diff (%) (8) = [(4) - (7)]/(4) |
|--|-----------|--------------|--------------|-------------------------|--------------------------|--------------|-------------------------|---|
| | | HI/HE (2) | LJ/LE (3) | Diff (4) = (2) - (3) | HI/HE (5) | LJ/LE (6) | Diff (7) = (5) - (6) | |
| Base Case | 2,416,384 | 1.34 | 1.01 | 0.33 | 1.19 | 0.97 | 0.22 | 34% |
| Different Cut-offs | | | | | | | | |
| P75/P25 | 2,416,384 | 1.43 | 0.96 | 0.47 | 1.25 | 0.93 | 0.32 | 32% |
| P90/P10 | 2,416,384 | 1.46 | 0.85 | 0.60 | 1.26 | 0.84 | 0.42 | 31% |
| Different Household Samples | | | | | | | | |
| Alternative outliers (excluding >P99 and <P1) | 2,551,773 | 1.44 | 1.03 | 0.41 | 1.25 | 0.98 | 0.27 | 36% |
| Excluding WIC | 2,397,366 | 1.34 | 1.01 | 0.33 | 1.19 | 0.97 | 0.22 | 34% |
| Excluding WIC, below SNAP income | 2,197,605 | 1.35 | 1.04 | 0.31 | 1.20 | 1.00 | 0.20 | 35% |
| Excluding WIC, below SNAP/CNP income | 2,119,385 | 1.35 | 1.05 | 0.30 | 1.20 | 1.00 | 0.20 | 35% |
| Excluding WIC, below SNAP/CNP income, with kids | 1,738,337 | 1.31 | 1.04 | 0.26 | 1.15 | 0.99 | 0.16 | 38% |
| Alternative Measures of Nutrition Quality | | | | | | | | |
| Absolute loss nutrient score | 2,416,384 | 0.57 | 0.46 | 0.11 | 0.53 | 0.45 | 0.07 | 34% |
| Random weight nutrient score | 276,893 | 2.74 | 2.23 | 0.51 | 2.26 | 1.92 | 0.34 | 34% |
| Expenditure score | 2,416,192 | 7.32 | 6.56 | 0.77 | 7.10 | 6.52 | 0.58 | 25% |
| Total calories (thousands) | 2,416,384 | 93.2 | 108.9 | -15.6 | 91.7 | 101.4 | -9.7 | 38% |
| Expenditure share on fruits & vegetables | 2,416,197 | 0.10 | 0.07 | 0.02 | 0.09 | 0.08 | 0.02 | 33% |
| Expenditure share on candy | 2,416,192 | 0.15 | 0.17 | -0.01 | 0.15 | 0.16 | -0.01 | 27% |
| Expenditure share on soda | 2,416,192 | 0.05 | 0.07 | -0.02 | 0.05 | 0.07 | -0.01 | 29% |

Notes: The figure above presents average raw and residualized household-level scores across households with different socioeconomic profiles. Residualized scores are obtained by subtracting tract fixed effects estimated in regressions of the log scores against demographic controls, including interacted income and education group fixed effects, month fixed effects, and tract dummies. See Footnote 12 for a description of how households are separated according to income and education levels. Our "base case" excludes outliers (more than twice the distance between the 50th and 90th percentiles from the median), uses a P50 cutoff, and includes WIC households, SNAP/CNP households, and households with kids.

rather than excluding household-months that are more than twice the distance between the 50th and 90th percentiles from the median. Next, to verify that our results are not driven by households that either report or are eligible for various types of food assistance that many alter shopping behaviors, we replicate our analysis on four additional sub-samples. These samples are designed to exclude reported WIC recipients, households that are eligible for either SNAP or free and reduced lunch, and households with kids who may receive some of their food at school.³⁹ The final column displays that our results are robust to the use of alternative household samples.

The last panel of Table 7 confirms that our results are robust to the use of alternative measures of nutritional quality. Our results are qualitatively robust to measuring nutritional quality using an absolute loss nutrient score, a nutrient score that includes random weight purchases, our expenditure score as defined in Appendix C, total calories, or the expenditure share on fruits and vegetables, candy, or soda.

5.2 Changing Retail Environments

In the analysis below, we exploit the panel nature of our data to examine how household purchases in our sample responded to changes in the availability of healthful foods in their area. Over the six years in our sample, we observe changes in the retail environments of households for three reasons: 1) the household moves to a different census tract with different access, 2) the stores in a household's neighborhood change the products they offer, and 3) stores enter and/or exit a household's neighborhood. Since household moves are endogenous, and sample households only report their residential location on an annual basis, we focus on households that reside in the same census tract throughout the sample.

We first consider how the healthfulness of incumbent household purchases responds to changes in retail environments driven by either store entry or changing product mixes. While this variation is not exogenous to the overall market in which these stores are located, these shifts in aggregate demand are more likely the result of households moving into or out of the neighborhood than shifts in the individual demand of incumbent households whose responses we are measuring. Using the resulting elasticities from this analysis, we estimate the maximal impact that resolving disparities in access could have on household consumption in the short run. Finally, since many state and federal policies targeting food deserts focus on store entry, we use an event study analysis to examine how households in our data respond to changes in access driven by store entry alone.

³⁹Refer to Appendix A.2 for details on the income cut-offs used to identify eligible households.

5.2.1 Time-Series Analysis

We estimate household responses to changes in retail environments by regressing household nutrient scores against two separate time-varying kernel densities that capture changes in local retail environments.⁴⁰ The first is the concentration index that we introduced in Section 3.2. Recall that these indexes are kernel densities that measure the driving time-weighted number of stores in a household’s vicinity. As the TDLinx data contains a snapshot of retail environments as of June each year, these concentration indexes are at the tract-year level.⁴¹ To measure the nutritional quality of the products offered in local stores, we construct kernel densities of store nutrient scores.⁴² To match the frequency of the TDLinx concentration indexes, we aggregate our monthly household and store nutrient score densities to the annual frequency using a simple average.

The results of this time-series analysis are provided in Table 8. In column (1), we see that even after controlling for the concentration and healthfulness of surrounding stores, household purchase quality is still increasing in income and education. Perhaps unsurprisingly, household nutrient scores are positively associated with both store concentration and the healthfulness of product offerings. We cannot interpret these coefficients causally, however. While column (1) controls for observable household characteristics, households may sort spatially by unobservable characteristics that are correlated with tastes for healthy foods. If households with stronger tastes for healthy foods sort into locations where these foods are more accessible, then our coefficients on store concentration and store nutrient scores will be biased upwards. If we assume that household preferences are fixed over the time period that we study (up to six years), we can control for this sorting behavior by including household fixed effects.^{43,44} In such a specification, the coefficients are identified from time-series variation in purchases and retail environments. In column (2), we observe the nutritional quality of household purchases responding to improvements in the nutrient composition of products sold by stores in their neighborhood but not to changes in the

⁴⁰As there is limited variation in price indexes across neighborhoods (Figures A.19 and A.20), we exclude prices from this analysis. The interested reader may refer to Appendix D for time-series results including price indexes.

⁴¹In this section, years are defined July-June to accord with the TDLinx snapshots.

⁴²We weight store scores using the same Gaussian kernel used to construct the concentration indexes (that is, a bandwidth of 10 minutes of driving time ignoring stores further than 40km from the tract centroid). Letting S_t denote the universe of stores in year t and N_{slt} the average nutrient score of store s in census tract l across months in year t , the weighted average nutrient score for census tract l in year t is given by $\frac{1}{\sum_{s=1}^{S_t} w_{st}} \sum_{s=1}^{S_t} w_{st} N_{slt}$, where w_{st} is defined as in Equation (1).

⁴³Since demographics are nearly constant across our sample period for a given household, we do not control for income, education, and other household demographics when we include household fixed effects.

⁴⁴For our preferred specifications that include household fixed effects, we use Conley Spatial HAC standard errors that allow for both temporal and spatial correlation. We used the `reg2hdspatial` package written by Thiemo Fetzer; available at <http://www.trfetzer.com/conley-spatial-hac-errors-with-fixed-effects>.

concentration of retail outlets in the household’s vicinity.

To explore whether the responsiveness of household purchases to changes in retail environments varies by household SES, we interact the access kernel densities with demeaned income and education in column (3). We see that the statistically insignificant average response of household-level nutrient scores to the concentration of stores in their vicinity does not mask any differences across households with different levels of income or education. We do observe a socioeconomic disparity in the responsiveness of household nutrient scores to local product offerings, with more educated households improving their consumption even more than less educated households when they are offered a more nutritionally-balanced mix of food products in their neighborhood stores.

Table 8: Response of Nutritional Quality of Household Purchases to Changes in Retail Access

| | Ln(Nutrient Score) | | |
|----------------------------|---------------------|---------------------|---------------------|
| | (1) | (2) | (3) |
| Ln(Income) | 0.041*** (0.007) | | |
| Education | 0.044*** (0.002) | | |
| Ln(Store Concentration) | 0.016*** (0.002) | 0.009 (0.009) | 0.010 (0.009) |
| * $\widetilde{Ln(Inc)}$ | | | 0.001 (0.002) |
| * \widetilde{Educ} | | | 0.000 (0.001) |
| Ln(Local Avg. Store Score) | 0.266*** (0.023) | 0.167*** (0.015) | 0.169*** (0.015) |
| * $\widetilde{Ln(Inc)}$ | | | -0.015 (0.014) |
| * \widetilde{Educ} | | | 0.015*** (0.005) |
| Observations | 244,528 | 244,503 | 244,503 |
| R^2 | 0.038 | 0.001 | 0.001 |
| Demographic Controls | Yes | No | No |
| Household FEs | No | Yes | Yes |

Notes: Standard errors are in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Observations are at the household-year level. Column (1) reports two-way clustered standard errors by tract and year; columns (2) and (3) report Conley Spatial HAC standard errors. A tilde represents that the variable has been demeaned. All regressions include year fixed effects and demographic controls including household size dummies, average head of household age, a dummy for marital status of household heads, dummies for households with either a female or male household head only, a dummy for the presence of children, and dummies for whether the household reports being white, black, Asian, or Hispanic. The sample only includes households that resided in the same census tract throughout their entire participation in the Homescan panel.

Despite being statistically different than zero, we note that the improvements in nutritional

consumption documented in Table 8 are small when compared to the existing socioeconomic disparities in nutritional consumption. To get a sense of what the magnitudes of the coefficients in Table 8 imply, we consider how a low-SES household would respond to a change in their retail environment equivalent to moving from the average low-SES to the average high-SES neighborhood. We focus on a household with income and education at the 25th percentile of each dimension, i.e., \$32,500 in annual income and 13 years of education. The elasticities of the nutrient score for such a household with respect to store concentration and the local average store score implied by the coefficients from the regression displayed in column (3) are 0.010 and 0.154, respectively.⁴⁵ Moving from the average low-SES to the average high-SES neighborhood translates to an increase of 0.66 and 0.04 in the log store concentration index and log driving time-weighted average of store nutrient scores, respectively. Combined with the estimated elasticities, these improvements in access imply that the nutrient score of a typical low-SES household would improve by 0.013 log units if they were to face the same store concentration and product availability as households living in the average high-SES neighborhood. Comparing this change to the socioeconomic disparity in raw household scores shown in Figure 2, we see that only 4.4% of the gap in nutrient scores would be removed by closing the gap in access to healthy foods. That is, even if low-SES households faced the products and concentration of stores found in high-SES neighborhoods, over 95% of the existing socioeconomic disparities in nutritional consumption would remain.

One concern with these results is that households may substitute calories from consumption away from home to consumption at home when their retail environments improve. If consumption at home is healthier than consumption away from home (as is the case in the FoodAPS data; see Table A.13) then the changes in the nutritional quality of food for consumption at home documented in Table 8 will understate changes in consumption more generally. While the FoodAPS data is not rich enough to explore how households substitute between consumption at home and consumption away from home when local access changes, we can combine estimates of the elasticity of total calories with respect to access in the Homescan data with the average nutrient scores at and away from home in the FoodAPS data to get an idea of how this substitution could impact the interpretation of our results. Replicating the analysis in Table 8 using total calories in place of nutrient scores, we find that the average low-SES household will increase their calories consumed at home by 1% if moved from the average low-SES to the average high-SES neighborhood (Table 9). Assuming that the average low-SES household (i) originally divides their calories 80-20 across food

⁴⁵Log income and education are demeaned in these regressions, so the elasticities are calculated as $\beta_0 + \beta_1 (13 - \overline{Educ}) + \beta_2 (\ln 32500 - \ln \overline{Inc})$, where β_0 , β_1 , and β_2 are the coefficients on a particular kernel density, the density interacted with demeaned education, and the density interacted with demeaned income, respectively. \overline{Educ} is the sample mean education level (14.3 years), and \overline{Inc} is the sample mean income (\$50,333).

at home and food away from home and then (ii) substitutes 0.8% ($0.8 \times 1\%$) of their total calories from food away from home to food at home when moved to the average high-SES neighborhood, the average nutrient scores for low-SES households at and away from home in Table A.13 imply that the household's overall consumption would improve by less than 1% from such a move.⁴⁶

Since the Scantrack data is more likely to track large retailers, our estimated elasticities could be biased. Suppose that bodegas and other small stores that we do not observe stop offering healthy foods when we observe a grocery store entering. If this is the case, we will underestimate the response of households to changes in their retail environments because we will overestimate the extent of changes in access. If, however, small retailers increase the healthy products they offer in order to compete with new entrants, we will overestimate household responses. As the majority of public funds are directed towards the entry of large retailers rather than small stores that would not be tracked by Nielsen, the general equilibrium effects that we estimate are suggestive of household responses to a large class of policy interventions. Furthermore, when using the estimated elasticities to measure how the nutritional consumption of the typical low-SES household would change if moved from the average low-SES to the average high-SES neighborhood, any bias in the estimated elasticities will be mitigated by any bias in the estimated disparities. If, for example, bodegas were offering relatively more healthy foods before a large retailer entered, then this limitation of the Scantrack data we would lead us to (i) underestimate the household elasticity but (ii) overestimate differences in access across neighborhoods.

Robustness As in Section 5.1, we want to ensure that our results are robust to the use of alternative household samples and alternative measures of household purchase quality. Additionally, given that our time-series results rely on our measures of access defined in Section 3.2, we further want to verify that our results are robust to the use of alternative bandwidths in our kernel density construction. The results from a variety of robustness checks are summarized in Table 9. For each robustness check, we run the regression in column (3) of Table 8 and use the estimated coefficients to recalculate the implied change in household consumption that would be observed by moving a household at the 25th percentile of both income and education from the average low-SES to the average high-SES neighborhood. The first column of Table 9 reports this implied log change in the household score, the second column reports the disparity between low-SES and high-SES households, and the third column reports the implied change as a percent of the existing disparity.

⁴⁶This calculation assumes that the household's nutrient scores at and away from home would not change with such a move. Even if the nutrient score both at and away from home increased by 5%, as is suggested from our analysis of food for consumption at home, the combination of (i) nutrient score improvement and (ii) calorie substitution from food away from home to food at home would improve the household's overall consumption by less than 6%.

The first row of Table 9 presents our base case, where we again see that our base specification is associated with a 4.4% reduction in the disparity between high-SES and low-SES households.

Table 9: Response of Household Purchases to Changes in Retail Access: Robustness

| | Total Implied Log Change (t-stat) (1) | Log Disparity Between HIHE & LILE Households (2) | % Change in HIHE - LILE Gap (3) = -(1)/(2) |
|---|---|---|--|
| Base Case | 0.013 (0.006) | 0.28 | -4.42% |
| <i>Different Household Samples</i> | | | |
| Underserved households | 0.010 (0.007) | 0.28 | -3.38% |
| Alternative outliers (excluding >P99 and <P1) | 0.009 (0.007) | 0.28 | -3.00% |
| Excluding WIC | 0.012 (0.006) | 0.28 | -4.20% |
| Excluding WIC, below SNAP income | 0.011 (0.007) | 0.28 | -3.72% |
| Excluding WIC, below SNAP/CNP income | 0.008 (0.006) | 0.28 | -2.94% |
| Excluding WIC, below SNAP/CNP inc, w/ kids | 0.010 (0.007) | 0.28 | -3.46% |
| <i>Alternative Measures of Nutrition Quality</i> | | | |
| Expenditure score | 0.013 (0.004) | 0.11 | -11.37% |
| Total calories | 0.011 (0.005) | -0.15 | 7.13% |
| Expenditure share on fruits & vegetables | 0.002 (0.001) | 0.02 | -9.51% |
| Expenditure share on soda | 0.000 (0.001) | -0.02 | 1.06% |
| <i>Different Kernel Densities</i> | | | |
| Driving time weight, 5 minute bandwidth | 0.010 (0.003) | 0.28 | -3.49% |
| Driving time weight, 15 minute bandwidth | 0.017 (0.016) | 0.28 | -6.00% |
| Distance weight, 3km bandwidth | 0.002 (0.002) | 0.28 | -0.61% |
| Distance weight, 5km bandwidth | 0.003 (0.005) | 0.28 | -1.01% |
| Transit time weight, 30 minute bandwidth | 0.006 (0.005) | 0.28 | -2.05% |
| Transit time weight, 45 minute bandwidth | 0.003 (0.003) | 0.28 | -0.93% |
| <i>Store Weights</i> | | | |
| Weekly sales revenue | 0.008 (0.010) | 0.28 | -2.88% |

Notes: The above table reports estimated changes in the nutritional quality of household purchases, replicating the regression in column (3) of Table (8) with various household samples, measures of nutritional quality, and kernel density bandwidths. Column (1) reports the estimated log change in the nutritional quality of a household at the 25th percentile for income and education were they to move from the average low-SES to the average high-SES neighborhood. Column (2) reports the observed log difference in nutritional quality between high-SES and low-SES households. Column (3) reports the estimated log change as a percent of the observed log difference between high-SES and low-SES households. Our “base case” excludes outliers (more than twice the distance between the 50th and 90th percentiles from the median), using a bandwidth of 10 minutes of driving time, and includes WIC households, SNAP/CNP households, and households with kids.

It is possible that a low-SES household originally residing in a neighborhood with below average access would respond more from moving to the average high-SES neighborhood than would a low-SES household originally living in the average low-SES neighborhood. To address the possibility of non-linear responses, we estimate the responsiveness of households in underserved neighborhoods to changes in their retail environments, where a neighborhood is defined as underserved if either its concentration index or nutrient score kernel density falls in the lowest quartile across

all census tracts. In the first panel of Table 9, we see that low-SES households in underserved neighborhoods actually respond less to improvements in their retail environments than comparable households in the average low-SES neighborhood. Overall, moving a low-SES household from an underserved neighborhood to the average high-SES neighborhood would reduce existing nutritional disparities by only 3.4%. Looking to the remainder of the “Different Household Samples” panel, we see that our results are robust to both the use of an alternative definition for outliers and to the exclusion of households that report or are eligible for various food assistance programs.

To make our results more comparable to previous studies, we replicate our analysis using alternative measures of nutritional quality in the second panel of Table 9. While the precise predicted impact of moving the typical low-SES household from the average low-SES to the average high-SES neighborhood unsurprisingly depends on the measure of nutritional quality used, the overall story is the same: equating access alone will not be sufficient to erase socioeconomic disparities in nutritional consumption. While we observe the greatest response when nutrition is measured using our expenditure score defined in Appendix C or expenditure shares on fruits and vegetables, nearly 90% of the disparities in these measures across low-SES and high-SES households remains when low-SES households are subject to the access of the average high-SES neighborhood.⁴⁷

While our base specification uses driving times to weight stores in constructing our access kernel densities, it is possible that distance or transit times are more relevant for households. In the second to last panel of Table 9, we replicate our analysis using access kernel densities constructed using a bandwidth of either 3 or 5km of distance or 30 or 40 minutes of transit time in place of a driving time bandwidth of 10 minutes. Using these alternative bandwidths, we predict a reduction in nutritional disparities in consumption of less than 1.5% by granting low-SES households access to the store concentration and product offerings of the average high-SES neighborhood. Using indexes that are again based on driving times but with a bandwidth of 15 minutes, we estimate an improvement in the nutritional consumption of a low-SES household that reduces the gap in nutritional consumption between low-SES and high-SES households by more than before, but still by less than 7%.

Finally, we note that our results are not being diluted by our equal weighting of stores of all sizes. In the final panel of Table 9, we see that even if we weight stores by their revenues in the construction of concentration indexes and local average store scores, we still see that improvements in household consumption are minimal. In fact, when weighting by store revenue when estimating

⁴⁷One interpretation of these results is that it takes time for households to learn how to eat according to the nutritional recommendations of the DGA, so household nutrient scores may take more time to respond. The healthfulness of fruits and vegetables is perhaps more salient to consumers from the outset, so the response of household expenditures on these products is closer to the long-run response that we observed in the cross-sectional results in Section 5.1.

household responses to changing retail environments, we find that moving a low-SES household from the average low-SES to the average high-SES neighborhood would result in less than a 3% reduction in disparities in nutritional consumption across these groups.

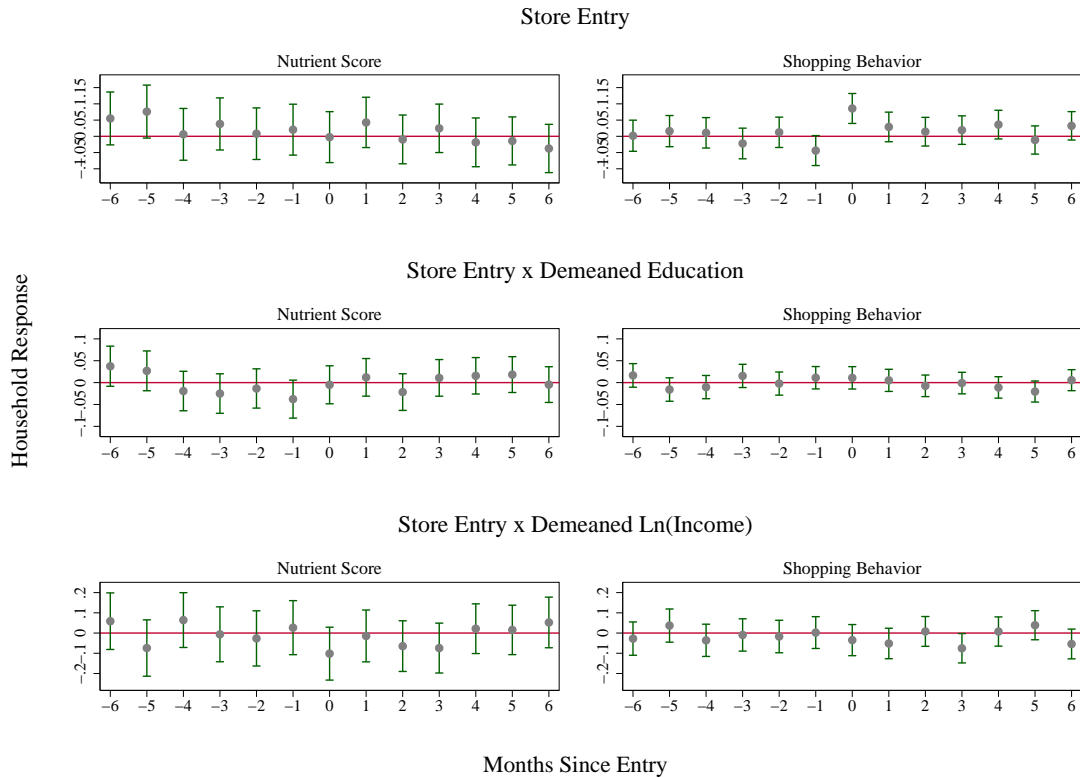
5.2.2 Event Study Analysis of Store Entry

Though some policies aimed at eradicating food deserts encourage incumbent stores to change their product offerings, most do so by encouraging store entry. While our results in Section 5.2 already suggest that the nutritional quality of household purchases does not respond to changes in the concentration of stores in the household's vicinity, it is worthwhile to consider how households respond to changes in their retail environments that are driven by store entry alone. We define a store as entering in a given month if (i) the store is first observed in the Scantrack data in that month, (ii) the store's parent company already appeared at least once in the Scantrack data prior to that month, and (iii) there was no store in the same sub-channel in the same census block in the TDLinx data in the previous month. We require the parent company to already be in the Scantrack data to avoid confusing sample growth with actual store entry. We require that no store in the same sub-channel be in the same census block in the TDLinx data in the month prior to avoid categorizing store re-branding due to merger and acquisition activity as store entry. We observe 2,106 entries between 2006 and 2011 in the Scantrack data that satisfy these requirements (66% of potential entries with location information).

To measure household responses to extensive margin adjustments in their retail environments, we use an event study specification. Specifically, we regress the log of household-month nutrient scores on household fixed effects, month-year fixed effects, and dummies for each of the six months before, the month of, and the six months after the entry of a grocery store within 2km of a household's census tract centroid. We plot the coefficients of the time-since-entry dummies in the first column of Figure 3. The top panel displays the average response across all households. We do not see any statistically significant response in the nutritional quality of the average household's purchases to store entry. The second and third panels display the gradient in the response with respect to household education and income, respectively. Here, we see that the null impact of store entry on the average household's nutrient score does not mask any significant heterogeneity across households with different levels of income and education: the nutritional consumption of high-SES and low-SES households alike does not respond to store entry.

The second column of plots in Figure 3 show that the lack of responsiveness of household nutrient scores to store entry is not due to the fact that household shopping behavior itself fails

Figure 3: Event Study Analysis of Store Entry



Notes: The above plots display the results from an event study analysis of store entry. The first column depicts the coefficient estimates on dummies for months before, during, and after store entry from a regression of log household-level nutrient scores on household fixed effects, month-year fixed effects, and dummies for each of the six months before, the month of, and the six months after the entry of a grocery store within 2km of a household's census tract centroid. The second column depicts the results from a regression of an indicator for whether the household shopped in a new store in that month on the same independent variables.

to respond. Here, we run the same event study specification using an indicator for whether the household visits a new store in a given month relative to the previous month as the dependent variable. In the first panel of the second column, the significantly positive coefficient in month zero indicates that households change the mix of stores in which they shop when a new store enters their neighborhood. As it returns to zero rather than becoming significantly negative, households who shop at the new store in period zero continue to shop at the store in each of the six months after entry. The coefficients on the time-since-entry dummies interacted with household education and income, displayed in the second and third panels in the second column, indicate that the likelihood to try a new store in the month of entry does not vary with these socioeconomic characteristics. Together, the results in Figure 3 indicate that while households are changing where they shop when a new store enters, they are not changing the healthfulness of the foods that they purchase.

These results are consistent with the null response of households to changes in the concentration of stores in their vicinity that was observed in Table 8. In fact, even if a store entering a

neighborhood 2km from the census tract centroid is associated with only a two minute drive, this store entry would only improve the neighborhood's driving time-weighted concentration index with a bandwidth of 10 minutes by 0.39. To put this number in context, the difference in concentration indexes across the average low-SES and high-SES neighborhood is 18.92 (Table A.8). Given the estimated reduction of only 4.4% in nutritional disparities from moving a household at the 25th percentile of both income and education from the average low-SES to the average high-SES neighborhood (Table 9), the potential impact of a single entry event alone is unsurprisingly minimal.

Given that we saw households responding to improvements in the product offerings of local stores in Table 8, one might expect households to respond more when the entering store improves the nutritional quality of the products available in their retail environment. To examine whether this is the case, we run the event study specification including interactions between the time since entry dummies and the nutrient score of the entering store relative to the average local nutrient score in the month before entry. Figure A.12 presents the coefficients on these interactions and we again see no effect. Households are no more likely to improve the healthfulness of what they purchase when an entering store's offerings are more healthful than that of incumbent stores. However, more educated households are more likely to change the mix of stores at which they shop when a healthier store enters.

Robustness One might suspect that improvements in access in underserved neighborhoods will be met with greater responses of household purchases. To see if this is the case, we replicate the analysis from Figure 3 looking only at households residing in tracts in the lowest quartile for either store concentration or nutrient score density. The results, presented in Figure A.11, are nearly identical to those estimated on the full sample.

Finally, we want to verify that our results are robust to store entry within different bandwidths of census tract centroids. In Figures A.13 and A.14, we show that results are very similar whether we consider stores entering within 1 or 4km, respectively, instead of 2km as in Figure 3. Although we do observe household nutrient scores responding in the month after entry when we considering stores entering within 4km, there is still no lasting impact of store entry on the nutritional quality of household purchases.

6 Discussion and Conclusion

Despite an absence of evidence drawing a causal link between disparities in retail access and disparities in nutritional consumption, much of the discussion surrounding food deserts assumes that equalizing access will eliminate nutritional disparities across different socioeconomic groups. Such an assumption underlies policies that aim to improve the quality of food purchases by increasing the availability of healthful products in areas with unhealthful consumption. Contrary to this assumption, our analysis indicates that the large socioeconomic disparities in nutritional consumption that we document across households are not driven by the relatively limited differences in access to healthy foods that we observe across neighborhoods with different socioeconomic compositions.

If differential access is entirely to blame for nutritional disparities, then any systematic differences in the nutritional quality of household purchases that we observe when looking across the entire US should disappear when we compare households living in the same neighborhood or shopping in the same store. On the contrary, even when we control for residential or retail location, we observe households with higher levels of income and education making purchases that are significantly closer to DGA recommendations than households with lower levels of income and education. These cross-sectional results indicate that most of the existing socioeconomic disparities in nutritional consumption cannot be reduced by improving access alone: even if spatial disparities in access across the US are entirely resolved, at least two thirds of existing disparities in nutritional consumption across households with different levels of income and education would remain.

We stress that even though socioeconomic disparities diminish when we control for residential or retail location, it is unlikely that resolving spatial disparities in access will reduce disparities across the entire US to the same extent. If households are sorting into retail environments on unobservables that are correlated with their taste for healthy foods, then the socioeconomic disparities that we observe for households living in the same location or shopping in the same store will be smaller, on average, than the socioeconomic disparities that would persist across the full cross-section of households if access were equated. Therefore, while our estimates indicate that equating access across the entire US could not reduce existing socioeconomic disparities by more than one third, it is likely that the true impact would be even smaller.

Policies that target access in the hope of improving the healthfulness of local consumption do so both by encouraging existing retailers to offer more healthful products and by stimulating store entry. These policies will only be effective insofar as the healthfulness of household purchases respond to changes in their retail environment. Contrary to this ideal, we find that the response of

a given household's purchases to changes in their local access is very limited. Moving the typical low-SES household to the typical high-SES neighborhood would only serve to reduce the gaps in nutritional consumption between these two groups by less than 5%. While our event study analysis indicates that households shop at entering stores, the quality of their nutritional consumption does not improve.

Despite the limited responsiveness of household purchases to the endogenous changes in access that we observe, it is again likely that, on average, households across the entire US will respond even less to changes in their retail environments. Improvements in access to healthy foods are more likely to occur in close proximity to sample households with growing tastes for these products, so the changes in the purchases of these households will reflect not only changes in access but also changes in tastes. This correlation between the time-variant component of demand and changes in access yield an upward-biased estimate of the effect of access-improving policies that are implemented independent of changes in local demand conditions. Therefore, while our estimates indicate that the nutritional quality of household purchases responds minimally to changes in retail environments, it is likely that the impact of policy-induced changes on nutritional consumption would be even smaller.

The results of our time-series analysis are smaller than our cross-sectional results. We find this unsurprising for two reasons. First, we expect differences in demand within a household over time to be more limited than differences in demand across households living in the same location, so the bound that we estimate using within-household variation should be tighter than the bound that we estimate using across-household variation. Second, our time-series analysis only captures short-run effects of improvements in access on household consumption, while our cross-sectional results may also capture long-run changes in consumption behaviors that result from living in an environment with greater access to healthful products. Viewed through this lens, our results indicate that access-improving policies will have larger effects in the long run than in the short run, although even in the long run such policies will eliminate at most one third of current socioeconomic disparities in nutritional consumption.

Taken together, our results provide strong evidence that policies that aim to reduce nutritional disparities by improving access to healthful foods alone will leave much of the disparity unresolved. Differences in demand across socioeconomic groups yield empirically relevant disparities above and beyond those that could also be attributed to the sorting of households by income, education, and unobservable tastes across residential locations or stores. Further analysis is required to understand which factors are most important in explaining why demand varies across socioeconomic groups with equal access.

References

- Aizer, A., J. Currie, P. Simon, and P. Vivier**, “Do low levels of blood lead reduce children’s future test scores?” *NBER Working Paper 22558*, 2016.
- Alcott, H., R. Diamond, and J.P. Dube**, “The geography of poverty and nutrition: Food deserts and food choices in the United States,” *Conference draft*, 2015.
- Anderson, M.L. and D.A. Matsa**, “Are restaurants really supersizing America?,” *American Economic Journal: Applied Economics*, 2011, pp. 152–188.
- Aussenberg, R.A.**, “SNAP and related nutrition provisions of the 2014 Farm Bill (P.L. 113-79),” *Congressional Research Service Report*, 2014.
- Bertail, Patrice and France Caillavet**, “Fruit and vegetable consumption patterns: A segmentation approach,” *American Journal of Agricultural Economics*, 2008, 90 (3), 827–842.
- Bitler, M. and S.J. Haider**, “An economic view of food deserts in the United States,” *Journal of Policy Analysis and Management*, 2011, 30 (1), 153–176.
- Carlson, A., M. Lino, and T.V. Fungwe**, “The low-cost, moderate-cost, and liberal food plans,” *United States Department of Agriculture, Center for Nutrition Policy and Promotion*, 2007.
- CDC**, “State initiatives supporting healthier food retail: An overview of the national landscape,” 2011.
- Chen, S., R.J. Florax, S. Snyder, and C.C. Miller**, “Obesity and access to chain grocers,” *Economic Geography*, 2010, 86 (4), 431–452.
- Chetty, R. and Lawrence Hendren N. Katz**, “The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment,” *American Economic Review*, 2016, 106 (4).
- , **N. Hilger, E. Saez, D. Schanzenbach, and D. Yagan**, “How does your kindergarten classroom affect your earnings? Evidence from Project STAR,” *Quarterly Journal of Economics*, 2011, 126 (4).
- Courtemanche, C. and A. Carden**, “Supersizing supercenters? The impact of Wal-Mart Supercenters on body mass index and obesity,” *Journal of Urban Economics*, 2011, 69 (2), 165–181.
- Currie, J. and R. Walker**, “Traffic congestion and infant health: Evidence from E-ZPass,” *American Economic Journal: Applied Economics*, 2011, 3 (1), 65–90.
- , **L. Davis, M. Greenstone, and R. Walker**, “Environmental health risks and housing values: Evidence from 1,600 toxic plant openings and closings,” *American Economic Review*, 2015, 105 (2), 678–709.
- , **S. DellaVigna, E. Moretti, and V. Pathania**, “The effect of fast food restaurants on obesity and weight gain,” *American Economic Journal: Economic Policy*, 2010, pp. 32–63.
- Curtis, C.**, “First Lady Michelle Obama on making a difference in cities with food deserts,” *The White House*, 2011.

- Cutler, D.M. and A. Lleras-Muney**, “Understanding differences in health behaviors by education,” *Journal of health economics*, 2010, 29 (1), 1–28.
- Dingel, Jonathan I.**, “The determinants of quality specialization,” *Mimeograph, Columbia University*, 2014.
- Dubois, P., R. Griffith, and A. Nevo**, “Do prices and attributes explain international differences in food purchases?” *American Economic Review*, 2014, 104 (3), 832–67.
- Eid, J., H.G. Overman, D. Puga, and M.A. Turner**, “Fat city: Questioning the relationship between urban sprawl and obesity,” *Journal of Urban Economics*, 2008, 63 (2), 385–404.
- Einav, L., E.S. Liebtag, and A. Nevo**, “On the accuracy of Nielsen Homescan data,” *USDA Economics Research Report*, 2008, (69).
- Elbel, B., A. Moran, L.B. Dixon, K. Kiszko, J. Cantor, C. Abrams, and T. Mijanovich**, “Assessment of a government-subsidized supermarket in a high-need area on household food availability and children’s dietary intakes,” *Public Health Nutrition*, 2015, *FirstView*, 1–10.
- Ellickson, Paul B.**, “Quality competition in retailing: A structural analysis,” *International Journal of Industrial Organization*, 2006, 24 (3), 521–540.
- Fajgelbaum, P., G.M. Grossman, and E. Helpman**, “Income distribution, product quality, and international trade,” *Journal of Political Economy*, August 2011, 119 (4), 721–765.
- Freedman, M. and A. Kuhns**, “Supply-side subsidies to improve food access and dietary outcomes: Evidence from the New Markets Tax Credit,” *Working Paper*, 2016.
- Handbury, J.**, “Are poor cities cheap for everyone? Non-homotheticity and the cost of living across U.S. cities,” *Mimeograph, The Wharton School at the University of Pennsylvania.*, 2013.
- Harding, M. and M. Lovenheim**, “The effect of prices on nutrition: Comparing the impact of product- and nutrient-specific taxes,” *NBER Working Paper 19781*, 2014.
- Helpman, E. and P.R. Krugman**, *Market structure and foreign trade: Increasing returns, imperfect competition, and the international economy*, MIT press, 1985.
- Hottman, Colin**, “Retail Markups, Misallocation, and Store Variety in the US,” *Mimeograph, Columbia University*, 2014.
- Jones, E.**, “An analysis of consumer food shopping behavior using supermarket scanner data: differences by income and location,” *American Journal of Agricultural Economics*, 1997, pp. 1437–1443.
- Larson, N.I., M.T. Story, and M.C. Nelson**, “Neighborhood environments: Disparities in access to healthy foods in the US,” *American Journal of Preventive Medicine*, 2009, 36 (1), 74–81.

- Ludwig, J., L. Sanbonmatsu, L. Gennetian, E. Adam, G.J. Duncan, L.F. Katz, R.C. Kessler, J.R. Kling, S.T. Lindau, R.C. Whitaker, and T.W. McDade**, “Neighborhoods, obesity, and diabetes—A randomized social experiment,” *New England Journal of Medicine*, 2011, 365, 1509–1519.
- Ogden, C.L., M.M. Lamb, M.D. Carroll, and K.M. Flegal**, “Obesity and socioeconomic status in adults: United States, 2005–2008,” *NCHS Data Brief No. 50*, 2010.
- Park, J.L., R.B. Holcomb, K.C. Raper, and O. Capps**, “A demand systems analysis of food commodities by US households segmented by income,” *American Journal of Agricultural Economics*, 1996, 78 (2), 290–300.
- Powell, L.M., S. Slater, D. Mirtcheva, Y. Bao, and F.J. Chaloupka**, “Food store availability and neighborhood characteristics in the United States,” *Preventive Medicine*, 2007, 44 (3), 189–195.
- Rahkovsky, I. and S. Snyder**, “Food choices and store proximity,” 2015, (ERR-195).
- Shaked, A. and J. Sutton**, “Relaxing price competition through product differentiation,” *The Review of Economic Studies*, 1982, pp. 3–13.
- and — , “Natural oligopolies,” *Econometrica*, 1983, pp. 1469–1483.
- Song, H.-J., J. Gittelsohn, M. Kim, S. Suratkar, S. Sharma, and J. Anliker**, “A corner store intervention in a low-income urban community is associated with increased availability and sales of some healthy foods,” *Public Health Nutrition*, 2009, 12 (11), 2060–2067.
- Todd, J.E., L. Mancino, E.S. Leibtag, and C. Tripodo**, “Methodology behind the quarterly food-at-home price database,” *United States Department of Agriculture, Economic Research Service*, 2010.
- USDA**, “Food Access Research Atlas, Documentation, Definitions,” *United States Department of Agriculture*, 2016.
- Volpe, R., A. Okrent, and E. Leibtag**, “The effect of supercenter-format stores on the healthfulness of consumers’ grocery purchases,” *American Journal of Agricultural Economics*, 2013, 95 (3), 568–589.
- and — , *Assessing the healthfulness of consumers’ grocery purchases*, BiblioGov, 2013.
- Waldfogel, J.**, “Preference externalities: An empirical study of who benefits whom in differentiated-product markets,” *The RAND Journal of Economics*, 2003, 34 (3), 557–568.
- Weatherspoon, D., J. Oehmke, A. Dembélé, M. Coleman, T. Satimanon, and L. Weatherspoon**, “Price and expenditure elasticities for fresh fruits in an urban food desert,” *Urban Studies*, 2013, 50 (1), 88–106.
- Zenk, S.N., A.M. Odoms-Young, C. Dallas, E. Hardy, A. Watkins, J. Hoskins-Wroten, and L. Holland**, “‘You have to hunt for the fruits, the vegetables’: Environmental barriers and adaptive strategies to acquire food in a low-income African American neighborhood,” *Health Education & Behavior*, 2011, p. 1090198110372877.

For Online Publication

A Data Appendix

A.1 Household Consumption

To examine the grocery purchases made by households, we use the Nielsen Homescan data. This dataset is collected by the National Consumer Panel (NCP), a joint venture between Nielsen and IRI, and provided by Nielsen through the USDA. As mentioned in Section 2, the Homescan data contains transaction-level purchase information for a representative panel of households across the entire US. While the number of participating households varies from year to year, we observe 114,286 unique households over our sample period (2006 through 2011).⁴⁸ The typical sample household participates in the NCP panel for 20 months.

Households in the panel use a scanner provided by the NCP to record all of their purchases at a wide variety of stores where food is sold. See Harding and Lovenheim (2014) for a detailed description of how households are recruited and encouraged to report purchases on a weekly basis. After scanning the Universal Product Code (UPC) of each item purchased, the household records the date, store name, quantity purchased, and price. To reduce measurement error, Nielsen replaces many of the prices recorded by households with prices reported in store-level data.

Items that do not have a standard UPC are included as “random weight” purchases. For these items, households record the quantities and prices of products in aggregated categories. In 2006 there were 43 such categories, including “candy,” “breads,” “cakes,” “beef,” “chicken,” and “fish.” For 2007 through 2011 the categories were more broadly defined, such as “baked goods,” “meat/poultry/fish,” “candy/nuts/seeds,” “fruits,” and “vegetables.” As we cannot know the precise nutritional information for random weight purchases, we exclude random weight items from our primary analysis. However, using the average nutritional characteristics of products with UPCs in each 2006 random weight category to infer the nutritional content of purchases, we verify that our results in 2006 are robust to the use of nutrient scores that include random weight items. Details of this imputation are provided in Section A.3. Reassuringly, the shares on random weight purchases do not vary systematically with household characteristics. We further note that even when random weight purchases are excluded, we still track fresh items with UPCs such as cartons of fresh strawberries.

⁴⁸The main sample used in our analysis contains 96,849 unique households. The methodology used to arrive at this sample is outlined below.

In addition to household-level purchase activity, the Homescan data also includes yearly information on demographics and residential location for each household in the panel. Households are asked to update their demographic information every year that they are in the sample, so the reported demographics should be relevant for the household's consumption decisions in that year. We use this demographic information to measure two dimensions of SES that are posited to impact a household's consumption decisions: income and education.

Households record their income in one of 19 categories: under 5,000; 5,000-7,999; 8,000-9,999; 10,000-11,999; 12,000-14,999; 15,000-19,999; 20,000-24,999; 25,000-29,999; 30,000-34,999; 35,000-39,999; 40,000-44,999; 45,000-49,999; 50,000-59,999; 60,000-69,999; 70,000-99,999; 100,000-124,999; 125,000-149,999; 150,000-199,999; 200,000+. We limit our analysis to households that have at least one household head working over 30 hours per week and report annual earnings of at least \$8,000. We assign households an income equal to the midpoint of their income category for each bounded category and an income of \$260,000 for the "\$200,000 and above" category. Where noted, we adjust the resulting household income for household size using the OECD equivalence scale. According to this scale, the first adult in the household receives a weight of 1, all other adults receive weights of 0.5, and each child receives a weight of 0.3 (<http://www.oecd.org/eco/growth/OECD-Note-EquivalenceScales.pdf>).

For education, households record the household head's education in one of six categories: grade school, some high school, high school graduate, some college, college graduate, or post-college graduate. Households in which either household head reports only a grade school education are excluded from our analysis (11,059 households are below either this education threshold or the income threshold defined above). We assign each household head a number of years of education assuming that some high school corresponds to 10 years, some college corresponds to 14 years, and post college corresponds to 18 years. For households with two household heads, we use their average years of education.

To create a consistent sample that can be used for all of our analyses, we exclude households with missing tract information (14% of remaining households, 13% of remaining household-month and household-store-month observations). For our household-store level analysis, we also exclude household-store-month observations with missing store information (27% of remaining household-store-month observations). Furthermore, to avoid adding noise by including inconsistent reporting, we only keep household-level observations where at least one purchase with non-zero calories is reported in each of three weeks in a given month (99% of remaining households, 72% of remaining household-month observations, 79% of remaining household-store-month observations). Next, as nutrient scores that are more than twice the distance between the 50th and 90th percentiles from

the median likely reflect measurement error, we further exclude these outliers from our preferred sample (less than 1% of remaining household observations, 7% of remaining household-month observations, 6% of remaining household-store-month observations). Finally, as households who qualify for either WIC, SNAP, or reduced school lunch pricing may have different consumption patterns as a result of institutional details, where noted we exclude households who qualify for various forms of food assistance (29% of remaining households, 28% of remaining household-month observations, 27% of remaining household-store-month observations). In particular, we exclude households who report WIC participation to Nielsen, households with monthly income below the threshold for SNAP based on household size (refer to <http://www.fns.usda.gov/snap/eligibility> for the cut-offs used), and households with annual income below the threshold for free and reduced school lunch based on household size (refer to <http://www.fns.usda.gov/school-meals/income-eligibility-guidelines> for the cut-offs used). As shown in Tables 7 and 9, our results are robust to an alternative definition of outliers and to the exclusion of households who qualify for food assistance. Summary statistics for the main household sample used in our analysis are provided in Table A.1.

One concern with using the Homescan data to examine socioeconomic disparities in consumption is that reporting diligence may vary systematically with household SES. Einav et al. (2008) study the credibility of the self-recorded data in the 2004 Homescan sample. They find that reporting errors in the Homescan data are on the same order of magnitude as those commonly found in earnings and employment-status data, although the reporting errors found in the Homescan sample are more pronounced for higher income and more educated households. It is possible that the incentives offered by Nielsen are too small to encourage wealthy households to consistently report all of their purchases. Across all households, however, Einav et al. (2008) find that purchase locations and quantities are reported more accurately than prices. Our results rely primarily on purchase locations and quantities, although our results are qualitatively consistent when we replicate our analyses using measures based on prices (see Appendix C for results based on recommended expenditure shares).

While representative of the US as a whole, another concern with the Homescan data is that it may not be representative of certain subpopulations. To explore this concern, in Figure A.1 we compare the distribution of household income in the Homescan sample to the distribution of household income in the five-year pooled (2007-2011) American Community Survey (ACS). It is clear from Figure A.1 that the Homescan sample under-represents households at either end of the income distribution. Where possible, we demonstrate that our results are consistent across the entire distribution of household income. That is, our results are neither being driven by the

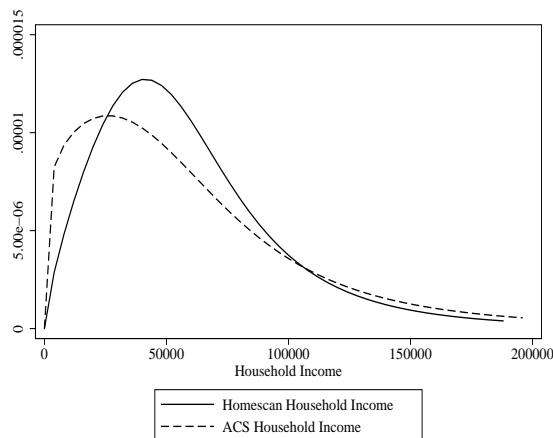
Table A.1: Summary Statistics: Household Demographics

| | By Socioeconomic Status | | | | | | | | |
|----------------------------------|-------------------------|----------------|----------------|---------------|---------------|----------------|---------------|----------------|----------------|
| | All [1] | HI/HE [2] | HI/LE [3] | LI/HE [4] | LI/LE [5] | HI [6] | LI [7] | HE [8] | LE [9] |
| Household level data | | | | | | | | | |
| Number of households | 96,849 | | | | | | | | |
| Mean (std dev) sample months | 25.0 (18.9) | | | | | | | | |
| Household-year level data | | | | | | | | | |
| Number of households | | | | | | | | | |
| 2006 | 34,944 | 11,483 | 5,968 | 6,380 | 11,113 | 17,451 | 17,493 | 17,863 | 17,081 |
| 2007 | 58,976 | 19,603 | 10,248 | 10,433 | 18,692 | 29,851 | 29,125 | 30,036 | 28,940 |
| 2008 | 57,548 | 19,005 | 9,997 | 10,181 | 18,365 | 29,002 | 28,546 | 29,186 | 28,362 |
| 2009 | 55,578 | 18,407 | 9,450 | 10,090 | 17,631 | 27,857 | 27,721 | 28,497 | 27,081 |
| 2010 | 54,880 | 18,384 | 10,255 | 9,066 | 17,175 | 28,639 | 26,241 | 27,450 | 27,430 |
| 2011 | 43,344 | 14,279 | 8,284 | 7,163 | 13,618 | 22,563 | 20,781 | 21,442 | 21,902 |
| Mean (std dev) of: | | | | | | | | | |
| Income (1000s) | | | | | | | | | |
| 2006 | 45.6 (28.3) | 69.2 (30.4) | 57.2 (20.4) | 27.8 (7.8) | 25.3 (8.2) | 65.1 (28.0) | 26.2 (8.2) | 54.4 (31.7) | 36.5 (20.5) |
| 2007 | 45.3 (27.1) | 67.7 (28.3) | 57.1 (18.6) | 27.7 (7.9) | 25.2 (8.3) | 64.1 (25.9) | 26.1 (8.3) | 53.8 (30.1) | 36.5 (20.0) |
| 2008 | 45.1 (26.7) | 67.2 (27.7) | 57.1 (18.6) | 27.9 (8.0) | 25.3 (8.5) | 63.7 (25.4) | 26.2 (8.4) | 53.5 (29.5) | 36.5 (20.0) |
| 2009 | 45.5 (27.9) | 68.2 (29.3) | 58.1 (20.5) | 27.9 (8.1) | 25.1 (8.5) | 64.8 (27.1) | 26.1 (8.5) | 53.9 (30.8) | 36.6 (21.0) |
| 2010 | 44.1 (22.0) | 63.1 (17.5) | 56.4 (15.2) | 27.8 (8.2) | 25.2 (8.4) | 60.7 (17.0) | 26.1 (8.4) | 51.4 (22.4) | 36.9 (19.0) |
| 2011 | 44.2 (22.3) | 63.6 (17.8) | 56.5 (15.5) | 27.6 (8.3) | 25.2 (8.5) | 61.0 (17.3) | 26.0 (8.5) | 51.6 (22.9) | 37.0 (19.2) |
| Education (years) | | | | | | | | | |
| 2006 | 14.1 (1.8) | 15.7 (1.1) | 12.9 (0.9) | 15.3 (1.0) | 12.5 (1.0) | 14.8 (1.7) | 13.5 (1.7) | 15.6 (1.1) | 12.6 (1.0) |
| 2007 | 14.1 (1.8) | 15.7 (1.1) | 12.9 (0.9) | 15.2 (1.0) | 12.5 (1.0) | 14.7 (1.7) | 13.5 (1.6) | 15.5 (1.1) | 12.6 (0.9) |
| 2008 | 14.1 (1.8) | 15.7 (1.1) | 12.9 (0.9) | 15.2 (1.0) | 12.5 (1.0) | 14.7 (1.7) | 13.5 (1.6) | 15.5 (1.1) | 12.6 (0.9) |
| 2009 | 14.1 (1.8) | 15.7 (1.1) | 12.9 (0.9) | 15.2 (1.0) | 12.5 (1.0) | 14.7 (1.7) | 13.5 (1.6) | 15.5 (1.1) | 12.6 (0.9) |
| 2010 | 14.1 (1.8) | 15.7 (1.1) | 12.9 (0.9) | 15.3 (0.9) | 12.5 (1.0) | 14.7 (1.7) | 13.5 (1.6) | 15.6 (1.0) | 12.7 (1.0) |
| 2011 | 14.1 (1.8) | 15.7 (1.1) | 12.9 (0.9) | 15.3 (0.9) | 12.5 (1.0) | 14.7 (1.7) | 13.5 (1.6) | 15.6 (1.0) | 12.7 (1.0) |

Notes: The above figure presents summary statistics of demographics for Homescan panelists included in our main estimation sample. Income and education are residualized from household size fixed effects, average head of household age, a dummy for marital status of household heads, dummies for households with either a female or male household head only, a dummy for the presence of children, and dummies for whether the household reports being white, black, Asian, or Hispanic. See Footnote 12 for a description of how households are separated according to income and education levels. The base sample of households represented here excludes outliers (more than twice the distance between the 50th and 90th percentiles from the median) and includes WIC households, SNAP/CNP households, and households with kids.

center of the distribution, where we observe more households, or by the tails, where our sample of households is more limited.

Figure A.1: Distribution of Household Income: Homescan versus ACS



Notes: The solid line depicts the fitted distribution of household income from the full 2010 Homescan sample; the dashed line depicts the fitted distribution of household income from the 2007-2011 ACS.

A.2 Retail Environments

The Homescan data only provides a limited picture of the retail environments in which households are making their purchase decisions. There are two problems with using the Homescan data to characterize retail environments: First, if no household in the Homescan sample shops at a given store, then we do not observe from the data that this store exists. Second, even if we do observe households shopping in a given store, we only observe the products that they actually purchase, not the full variety of products offered. Because of these limitations, we use two additional datasets, both maintained by Nielsen, to obtain a more comprehensive picture of the retail environments that households face.

In order to observe the full set of stores available to households, we use a sample of the Nielsen TDLinX data provided to us through the USDA. The TDLinX data contains the names and geocoded locations of all food stores in the US.⁴⁹ Our sample contains information on all stores in the grocery, convenience, drug, wholesale club, and mass merchandiser categories. There are 252,884 stores across these five categories. As dollar stores may attract customers from different socioeconomic profiles than other mass merchandisers, we separate dollar stores from mass merchandisers by making a new channel code for mass merchandisers with a sub-channel code description containing “Dollar Store.” Our sample does not include information on stores in the Superette, gas

⁴⁹TDLinX materials state that the data provides “universal coverage and unique codes for every store in retail trade channels” (see http://www.nielsen.com/content/dam/nielsen/en_us/documents/pdf/Fact%20Sheets%20III/Nielsen%20TDLinX.pdf).

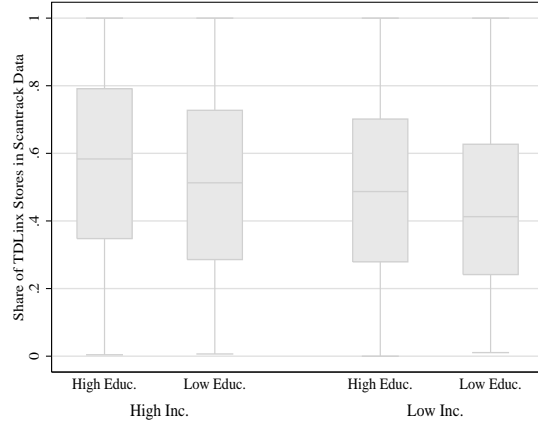
station, liquor store, or cigarette outlet categories. To the extent that access to these primarily non-food stores varies systematically across neighborhoods, our results in Section 3.2 could either overstate or understate differences in access to a wider range of retailers. Our cross-sectional analysis in Section 5.1, however, does not rely on the access disparities that we document using the TDLinx data.

While the TDLinx data tells us about the number and types of stores that households have access to, it provides us with no direct information about product offerings within these stores. To see the full set of food products available at a subset of stores, we use the Nielsen Scantrack data. This data is provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business; refer to <http://research.ChicagoBooth.edu/nielsen> for information on availability and access. The Scantrack data contains weekly sales and quantities at the UPC level. This information is collected by point-of-sale systems that are located in over 30,000 retailers across the US. Stores are divided into four categories in the Scantrack data: grocery, convenience, drug, and mass merchandiser. Unlike the Homescan data, the Scantrack data does not track random weight purchases. It is therefore possible that our results in Section 3.2 are not representative of differences in access to food products without a UPC. We discuss this limitation where relevant. Furthermore, we note that many fresh products, such as cartons of fresh strawberries, have UPCs. Therefore, the Scantrack data still tracks a variety of fresh items.

Despite this detailed information on prices and product offerings, the Scantrack data covers a more limited range of retail outlets than the TDLinx data and only provides us with the county, not the precise geo-coded location, of each store. Where possible, we obtain the geo-coded locations of Scantrack stores by using a concordance between the Homescan, Scantrack, and TDLinx data provided by the USDA. For every store in which a Homescan panelist is observed shopping, the concordance provides both the Scantrack and the TDLinx store bearing the reported name that is closest to the household's residence. To first verify the quality of the concordance, we merge in store-level information provided in both the Scantrack and the TDLinx data. We drop TDLinx-Scantrack matches for which the 3-digit zip codes or county codes do not match across the identified TDLinx and Scantrack stores (17.15% of store-year level observations) or for which the retailer or parent codes do not match the reported store name (40.07% of remaining store-year level observations). Finally, since the methodology used to create the concordance results in multiple TDLinx stores being matched to the same Scantrack store, we select the TDLinx store that is most frequently linked to a given Scantrack store over all years from the remaining TDLinx-Scantrack matches. This methodology allows us to extract the geo-coded locations of 62.2% of the Scantrack stores.

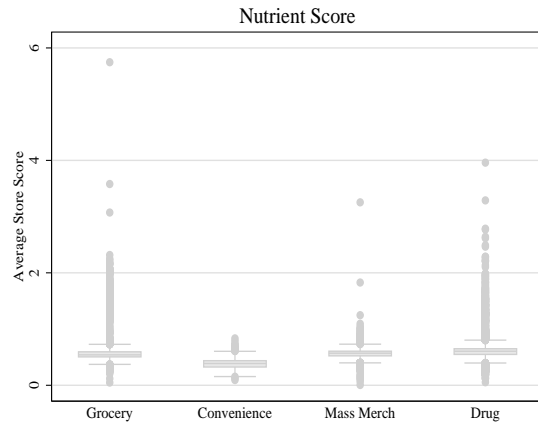
One concern with the Scantrack data is that participation of retailers may systematically vary either across neighborhoods or across store types. As shown in Figure A.2, the average share of TDLinx stores appearing in our geo-coded Scantrack sample is not statistically different across tracts with different socioeconomic profiles. However, there are differences in the percent of TDLinx stores appearing in our geo-coded Scantrack sample across store types. This is in part due to better coverage for certain store types in the Scantrack data and in part due to our geo-coding procedure. While the Scantrack data contains 75% fewer grocery stores than the TDLinx data, we are able to extract the geo-coded location of approximately 90% of Scantrack grocery stores. The numbers are similar for drug stores: the Scantrack data contains 65% fewer drug stores than does the TDLinx data, but we are able to extract the geo-coded locations of 70% of drug stores in the Scantrack sample. The coverage for mass merchandisers and convenience stores is weaker. For convenience stores this is primary due to the fact that the Scantrack sample only contains 2% as many convenience stores as does the TDLinx data, whereas for mass merchandisers it is primarily because we are only able to extract the geo-coded locations of 25% of mass merchandisers in the Scantrack sample. While our concentration indexes only use the TDLinx data and thus speak to broader disparities in access, our estimates of disparities in access to nutritious food speak primarily to differences in access across grocery and drug stores. We note, however, that given Scantrack's poor coverage of convenience stores, it is likely that our estimated disparities actually overstate true differences in access to nutritious products across neighborhoods. As shown in Table 4 and Figure A.3, convenience stores are both more likely to be located in high-SES neighborhoods and to offer a range of products that deviates from recommendations more than other store types.

Figure A.2: Share of TDLinx Stores Appearing in Geo-coded Scantrack Sample Across Tracts



Notes: The figure above presents the average share of TDLinx stores included in the Scantrack sample across tracts with different socioeconomic compositions. Stores are weighted by sales in constructing the shares. See Footnote 14 for a description of how tracts are separated according to income and education levels. These results are for January 2010; they are representative of other months in the Scantrack sample and other years in the TDLinx sample.

Figure A.3: Store Nutrient Scores Across Channels



Notes: The above figure presents distributions of store-level nutrient scores by channel. Stores in the Scantrack data are divided into four channels: grocery, convenience, mass merchandiser, and drug. These results are for January 2010; they are representative of the other months in the Scantrack sample.

A.3 Nutritional Information

To obtain nutritional information for the products purchased by Homescan panelists and sold in Scantrack stores, we use IRI’s nutritional database. The IRI Nutrition Database provides nutritional information for over 700,000 unique UPCs throughout the entire length of our sample. As described in Section 2, the database contains information on the quantity of macro-nutrients and vitamins per serving, serving size in weight, and the number of servings per container at the UPC

level. IRI collects this information directly from product labels. Since product characteristics can change without a change in the product's UPC, IRI revises its database when an updated version of an existing product is received and includes a time stamp of when the change was made. We use a version of the database that includes a snapshot of the market as of July 30th each year. We assume that these product characteristics are relevant for that calendar year.

We merge the IRI database with the Homescan and Scantrack data to uncover the full nutritional profiles of products we observe being purchased by households and sold in stores. These merges are not perfect: only 45% of the UPCs in the Homescan data and 57% of the UPCs in the Scantrack data are in the IRI nutrition database. We impute nutritional information for products not in the IRI data using the average nutritional information for UPCs in the same product module and product group with the same values for all other relevant characteristics, including brand, flavor, form, formula, style, and type. This same procedure is used to impute the nutritional information for random weight purchases in 2006 in the Homescan data. In addition, the nutritional profiles for 87% of UPCs are not available in every year from 2006 to 2011. For such UPCs, we impute the nutritional information in each missing year using the nutritional information for the same UPC in the prior year. If the nutrient profile of a UPC is missing in 2006, we impute the nutritional information using the information from the first year the UPC's nutritional information is available.

To assess the quality of our nutritional imputation, we compare the nutrient scores of bundles with no missing nutritional information to counterfactual nutrient scores in which some of the nutritional information is imputed. While we can also compare true nutritional information to counterfactual, imputed nutritional information on a nutrient by nutrient basis, comparing true and imputed nutrient scores is a more parsimonious way to evaluate the strength of the imputation. To do so, we set the nutritional information to missing for a random sample of 20% of the UPCs in the IRI database. We then use the nutritional information for the remaining 80% of UPCs to impute the nutritional information for the UPCs with "missing" nutritional information. Finally, we randomly draw with replacement 1,000 bundles of 400 products each and compute the nutrient score of the bundle using either the true or the partially-imputed nutritional information. The results of this exercise demonstrate that the imputation works quite well: the average absolute difference between the nutrient scores for a given bundle is only 6.8% of the average nutrient score across all bundles with a standard deviation of 8.7%.

We are not concerned that the imputed nutritional information biases our results for two reasons. First, the percent of purchased UPCs with imputed nutritional information does not vary systematically with household characteristics. Therefore, even if the imputed nutritional information introduces noise, it does so equally across household socioeconomic profiles. Second, all of

our results are robust to using measures of nutritional quality that do not rely on the precise nutritional information of each UPC. As shown in Section C, we obtain the same results if we use a measure of nutritional quality based on recommended expenditure shares for food categories as opposed to recommendations for particular nutrients.

A.4 Driving and Transit Times

To estimate how long it would take a household residing in a given census tract to access local stores, we scrape driving and transit times from Google Maps. In particular, we pull the driving and transit times between the centroid of each census tract and all stores in our TDLinx sample within 40km (these times are current as of April 2015). While coverage on Google Maps is good, it is not perfect: we are able to obtain driving (transit) times for 67% (61%) of store-tract pairs under 40km. Across census tracts, the average share of stores within 40km for which driving (transit) times are non-missing is 63% (37%). Reassuringly, there is no statistically significant difference in the share of non-missing driving and transit times across tracts with different socioeconomic profiles. We further note that our results are robust to the use of distance weights in place of driving or transit time weights when calculating access kernel densities, and our coverage of distance for store-tract pairs is 100%.

A.5 Neighborhood Demographics

While the Homescan data contains demographics for sample households, it only provides us with a limited picture of tract-level demographics. To measure the distribution of income and education in the neighborhoods in which Nielsen households reside and Nielsen stores are located, we use the five-year pooled (2007-2011) American Community Survey (ACS).

While the Nielsen datasets use 2000 census tract boundaries, the 2007-2011 ACS uses census tract boundaries as defined in 2010. We construct demographics from the ACS for the census tract boundaries used by Nielsen as follows. First, using the crosswalk provided by the Census at https://www.census.gov/geo/maps-data/data/tract_rel_download.html, we compute population shares that represent the share of the population from a given 2000 census tract residing in every overlapping 2010 census tract boundary in 2010. That is, letting $pop_{i(y)}^{10}$ denote the population in 2010 in census tract i whose boundary was defined in year y , we compute $w_{i(00),j(10)} = \frac{pop_{i(00)}^{10} \cap pop_{j(10)}^{10}}{pop_{i(00)}^{10}}$ for all tract boundaries such that $pop_{i(00)}^{10} \cap pop_{j(10)}^{10} \neq 0$. For a given 2000 census tract, we then compute the share of college-educated residents as reported in the 2007-2011 ACS by taking a population-weighted average of college-educated shares across 2010 tract

boundaries that overlap with the 2000 tract boundary in question. If, for example, a 2000 census tract was split across two census tracts in 2010, the share of residents with at least a college degree in the 2000 census tract would be given by $educ_{i(00)}^{ACS} = w_{i(00),j(10)} \cdot educ_{j(10)}^{ACS} + w_{i(00),j'(10)} \cdot educ_{j'(10)}^{ACS}$, where $educ_{i(y)}^{ACS}$ denotes the share of college-educated residents as measured in the 2007-2011 ACS in census tract i defined in year y . We compute tract-level median income using a similar procedure. Using the share of residents in each of 16 binned categories as provided in the ACS for every 2010 census tract boundary, we take a population-weighted average to compute analogous shares for 2000 tract boundaries. Median income is then set to the midpoint of the income bin for which at least 50% of the population in the 2000 census tract boundary has income either in or below.

B Supplementary Tables and Figures

B.1 Supplementary Tables

Table A.2: Three Sample Bundles

| Healthy Products | Amount (OZ) | | | Unhealthy Products | Amount (OZ) | | |
|-------------------------|-------------|-------|-----|--------------------|-------------|-------|--------|
| | Good | Mixed | Bad | | Good | Mixed | Bad |
| Cereal- Ready to eat | 12.25 | 6.125 | 0 | Potato Chips | 0 | 5.5 | 11 |
| Russet Potatoes | 160 | 80 | 0 | Milk - 2% Fat | 0 | 64 | 128 |
| Broccoli Florets | 12 | 6 | 0 | American Cheese | 0 | 6 | 12 |
| Carrots | 16 | 8 | 0 | Bacon | 0 | 8 | 16 |
| Kidney Beans | 30 | 15 | 0 | Breakfast Scramble | 0 | 12 | 24 |
| Iceberg Lettuce | 16 | 8 | 0 | Butter Grade AA | 0 | 4 | 8 |
| Strawberries | 16 | 8 | 0 | Coca Cola | 0 | 72 | 144 |
| Orange Juice - No Pulp | 64 | 32 | 0 | Oreo Cookies | 0 | 1.125 | 2.25 |
| Low-fat Yogurt | 36 | 18 | 0 | Mayo | 0 | 1.875 | 3.75 |
| Boneless Chicken Breast | 48 | 24 | 0 | Frozen Pizza | 0 | 56.60 | 113.20 |
| Tuna - Chunk Light | 20 | 10 | 0 | | | | |
| Creamy Peanut Butter | 18 | 9 | 0 | | | | |
| Egg - Grade A Large | 24 | 12 | 0 | | | | |

Notes: The above table shows the composition of three sample bundles. To determine the food products included in each of these bundles, we select among the most widely purchased UPCs in each CNPP food category.

Table A.3: Summary Statistics: Nutritional Quality of Household Purchases

| | By Socioeconomic Status | | | | | | | | |
|---|-------------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|
| | All [1] | HI/HE [2] | HI/LE [3] | LI/HE [4] | LI/LE [5] | HI [6] | LI [7] | HE [8] | LE [9] |
| Household-month level data | | | | | | | | | |
| Number of household-months (millions) | 2.42 | 0.78 | 0.42 | 0.42 | 0.78 | 1.21 | 1.21 | 1.21 | 1.21 |
| Mean (std dev) of: | | | | | | | | | |
| Nutrient score | 1.17 (1.4) | 1.34 (1.6) | 1.10 (1.3) | 1.21 (1.4) | 1.01 (1.2) | 1.26 (1.5) | 1.08 (1.3) | 1.29 (1.5) | 1.04 (1.2) |
| Expenditure score | 6.94 (2.9) | 7.32 (3.2) | 6.83 (2.7) | 7.03 (3.0) | 6.56 (2.6) | 7.15 (3.0) | 6.72 (2.7) | 7.22 (3.1) | 6.65 (2.6) |
| Calories (1000s) | 100.87 (66.8) | 93.24 (63.1) | 101.27 (65.0) | 99.81 (66.6) | 108.87 (70.6) | 96.06 (63.9) | 105.69 (69.3) | 95.55 (64.4) | 106.20 (68.8) |
| Total fat per 100 calories | 4.00 (1.1) | 3.92 (1.1) | 4.04 (1.0) | 3.99 (1.1) | 4.07 (1.0) | 3.96 (1.1) | 4.04 (1.0) | 3.95 (1.1) | 4.06 (1.0) |
| Percent of expenditure on: | | | | | | | | | |
| Fruits and vegetables | 8.54 (7.5) | 9.71 (8.1) | 8.43 (7.2) | 8.62 (7.5) | 7.38 (6.7) | 9.26 (7.9) | 7.81 (7.0) | 9.33 (8.0) | 7.75 (6.9) |
| Soda | 5.94 (7.8) | 5.20 (7.2) | 6.02 (7.7) | 5.72 (7.7) | 6.75 (8.4) | 5.48 (7.4) | 6.39 (8.2) | 5.38 (7.3) | 6.50 (8.2) |
| Household-month-store level data | | | | | | | | | |
| Number of household-store-months (millions) | 5.52 | 1.79 | 0.97 | 0.97 | 1.79 | 2.76 | 2.76 | 2.76 | 2.76 |
| Mean (std dev) of: | | | | | | | | | |
| Nutrient score | 0.68 (0.9) | 0.74 (1.0) | 0.66 (0.9) | 0.68 (0.9) | 0.62 (0.8) | 0.72 (0.9) | 0.64 (0.8) | 0.72 (1.0) | 0.64 (0.8) |
| Expenditure score | 4.76 (3.7) | 5.01 (4.0) | 4.68 (3.6) | 4.83 (3.9) | 4.51 (3.5) | 4.89 (3.9) | 4.62 (3.6) | 4.95 (4.0) | 4.57 (3.5) |
| Calories (1000s) | 27.08 (37.2) | 25.38 (34.8) | 27.77 (37.5) | 26.31 (36.4) | 28.84 (39.6) | 26.22 (35.8) | 27.95 (38.5) | 25.71 (35.4) | 28.46 (38.9) |
| Total fat per 100 calories | 3.70 (2.0) | 3.64 (2.0) | 3.73 (2.0) | 3.70 (2.0) | 3.76 (2.0) | 3.67 (2.0) | 3.74 (2.0) | 3.66 (2.0) | 3.75 (2.0) |
| Percent of expenditure on: | | | | | | | | | |
| Fruits and vegetables | 7.66 (14.6) | 8.45 (15.4) | 7.53 (14.2) | 7.82 (15.0) | 6.85 (13.7) | 8.13 (15.0) | 7.19 (14.2) | 8.23 (15.2) | 7.09 (13.9) |
| Soda | 6.44 (16.5) | 5.52 (15.1) | 6.56 (16.5) | 6.22 (16.3) | 7.40 (17.7) | 5.89 (15.6) | 6.98 (17.3) | 5.77 (15.5) | 7.10 (17.3) |

Notes: The above figure presents summary statistics for the nutritional quality of purchases made by Homescan panelists in our main estimation sample. See Footnote 12 for a description of how households are separated according to income and education levels. The base sample of households represented here excludes outliers (more than twice the distance between the 50th and 90th percentiles from the median) and includes WIC households, SNAP/CNP households, and households with kids.

Table A.4: Correlations between Measures of Nutritional Quality (Household-Month Purchases)

| | Nutrient Score | Expenditure Score | Total Calories | Fat Per Calorie | Soda Exp. Share | Fruit/Veg. Exp. Share |
|-----------------------|----------------|-------------------|----------------|-----------------|-----------------|-----------------------|
| Nutrient Score | 1 | 0.20 | -0.04 | -0.36 | -0.06 | 0.17 |
| Expenditure Score | 0.20 | 1 | 0.13 | -0.03 | -0.16 | 0.56 |
| Total Calories | -0.04 | 0.13 | 1 | 0.08 | 0.03 | -0.09 |
| Fat Per Calorie | -0.36 | -0.03 | 0.08 | 1 | -0.22 | -0.09 |
| Soda Exp. Share | -0.06 | -0.16 | 0.03 | -0.22 | 1 | -0.19 |
| Fruit/Veg. Exp. Share | 0.17 | 0.56 | -0.09 | -0.09 | -0.19 | 1 |

Table A.5: Correlations between Measures of Nutritional Quality (Household-Store-Month Purchases)

| | Nutrient Score | Expenditure Score | Total Calories | Fat Per Calorie | Soda Exp. Share | Fruit/Veg. Exp. Share |
|-----------------------|----------------|-------------------|----------------|-----------------|-----------------|-----------------------|
| Nutrient Score | 1 | 0.26 | 0.19 | -0.22 | -0.02 | 0.15 |
| Expenditure Score | 0.26 | 1 | 0.29 | -0.06 | -0.14 | 0.52 |
| Total Calories | 0.19 | 0.29 | 1 | 0.12 | -0.02 | 0.00 |
| Fat Per Calorie | -0.22 | -0.06 | 0.12 | 1 | -0.28 | -0.15 |
| Soda Exp. Share | -0.02 | -0.14 | -0.02 | -0.28 | 1 | -0.10 |
| Fruit/Veg. Exp. Share | 0.15 | 0.52 | 0.00 | -0.15 | -0.10 | 1 |

Table A.6: Household Characteristics and Nutritional Quality of Purchases: Full Results

| | Ln(Nutrient Score) | | | | | |
|----------------------|----------------------|-----------------------|-----------------------|------------------------|-----------------------|----------------------|
| | Homescan | | | | | FoodAPS |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Ln(Income) | 0.109*** (0.0040) | | 0.0482*** (0.0042) | 0.0303*** (0.0026) | 0.00262 (0.0049) | 0.0661** (0.0026) |
| Education | | 0.0519*** (0.0013) | 0.0456*** (0.0014) | 0.0854*** (0.0026) | 0.0518*** (0.0044) | 0.124*** (0.0250) |
| Ln(Avg. HH Head Age) | -0.0127 (0.011) | 0.00730 (0.011) | 0.0154 (0.011) | 0.00346 (0.0026) | 0.0151** (0.0046) | -0.0267 (0.0210) |
| HH Heads Married | 0.111*** (0.010) | 0.113*** (0.010) | 0.105*** (0.010) | 0.0502*** (0.0049) | 0.0233* (0.0093) | 0.0274 (0.0285) |
| Female HH Head Only | 0.0904*** (0.012) | 0.0433*** (0.012) | 0.0542*** (0.012) | 0.0225*** (0.0051) | 0.00930 (0.0095) | 0.0524 (0.0302) |
| Male HH Head Only | -0.0191 (0.014) | -0.0583*** (0.015) | -0.0566*** (0.014) | -0.0165*** (0.0042) | -0.0130 (0.0077) | -0.0235 (0.0261) |
| Kids Present | 0.107*** (0.0077) | 0.0858*** (0.0076) | 0.0910*** (0.0076) | 0.0375*** (0.0031) | 0.0248*** (0.0062) | 0.101** (0.0321) |
| Race: White | 0.0773*** (0.012) | 0.0818*** (0.012) | 0.0795*** (0.012) | 0.0282*** (0.0042) | 0.0117 (0.0079) | 0.0529 (0.0345) |
| Race: Black | -0.114*** (0.014) | -0.115*** (0.014) | -0.119*** (0.014) | -0.0331*** (0.0038) | -0.0213** (0.0072) | 0.00515 (0.0323) |
| Race: Asian | 0.0126 (0.018) | -0.0151 (0.018) | -0.0228 (0.018) | -0.00350 (0.0028) | 0.00659 (0.0052) | -0.00641 (0.0248) |
| Hispanic | 0.0270* (0.011) | 0.0323** (0.011) | 0.0303** (0.011) | 0.00644** (0.0024) | 0.00432 (0.0043) | 0.0687** (0.0231) |
| Observations | 2,416,384 | 2,416,384 | 2,416,384 | 2,416,384 | 276,893 | 3,934 |
| R^2 | 0.015 | 0.020 | 0.020 | 0.020 | 0.007 | 0.055 |
| Standardized | No | No | No | Yes | Yes | Yes |
| Random Weight | No | No | No | No | Yes | N/A |

Notes: Standard errors are in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Columns (1)-(5) use the Homescan data; column (6) uses the FoodAPS data. Column (5) uses nutrient scores for 2006 only that include random weight purchases. When the Homescan data is used, observations are at the household-month level, standard errors are clustered by household, and year-month fixed effects are included. Observations in the FoodAPS data are at the household level. All regressions include household size dummies and expenditure weights.

Table A.7: Household Characteristics and Nutritional Quality of Purchases: Individual Nutrients

| Healthful Nutrients | | | | | |
|-----------------------|-----------------------|-------------------------|------------------------|-------------------------|-----------------------|
| | Fiber | Iron | Calcium | Vitamin A | Vitamin C |
| Ln(Income) | 0.022*** (0.00063) | 0.0013 (0.00071) | 0.018*** (0.0016) | 0.021*** (0.0012) | 0.013*** (0.00052) |
| Education | 0.022*** (0.00061) | 0.017*** (0.00068) | 0.036*** (0.0015) | 0.024*** (0.00099) | 0.014*** (0.00050) |
| Observations | 2,416,384 | 2,416,384 | 2,416,384 | 2,416,384 | 2,416,384 |
| R^2 | 0.048 | 0.013 | 0.031 | 0.013 | 0.024 |
| Unhealthful Nutrients | | | | | |
| | Total Fat | Sat. Fat | Sodium | Cholesterol | |
| Ln(Income) | -0.015*** (0.0025) | 0.0031*** (0.00068) | 0.0011* (0.00053) | -0.0014*** (0.00016) | |
| Education | -0.041*** (0.0025) | -0.0075*** (0.00067) | -0.011*** (0.00051) | -0.0015*** (0.00016) | |
| Observations | 2,416,384 | 2,416,384 | 2,416,384 | 2,416,384 | |
| R^2 | 0.026 | 0.010 | 0.018 | 0.009 | |

Notes: Standard errors are in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variable in each regression is the normalized deviation of a household's per calorie consumption of a given nutrient in a given month from the recommended consumption. Standard errors are clustered by household. All variables are standardized by the variable's standard deviation. All regressions include year-month fixed effects and controls for household demographics, including household size dummies, average head of household age, a dummy for marital status of household heads, dummies for households with either a female or male household head only, a dummy for the presence of children, and dummies for whether the household reports being white, black, Asian, or Hispanic.

Table A.8: Summary Statistics: Tract Demographics, Store Density, Product Availability, and Pricing

| | By Socioeconomic Status | | | | | | | | |
|---|-------------------------|--------------------|-------------------|-------------------|-------------------|--------------------|-------------------|--------------------|--------------------|
| | All [1] | HI/HE [2] | HI/LE [3] | LI/HE [4] | LI/LE [5] | HI [6] | LI [7] | HE [8] | LE [9] |
| <i>Tract level data</i> | | | | | | | | | |
| Number of tracts with Homescan panelists | 47,021 | 20,293 | 7,350 | 4,358 | 15,020 | 27,643 | 19,378 | 24,651 | 22,370 |
| Mean (std dev) of: | | | | | | | | | |
| Median income | 55,279 (25,587) | 75,304 (24,807) | 55,270 (8,924) | 36,267 (6,960) | 33,746 (7,410) | 69,977 (23,479) | 34,313 (7,386) | 68,403 (27,146) | 40,818 (12,854) |
| College-educated share | 0.27 (0.17) | 0.41 (0.15) | 0.16 (0.04) | 0.32 (0.11) | 0.12 (0.05) | 0.34 (0.17) | 0.17 (0.11) | 0.39 (0.15) | 0.13 (0.05) |
| Number of tracts with stores within 40km: | | | | | | | | | |
| All tracts | 46,990 | 20,270 | 7,355 | 4,353 | 15,012 | 27,625 | 19,365 | 24,623 | 22,367 |
| Tracts with driving times | 36,951 | 15,921 | 5,852 | 3,395 | 11,783 | 21,773 | 15,178 | 19,316 | 17,635 |
| Tracts with transit times | 24,895 | 13,000 | 3,170 | 2,463 | 6,262 | 16,170 | 8,725 | 15,463 | 9,432 |
| Mean (std dev) in 2011 of: | | | | | | | | | |
| Store concentration (Weight, bandwidth): | | | | | | | | | |
| Driving time, 10 min | 89.9 (140) | 101.1 (144) | 64.8 (106) | 106.9 (169) | 82.2 (138) | 91.4 (135) | 87.7 (146) | 102.2 (148) | 76.4 (128) |
| Driving time, 5 min | 18.7 (30) | 18.4 (28) | 12.7 (21) | 25.8 (39) | 19.9 (33) | 16.9 (26) | 21.2 (34) | 19.7 (30) | 17.5 (29) |
| Distance, 3km | 25.0 (59) | 26.0 (60) | 14.7 (35) | 37.0 (78) | 25.3 (61) | 23.0 (55) | 27.9 (65) | 28.0 (64) | 21.8 (54) |
| Distance, 5km | 57.5 (134) | 63.8 (142) | 35.4 (87) | 79.0 (170) | 53.8 (130) | 56.2 (130) | 59.4 (140) | 66.5 (147) | 47.7 (118) |
| Transit time, 30 min | 112 (235) | 108 (240) | 85 (161) | 137 (289) | 122 (233) | 104 (227) | 127 (250) | 113 (248) | 110 (212) |
| Transit time, 45 min | 258 (477) | 264 (486) | 211 (369) | 280 (559) | 262 (472) | 253 (465) | 267 (498) | 266 (498) | 245 (441) |
| Distance (km) to Nearest: | | | | | | | | | |
| Store | 1.46 (2.2) | 1.26 (1.7) | 2.21 (2.8) | 0.87 (1.9) | 1.52 (2.5) | 1.51 (2.1) | 1.37 (2.4) | 1.19 (1.7) | 1.75 (2.6) |
| Grocery store | 2.64 (3.7) | 2.13 (2.7) | 4.05 (4.6) | 1.56 (2.9) | 2.95 (4.4) | 2.64 (3.4) | 2.64 (4.1) | 2.03 (2.8) | 3.31 (4.5) |
| Healthy grocery store | 2.82 (4.0) | 2.27 (2.8) | 4.28 (4.8) | 1.73 (3.3) | 3.16 (4.7) | 2.80 (3.6) | 2.84 (4.5) | 2.17 (2.9) | 3.53 (4.8) |
| Total number of stores within: | | | | | | | | | |
| 0.5km | 1.37 (4.3) | 1.11 (4.1) | 0.69 (2.4) | 2.46 (6.2) | 1.72 (4.7) | 1.00 (3.7) | 1.89 (5.1) | 1.35 (4.5) | 1.38 (4.1) |
| 0.5 to 1km | 4.00 (10.0) | 3.50 (9.7) | 2.33 (5.8) | 6.60 (13.7) | 4.74 (10.6) | 3.19 (8.8) | 5.16 (11.4) | 4.04 (10.6) | 3.95 (9.3) |
| 1 to 2km | 13.1 (32) | 12.4 (30) | 7.6 (17) | 21.1 (43) | 14.5 (35) | 11.1 (27) | 16.0 (37) | 13.9 (33) | 12.3 (30) |
| 2 to 4km | 42.4 (104) | 44.2 (105) | 24.8 (61) | 62.6 (135) | 42.6 (108) | 39.0 (96) | 47.1 (115) | 47.5 (111) | 36.7 (95) |
| 4 to 8km | 130 (317) | 150 (346) | 80 (210) | 172 (390) | 116 (292) | 131 (317) | 129 (318) | 154 (354) | 104 (268) |
| 8 to 16km | 365 (817) | 458 (924) | 259 (628) | 406 (925) | 279 (688) | 405 (859) | 308 (749) | 449 (924) | 272 (669) |
| 16 to 32km | 824 (1552) | 1,122 (1768) | 655 (1314) | 675 (1409) | 548 (1295) | 998 (1672) | 577 (1323) | 1,043 (1719) | 583 (1302) |
| Average vicinity store: | | | | | | | | | |
| Nutrient score (availability) | 0.61 (0.1) | 0.63 (0.1) | 0.61 (0.1) | 0.61 (0.1) | 0.60 (0.1) | 0.62 (0.1) | 0.61 (0.1) | 0.62 (0.1) | 0.60 (0.1) |
| Price index | 1.06 (0.0) | 1.07 (0.0) | 1.06 (0.0) | 1.06 (0.0) | 1.05 (0.0) | 1.07 (0.0) | 1.05 (0.0) | 1.07 (0.0) | 1.05 (0.0) |
| Healthy-to-unhealthy price ratio | 1.00 (0.0) | 1.00 (0.0) | 1.00 (0.0) | 1.00 (0.0) | 1.00 (0.0) | 1.00 (0.0) | 1.00 (0.0) | 1.00 (0.0) | 1.00 (0.0) |

Notes: The above figure presents summary statistics for demographics, store density, product availability, and pricing for the tracts in which Nielsen Homescan panelists reside. Household demographics are taken from the 2007-2011 ACS; store-level information is taken from the Nielsen TDLinx and Scantrack data in 2011. See Footnote 14 for a description of how tracts are separated according to income and education levels. A grocery store is considered "healthy" if its store-level nutrient score is above the median across all stores.

Table A.9: Summary Statistics: Store-Level Product Availability and Pricing

| | By Socioeconomic Status | | | | | | | | |
|--|-------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| | All [1] | HI/HE [2] | HI/LE [3] | LI/HE [4] | LI/LE [5] | HI [6] | LI [7] | HE [8] | LE [9] |
| Store level data | | | | | | | | | |
| Number of stores | | | | | | | | | |
| TD Linx | 284,152 | 95,235 | 39,156 | 29,996 | 119,765 | 134,391 | 149,761 | 125,231 | 158,921 |
| TD Linx-Kilts merged | 21,713 | 10,784 | 2,385 | 2,548 | 5,996 | 13,169 | 8,544 | 13,332 | 8,381 |
| Mean (std dev) in January 2011 of: | | | | | | | | | |
| Nutrient score (availability) | 0.61 (0.2) | 0.62 (0.2) | 0.60 (0.1) | 0.60 (0.2) | 0.59 (0.2) | 0.62 (0.2) | 0.59 (0.2) | 0.62 (0.2) | 0.59 (0.2) |
| Nutrient score (sales) | 0.79 (0.3) | 0.86 (0.3) | 0.74 (0.2) | 0.76 (0.3) | 0.71 (0.2) | 0.84 (0.3) | 0.73 (0.2) | 0.84 (0.3) | 0.72 (0.2) |
| Expenditure score (availability) | 6.77 (2.1) | 6.96 (2.1) | 6.63 (2.1) | 6.64 (2.1) | 6.55 (2.1) | 6.90 (2.1) | 6.58 (2.1) | 6.90 (2.1) | 6.57 (2.1) |
| Expenditure score (sales) | 6.87 (3.4) | 7.22 (3.6) | 6.49 (3.2) | 6.76 (3.3) | 6.44 (3.1) | 7.09 (3.5) | 6.54 (3.2) | 7.13 (3.5) | 6.46 (3.2) |
| Price indexes | | | | | | | | | |
| Aggregate price index | 1.059 (0.07) | 1.064 (0.08) | 1.056 (0.06) | 1.059 (0.07) | 1.049 (0.06) | 1.063 (0.07) | 1.052 (0.06) | 1.063 (0.07) | 1.051 (0.06) |
| Healthy-to-unhealthy ratio (all products) | 1.001 (0.04) | 1.003 (0.04) | 1.000 (0.04) | 1.003 (0.04) | 0.998 (0.05) | 1.002 (0.04) | 1.000 (0.05) | 1.003 (0.04) | 0.999 (0.05) |
| Healthy-to-unhealthy ratio (storable products only) | 1.002 (0.04) | 1.002 (0.04) | 1.002 (0.04) | 1.003 (0.04) | 1.000 (0.05) | 1.002 (0.04) | 1.001 (0.05) | 1.002 (0.04) | 1.001 (0.04) |

Notes: The above figure presents summary statistics for the number of stores in the Nielsen TDLinx data in January 2011 and the store count, product availability, and pricing for the stores in the merged TDLinx-RMS dataset. See Footnote 14 for a description of how tracts are separated according to income and education levels.

Table A.10: Correlations between Measures of Access to Stores and Nutrition

| | Concentration Index | | | Distance to nearest | | | Average local |
|---------------------------------|---------------------|------------|----------|---------------------|-----------|----------------|---------------|
| | Distance | Tran. Time | Dr. Time | Store | Groc. St. | Healthy G. St. | Nutr. Score |
| Concentration Indexes | | | | | | | |
| Distance, bandwidth = 10km | 1 | 0.95 | 0.95 | -0.25 | -0.26 | -0.26 | 0.81 |
| Transit time, bandwidth = 30min | 0.95 | 1 | 0.92 | -0.24 | -0.25 | -0.25 | 0.79 |
| Driving time, bandwidth = 10min | 0.95 | 0.92 | 1 | -0.30 | -0.31 | -0.32 | 0.75 |
| Distance to Nearest | | | | | | | |
| Store | -0.25 | -0.24 | -0.30 | 1 | 0.73 | 0.70 | -0.14 |
| Grocery store | -0.26 | -0.25 | -0.31 | 0.73 | 1 | 0.95 | -0.15 |
| Healthy grocery store | -0.26 | -0.25 | -0.32 | 0.70 | 0.95 | 1 | -0.17 |
| Average local nutrient score | 0.87 | 0.81 | 0.79 | -0.16 | -0.16 | -0.18 | 1 |

Notes: A grocery store is considered "healthy" if its store-level nutrient score is above the median across all stores.

Table A.11: Nutritional Quality of Household Purchases: Controlling for Access

| | Ln(Nutrient Score) | | | | |
|--------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | Geographic Controls | | | Store Controls | |
| | (1) | (2) | (3) | (4) | (5) |
| Ln(Income) | 0.0683*** (0.0025) | 0.0488*** (0.0025) | 0.0285*** (0.0025) | 0.0437*** (0.0020) | 0.0283*** (0.0019) |
| R^2 | 0.015 | 0.034 | 0.174 | 0.012 | 0.119 |
| Education | 0.0971*** (0.0024) | 0.0885*** (0.0024) | 0.0752*** (0.0026) | 0.0556*** (0.0020) | 0.0476*** (0.0018) |
| R^2 | 0.020 | 0.039 | 0.176 | 0.013 | 0.120 |
| Observations | 2,416,384 | 2,416,384 | 2,416,384 | 5,518,666 | 5,518,666 |
| Access FEs | None | County | Tract | None | Store |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Standard errors are in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. In the first three columns, observations are at the household-month level. In the last two columns, observations are at the household-store-month level. Standard errors are clustered by household. All variables are standardized by the variable's standard deviation. All regressions include year-month fixed effects and controls for household demographics, including household size dummies, average head of household age, a dummy for marital status of household heads, dummies for households with either a female or male household head only, a dummy for the presence of children, and dummies for whether the household reports being white, black, Asian, or Hispanic. Columns (1)-(3) include expenditure weights; columns (4)-(5) include expenditure-share weights.

Table A.13: Nutritional Quality of Food at Home vs. Food Away from Home in FoodAPS

| | Nutrient Scores | | | |
|-----------------------------------|-----------------|---------|---------|---------|
| | HI/HE | HI/LE | LI/HE | LI/LE |
| <i>Food at home</i> | | | | |
| Average | 1.102 | 0.93 | 0.87 | 0.795 |
| Standard error | (1.036) | (0.794) | (0.797) | (0.732) |
| Observations | 769 | 859 | 507 | 1,409 |
| <i>Food away from home</i> | | | | |
| Average | 0.724 | 0.666 | 0.679 | 0.63 |
| Standard error | (0.382) | (0.384) | (0.381) | (0.372) |
| Observations | 838 | 906 | 496 | 1,401 |

Notes: The above table presents raw averages of household nutrient scores in the FoodAPS data for households with different demographic profiles. Results are quantitatively robust to only using household who report both food at home and food away from home.

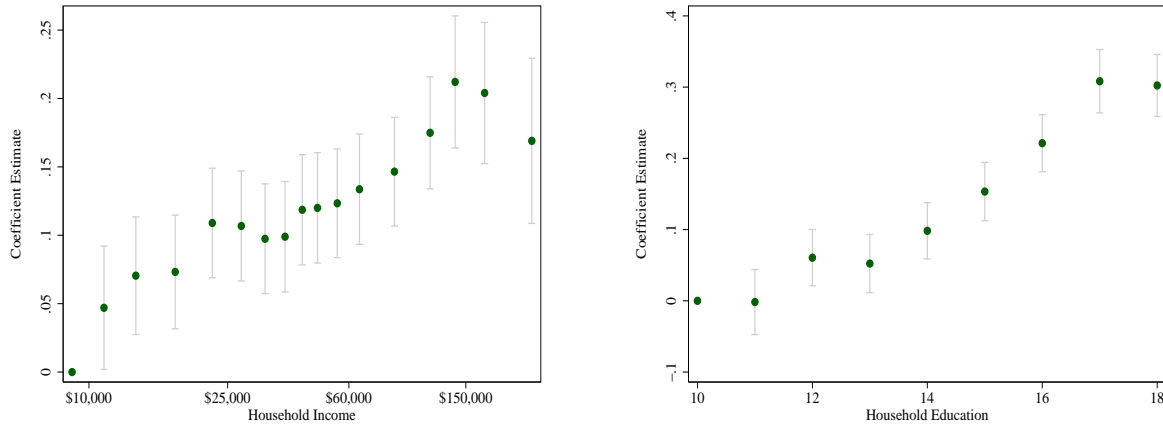
Table A.12: Socioeconomic Disparities in Nutritional Consumption within Stores: Robustness

| | N (1) | Raw | | | Residualized (Store FEs) | | | Diff-in-Diff (%) (8) = [(4) - (7)]/(4) |
|---|-----------|--------------|--------------|-------------------------|--------------------------|--------------|-------------------------|---|
| | | HI/HE (2) | LJ/LE (3) | Diff (4) = (2) - (3) | HI/HE (5) | LJ/LE (6) | Diff (7) = (5) - (6) | |
| Base case | 5,518,668 | 0.74 | 0.62 | 0.12 | 0.68 | 0.59 | 0.09 | 25% |
| <i>Different Cut-offs</i> | | | | | | | | |
| P75/P25 | 5,518,668 | 0.78 | 0.60 | 0.17 | 0.71 | 0.58 | 0.13 | 24% |
| P90/P10 | 5,518,668 | 0.77 | 0.57 | 0.20 | 0.69 | 0.55 | 0.14 | 28% |
| <i>Different Household Samples</i> | | | | | | | | |
| Alternative outliers (excluding >P99 and <P1) | 5,728,431 | 0.82 | 0.66 | 0.16 | 0.74 | 0.62 | 0.12 | 23% |
| Excluding WIC | 5,480,125 | 0.74 | 0.62 | 0.12 | 0.68 | 0.59 | 0.09 | 25% |
| Excluding WIC, below SNAP income | 5,077,992 | 0.75 | 0.63 | 0.12 | 0.69 | 0.60 | 0.09 | 25% |
| Excluding WIC, below SNAP/CNP income | 4,911,686 | 0.74 | 0.63 | 0.12 | 0.68 | 0.60 | 0.09 | 25% |
| Excluding WIC, below SNAP/CNP income, with kids | 4,016,464 | 0.72 | 0.62 | 0.10 | 0.66 | 0.59 | 0.07 | 26% |
| <i>Alternative Measures of Nutrition Quality</i> | | | | | | | | |
| Absolute loss nutrient score | 5,518,668 | 0.42 | 0.37 | 0.05 | 0.40 | 0.36 | 0.04 | 24% |
| Expenditure score | 5,518,668 | 5.01 | 4.51 | 0.50 | 4.74 | 4.29 | 0.45 | 11% |
| Expenditure share on fruits & vegetables | 5,518,668 | 0.08 | 0.07 | 0.02 | 0.08 | 0.07 | 0.01 | 19% |
| Expenditure share on candy | 5,518,668 | 0.22 | 0.23 | -0.01 | 0.22 | 0.23 | 0.00 | 49% |
| Expenditure share on soda | 5,518,668 | 0.06 | 0.07 | -0.02 | 0.06 | 0.07 | -0.02 | 20% |

Notes: The above table presents average raw and residualized household-store-level scores across households with different socioeconomic profiles. Residualized scores are obtained by subtracting store fixed effects estimated in regressions of the log scores against demographic controls, including interacted income and education group fixed effects, year-month fixed effects, and store dummies. See Footnote 12 for a description of how households are separated according to income and education levels. Our "base case" excludes outliers (more than twice the distance between the 50th and 90th percentiles from the median), uses a P50 cutoff, and includes WIC households, SNAP/CNP households, and households with kids.

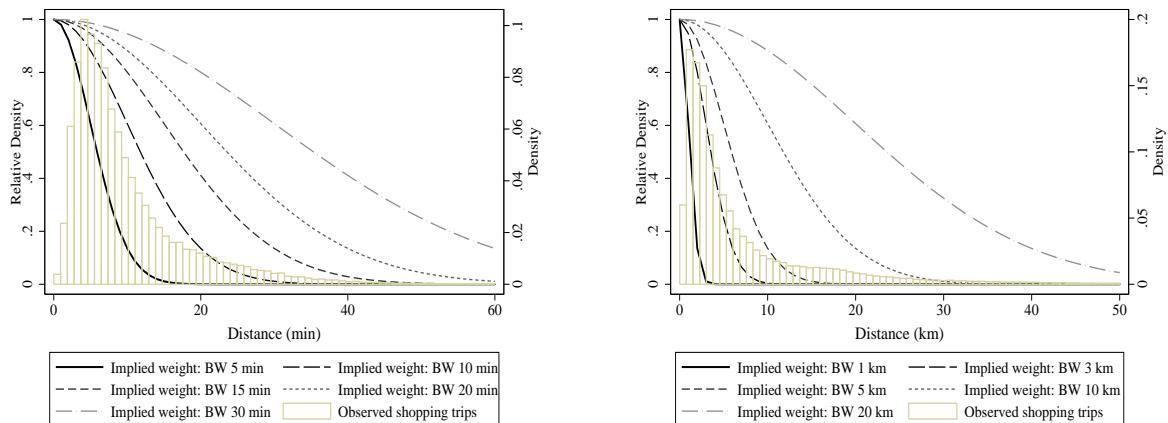
B.2 Supplementary Figures

Figure A.4: Household Nutrient Scores by Income and Education



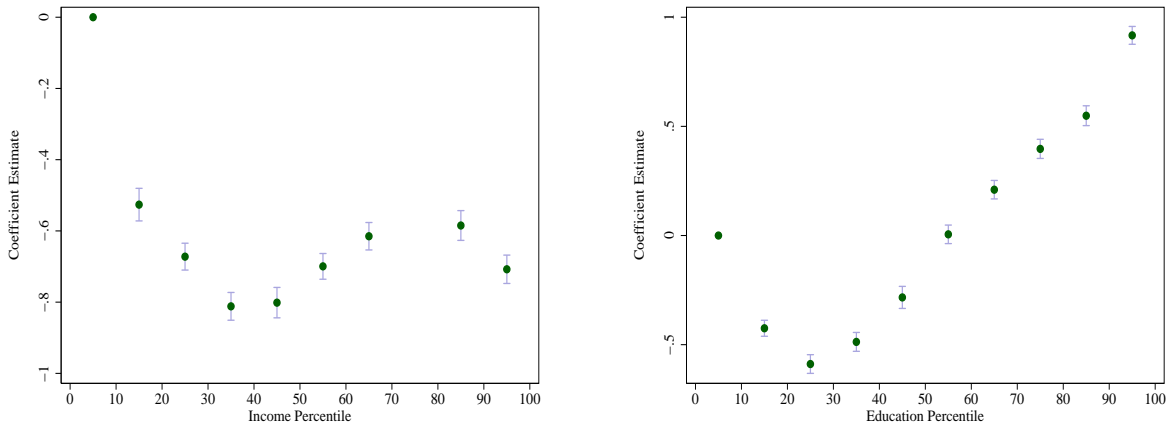
Notes: The above plots depict the association between household demographics and the nutritional quality of household purchases. Observations are at the household-month level. The dots in the left plot are the coefficient estimates on income dummies from an expenditure-weighted regression of log household nutrient scores on income dummies, education, other household demographics, and month-year fixed effects. The dots in the right plot are the coefficient estimates on education dummies from an expenditure-weighted regression of log household nutrient scores on education dummies, log income, other household demographics, and month-year fixed effects. The bars depict 95% confidence intervals.

Figure A.5: Implied Bandwidth Weights vs. Distribution of Household Shopping Trips



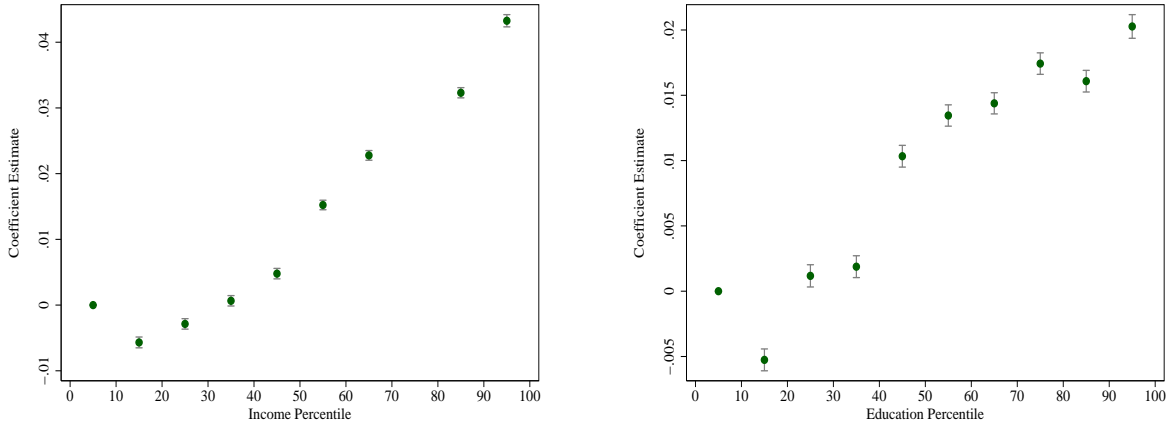
Notes: The above plots depict the relative distribution of observed shopping trips and the implied weights from various Gaussian kernel density bandwidths. The left (right) subplot uses driving times (distances). The bars depict the propensity of households to visit stores at various driving times/distances from their census tract centroid relative to stores at their census tract centroid (i.e. a driving time of 0 minutes and a distance of 0km). Similarly, the dashed lines represent the relative propensities implied by a Gaussian kernel with various bandwidths.

Figure A.6: Store Concentration Indexes by Tract Income and Education



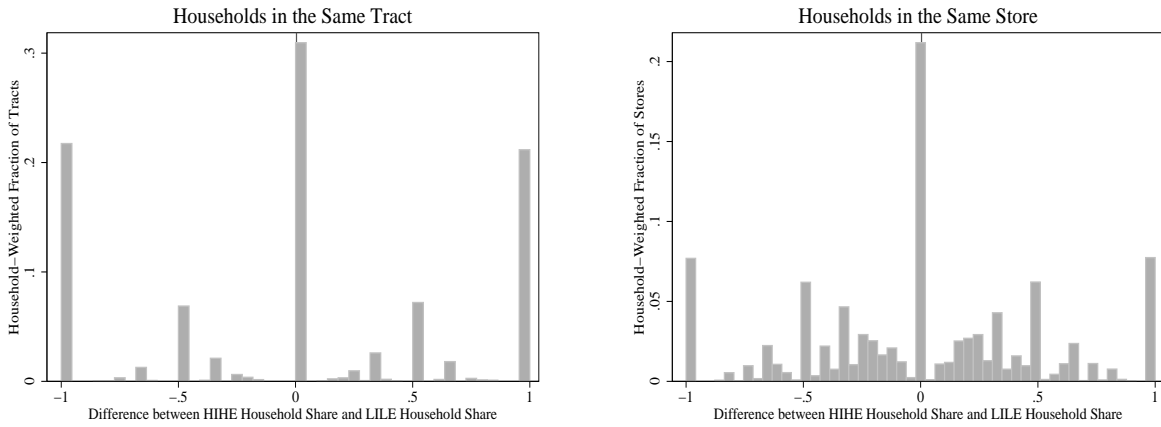
Notes: The above plots depict the association between neighborhood demographics and the concentration of stores. Observations are at the tract-year level. The dots in the left plot are the coefficient estimates on income dummies from a regression of log store concentration indexes on income dummies controlling for log education. The dots in the right plot are the coefficient estimates on education dummies from a regression of log store concentration indexes on education dummies controlling for log income. Income dummies are constructed using tract-level median household income from the ACS; education dummies are constructed using tract-level shares of college-educated residents from the ACS. The bars depict 95% confidence intervals.

Figure A.7: Store Nutrient Scores by Tract Income and Education: Available Products



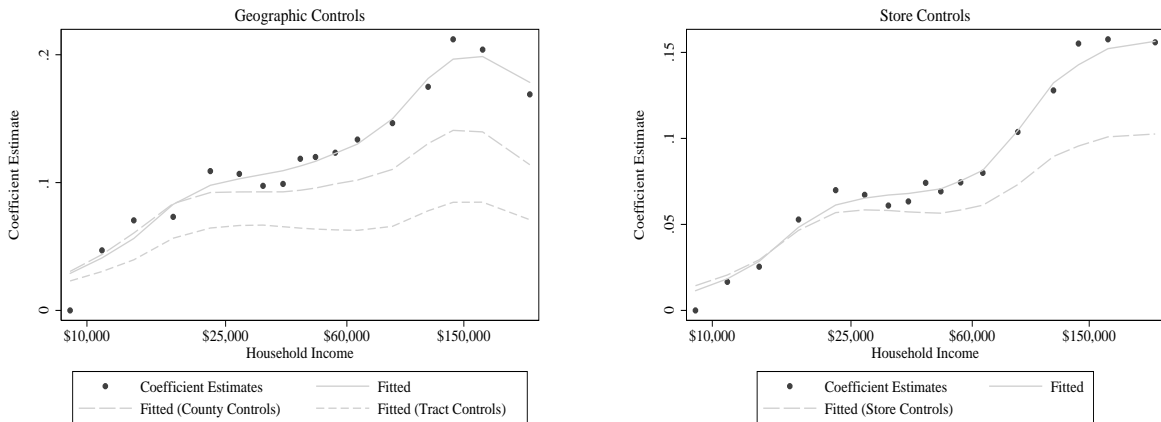
Notes: The above plots depict the association between neighborhood demographics and store nutrient scores. Observations are at the tract-month level. The dots in the left plot are the coefficient estimates on income dummies from a regression of log kernel densities of the nutrient scores of stores surrounding a tract on income dummies and log education. The dots in the right plot are the coefficient estimates on education dummies from a regression of log kernel densities of the nutrient scores of stores surrounding a tract on education dummies and log income. Income dummies are constructed using tract-level median household income from the ACS; education dummies are constructed using tract-level shares of college-educated residents from the ACS. The bars depict 95% confidence intervals.

Figure A.8: Overlap in Household SES in the Same Retail Environment



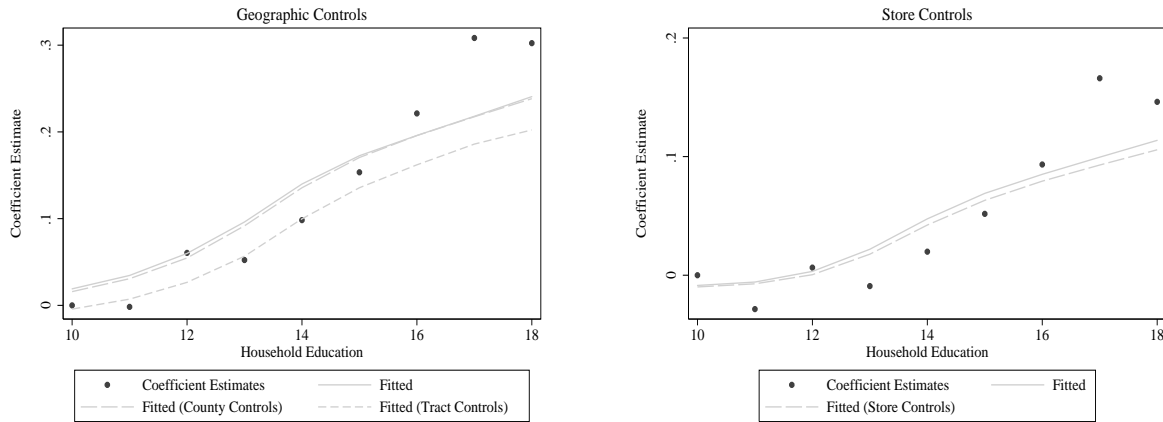
Notes: The plot on the left (right) is a household-weighted histogram of the difference in the share of sample households with above median income and education (HIHE) from the share of sample households with below median income and education (LILE) in the Homescan sample for a given tract (store) in January 2011. Each bar reflects the fraction of sample households represented by tracts (stores) whose HIHE-LILE household share gap is equal to the value on the x-axis.

Figure A.9: Income Effects with Geographic and Store Controls



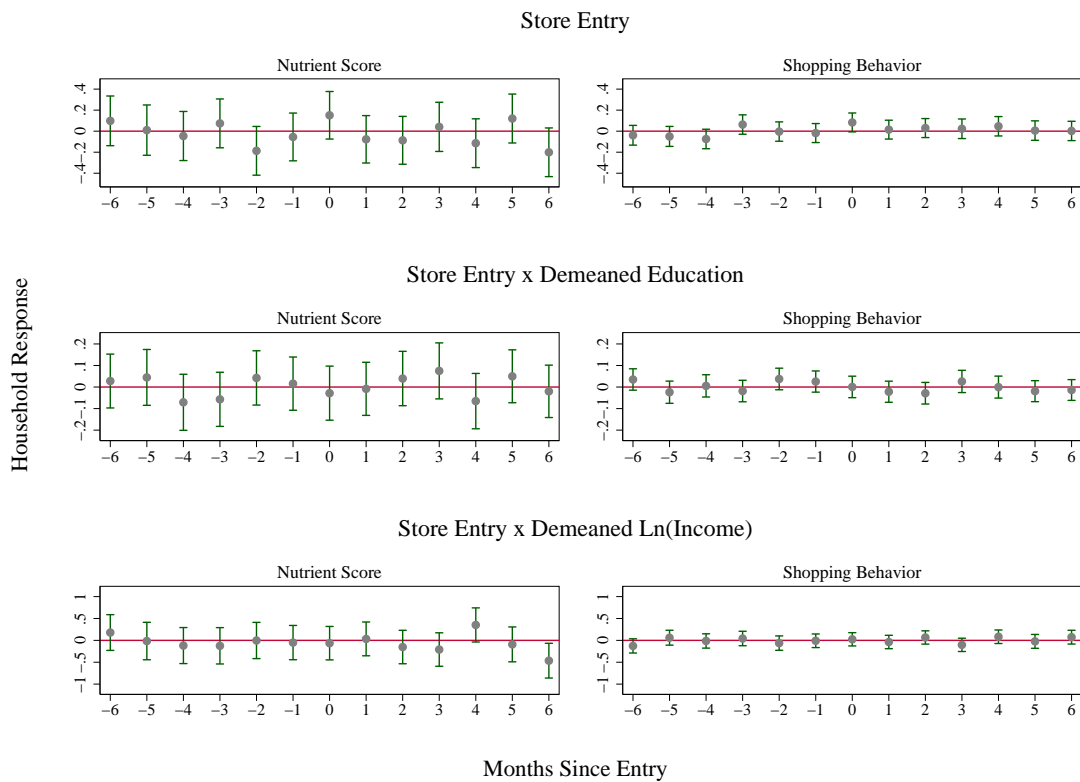
Notes: The above plots depict how the association between income and the nutritional quality of household purchases changes when we control for access using either location or store controls. Observations in the left subplot are at the household-month level; observations in the right subplot are at the household-store-month level. The dots in each plot are the coefficient estimates on income dummies from an expenditure-weighted regression of log household nutrient scores on income dummies, education, other household demographics, and month-year fixed effects. The solid line depicts the smoothed kernel of these estimates. The dashed lines reflect the smoothed kernels of the coefficients on income dummies from the same regression with the addition of either geographic or store controls.

Figure A.10: Education Effects with Geographic and Store Controls



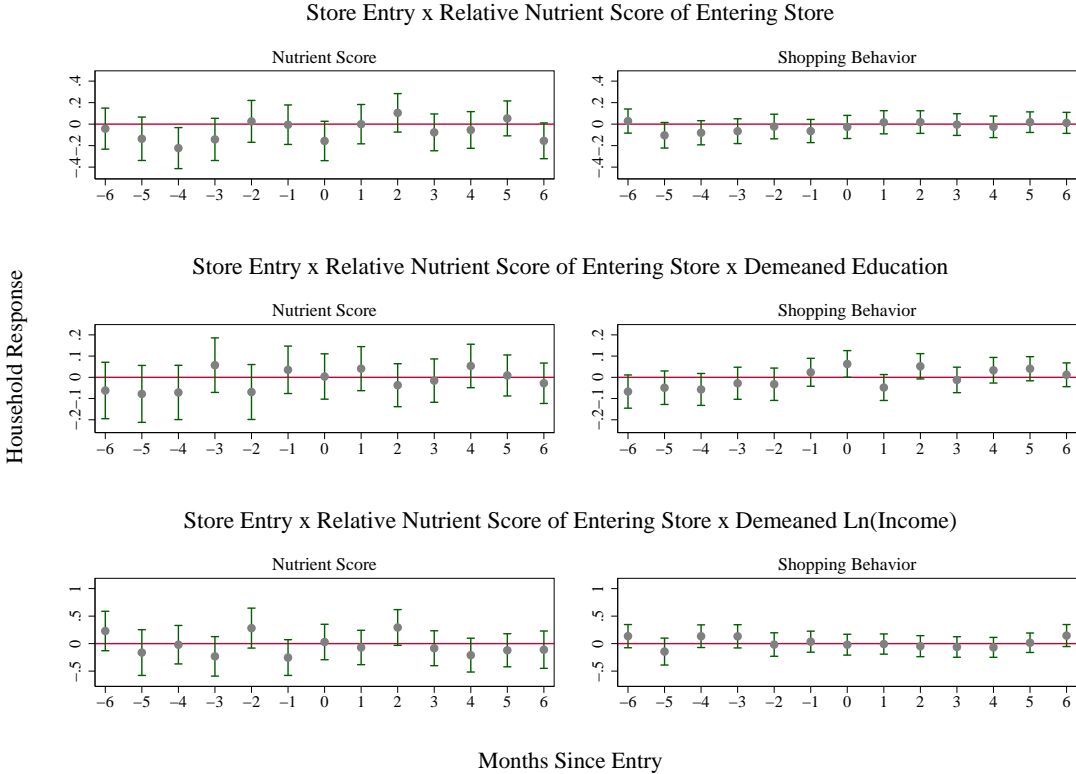
Notes: The above plots depict how the association between education and the nutritional quality of household purchases changes when we control for access using either location or store controls. Observations in the left subplot are at the household-month level; observations in the right subplot are at the household-store-month level. The dots in each plot are the coefficient estimates on education dummies from an expenditure-weighted regression of log household nutrient scores on education dummies, log income, other household demographics, and month-year fixed effects. The solid line depicts the smoothed kernel of these estimates. The dashed lines reflect the smoothed kernels of the coefficients on education dummies from the same regression with the addition of either geographic or store controls.

Figure A.11: Event Study Analysis of Store Entry: Households in Underserved Neighborhoods



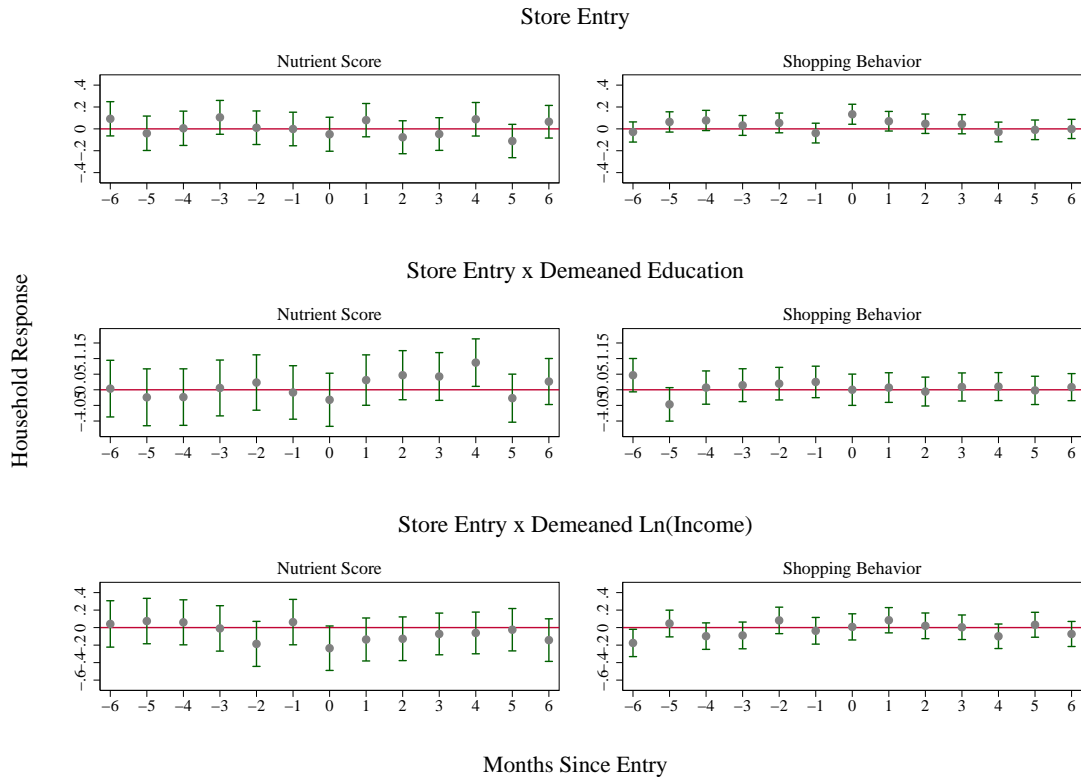
Notes: The above plots display the results from an event study analysis of store entry. The first column depicts the coefficient estimates on dummies for months before, during, and after store entry from a regression of log household-level nutrient scores on household fixed effects, month-year fixed effects, and dummies for each of the six months before, the month of, and the six months after the entry of a grocery store within 2km of a household's census tract centroid. The second column depicts the results from a regression of an indicator for whether the household shopped in a new store in that month on the same independent variables. A tract is defined as being underserved if it falls in the lowest quartile for either its concentration index or its nutrient score density.

Figure A.12: Event Study Analysis of Store Entry: Accounting for Store Quality



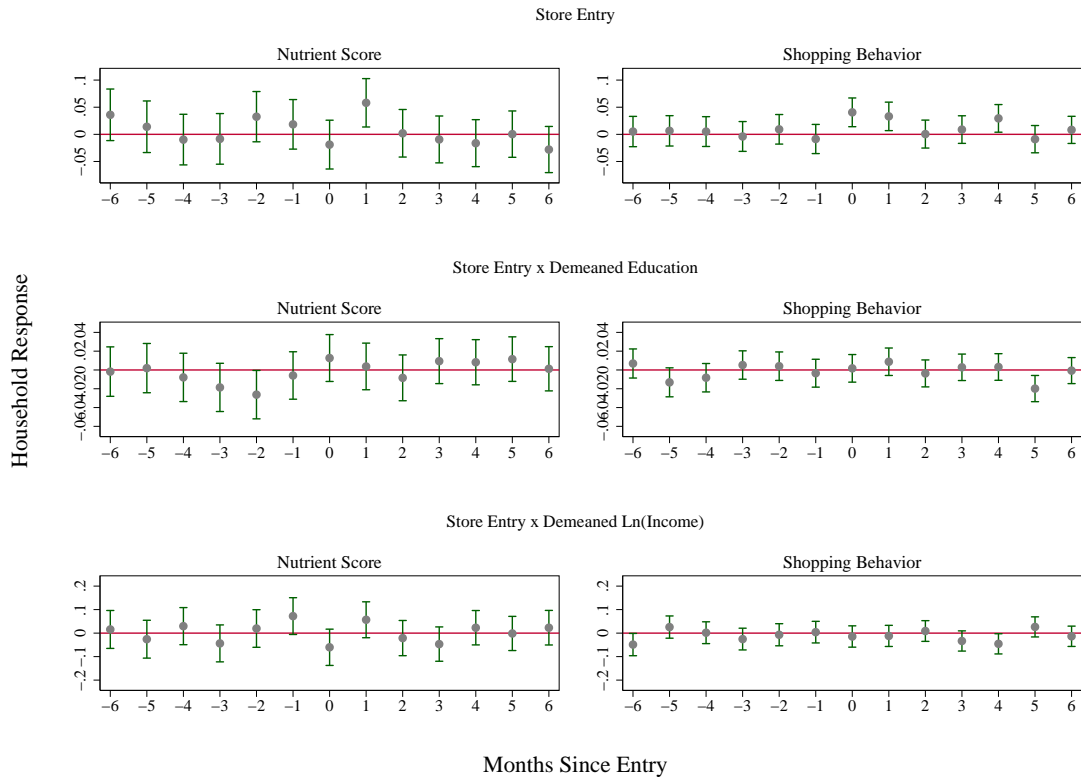
Notes: The above plots display the results from an event study analysis of store entry. The first column depicts the coefficient estimates on dummies for months before, during, and after store entry from a regression of log household-level nutrient scores interacted with the ratio of the nutrient score of the entering store in the month of entry relative to the weighted average nutrient score in the vicinity in the month before entry on household fixed effects, month-year fixed effects, and dummies for each of the six months before, the month of, and the six months after the entry of a grocery store within 2km of a household’s census tract centroid. The second column depicts the results from a regression of an indicator for whether the household shopped in a new store in that month on the same independent variables.

Figure A.13: Event Study Analysis of Store Entry: Within 1km



Notes: The above plots display the results from an event study analysis of store entry. The first column depicts the coefficient estimates on dummies for months before, during, and after store entry from a regression of log household-level nutrient scores on household fixed effects, month-year fixed effects, and dummies for each of the six months before, the month of, and the six months after the entry of a grocery store within 1km of a household's census tract centroid. The second column depicts the results from a regression of an indicator for whether the household shopped in a new store in that month on the same independent variables.

Figure A.14: Event Study Analysis of Store Entry: Within 4km



Notes: The above plots display the results from an event study analysis of store entry. The first column depicts the coefficient estimates on dummies for months before, during, and after store entry from a regression of log household-level nutrient scores on household fixed effects, month-year fixed effects, and dummies for each of the six months before, the month of, and the six months after the entry of a grocery store within 4km of a household's census tract centroid. The second column depicts the results from a regression of an indicator for whether the household shopped in a new store in that month on the same independent variables.

C Alternative Measure of Nutritional Quality: The Expenditure Score

In the following section, we reproduce our results using an alternative measure of nutritional quality. In particular, we introduce a second index, the “expenditure score,” that examines how products purchased by households and offered in stores deviate from recommendations for food group expenditure shares. After defining the expenditure score both for households and stores in Section C.1, we reproduce our main figures and tables using the expenditure score in place of the nutrient score in Section C.2. Note that results for our cross-section and time-series analysis using the expenditure score in place of the nutrient score are provided in robustness Tables 7 and 9, respectively.

C.1 Definition

Household Purchase Quality The USDA Center for Nutrition Policy and Promotion (CNPP) designs food plans for consumers based on recommendations from the Dietary Guidelines for Americans. Our second index, the “expenditure score,” examines how a household’s grocery purchases on each food group deviate from the CNPP’s recommended expenditure shares in the “thrifty food plan” (TFP). The expenditure index follows the measure used by Volpe et al. (2013).

The expenditure score for the grocery purchases recorded by household h in month t is defined as

$$Expenditure\ Score_{ht} = \left[\sum_{c \in C_{Healthful}} (sh_{cht} - sh_{ch}^{TFP})^2 | sh_{cht} < sh_{ch}^{TFP} + \sum_{c \in C_{Unhealthful}} (sh_{cht} - sh_{ch}^{TFP})^2 | sh_{cht} > sh_{ch}^{TFP} \right]^{-1}$$

where c indexes TFP food categories, sh_{cht} denotes the percent of household h ’s grocery expenditures in month t spent on products in category c , and sh_{ch}^{TFP} is the category c expenditure share, also in percent units, that the TFP recommends for a household with the same gender-age profile as household h .

We use the recommended individual expenditure shares from the TFP outlined in Carlson et al. (2007) to construct recommended household expenditure shares, sh_{ch}^{TFP} . The recommended category c expenditure share for each household member i , denoted by sh_{ci}^{TFP} , is determined by his/her age and gender profile. We assign weights to each household member following the OECD equiva-

Table A.14: Healthful and Unhealthful Food Categories

| Healthful | Unhealthful |
|--|--|
| Whole grain products | Non-whole grain breads, cereals, rice, |
| Potato products | pasta, pies, pastries, snacks, and flours |
| Dark green vegetables | Whole milk products |
| Orange vegetables | Cheese |
| Canned and dry beans, lentils, and peas | Beef, pork, veal, lamb, and game |
| Other vegetables | Bacon, sausage, and luncheon meats |
| Whole fruits | Fats and condiments |
| Fruit juices | Soft drinks, sodas, fruit drinks, and ades |
| Reduced fat, skim milk, and low-fat yogurt | Sugars, sweets, and candies |
| Chicken, turkey, and game birds | Soups |
| Eggs and egg mixtures | Frozen or refrigerated entrées |
| Fish and fish products | |
| Nuts, nut butters, and seeds | |

Notes: We determine which TFP food categories are healthful and unhealthful using the recommendations from the Quarterly Food-at-Home Price Database (QFAHPD) indicators for which of 52 food groups are healthful and unhealthful. The QFAHPD categories were created by the USDA (see Todd et al. (2010) for more details).

lence scale and calculate the food expenditure weights as $w_{adult} = \frac{\frac{1+(n_{adult}-1) \times 0.5}{n_{adult}}}{1+(n_{adult}-1) \times 0.5 + n_{children} \times 0.3}$ and $w_{child} = \frac{0.3}{1+(n_{adult}-1) \times 0.5 + n_{children} \times 0.3}$. The recommended category c expenditure share for household h is a weighted average of the recommended category c expenditure shares for each household member, i.e., $sh_{ch}^{TFP} = \sum_i w_i sh_{ci}^{TFP}$. Our results are robust to using the recommended individual expenditure shares from the low-cost, moderate-cost, or liberal food plans instead of those from the TFP.

The TFP provides recommendations for individual-level expenditure shares in 24 food categories. We matched the TFP food groups with Nielsen products using the Quarterly Food-at-Home Price Database (QFAHPD) developed by Todd et al. (2010). In particular, we aggregate the 52 QFAHPD food groups to the 24 TFP food categories using the correspondence created by Volpe and Okrent (2013). In doing so, we find that two TFP food categories, cheese and meat, contain both healthful and unhealthful food groups. Since the vast majority of cheese and meat purchases are of UPCs that fall into the unhealthful QFAHPD food groups, we assume that the aggregate TFP cheese and meat categories are unhealthful. All of our results are robust to assuming that these food categories are instead healthful. Refer to Table A.14 for the full list of healthful and unhealthful food categories that we use.

The expenditure score penalizes households for having a lower-than-recommended expenditure share for healthful food categories ($c \in C_{Healthful}$) and for having a higher-than-recommended expenditure share for unhealthful categories ($c \in C_{Unhealthful}$). As there are no clear guidelines

for which food categories are most important for health, the index construction gives equal weight to all categories. For example, an underconsumption of whole fruits and an overconsumption of frozen or refrigerated entrees are treated the same. We follow Volpe et al. (2013) and take the inverse of the squared loss function so that higher scores are indicative of healthfulness.

The expenditure and nutrient scores consider the healthfulness of consumer purchases from two complementary perspectives, and each measure has its strengths and its weaknesses.⁵⁰ Since consumers choose foods rather than nutrients, the expenditure score is more closely related to consumer demand. Furthermore, expenditures on specific food groups, such as fruits and vegetables, are used by many other studies, and thus the expenditure score is more comparable to previous research.⁵¹ Finally, the expenditure score takes into account other nutrients, such as zinc and potassium, which are not displayed on the nutritional facts panel and are therefore not included in the nutrient score. The expenditure score, on the other hand, distinguishes between products in the same food category, e.g. frozen fish fillets versus fish sticks, that will be missed by the nutrient score. The nutrient score is also not sensitive to systematic variations in the price of foods purchased by different socioeconomic groups. If, for example, low-income and high-income consumers purchase identical quantities of cheese, but high-income consumers purchase more expensive varieties, then for all else equal, expenditure scores will differ by income. The nutrient score, on the other hand, will reflect that both groups have similar diets.⁵²

Available Product Quality The expenditure score for store s in month t can be written as

$$Expenditure\ Score_{st} = \left[\sum_{c \in C_{Healthful}} (sh_{cst} - sh_{ch}^{TFP})^2 | sh_{cst} < sh_{ch}^{TFP} + \sum_{c \in C_{Unhealthful}} (sh_{cst} - sh_{ch}^{TFP})^2 | sh_{cst} > sh_{ch}^{TFP} \right]^{-1}$$

⁵⁰The household-level expenditure and nutrient scores are positively correlated (correlation coefficient of 0.20).

⁵¹The correlation between our expenditure score and expenditure shares on fruits and vegetables is 0.56.

⁵²To address the sensitivity of expenditure scores to prices, we recompute household food category expenditures using the average price per module instead of the actual price paid. Expenditure scores based on this alternative measure of expenditures are comparable to expenditure scores calculated using observed expenditures.

where c again indexes TFP food categories. sh_{cst} is the representative household’s predicted category c expenditure share in store s in month t , calculated as

$$sh_{cst} = \sum_{u \in U_{cst}} \left(\frac{v_{ut}}{\sum_{u \in U_{st}} v_{ut}} \right)$$

Here, U_{cst} is the set of TFP-category c UPCs with positive sales in store s in month t , U_{st} is the set of all UPCs with positive sales in store s in month t , and v_{ut} is the total value of sales of UPC u across all stores in the national Scantrack sample in month t . We look at the distance of this representative household’s category expenditure shares from the TFP’s recommended category expenditure shares for a “typical” household, consisting of a male of age 19-50, a female of age 19-50, one child of age 6-8, and one child of age 9-11. We denote the recommended expenditure share in category c for this modal household by sh_{ch}^{CNPP} .⁵³

C.2 Results

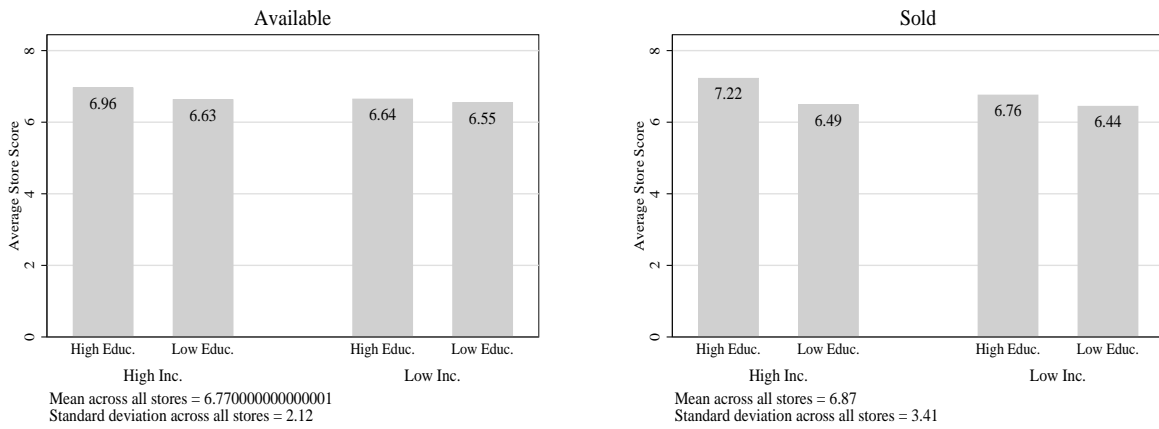
Table A.15: Household Characteristics and Nutritional Quality of Purchases

| | Ln(Expenditure Score) | | | |
|--------------|-----------------------|------------------------|------------------------|-----------------------|
| | (1) | (2) | (3) | (4) |
| Ln(Income) | 0.0518*** (0.0018) | | 0.0301*** (0.0020) | 0.0465*** (0.0030) |
| Education | | 0.0205*** (0.00058) | 0.0166*** (0.00063) | 0.0763*** (0.0029) |
| Observations | 2,416,234 | 2,416,234 | 2,416,234 | 2,416,234 |
| R^2 | 0.066 | 0.069 | 0.071 | 0.071 |
| Standardized | No | No | No | Yes |

Notes: Standard errors are in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Observations are at the household-month level. Standard errors are clustered by household. All regressions include year-month fixed effects and controls for household demographics, including household size dummies, average head of household age, a dummy for marital status of household heads, dummies for households with either a female or male household head only, a dummy for the presence of children, and dummies for whether the household reports being white, black, Asian, or Hispanic.

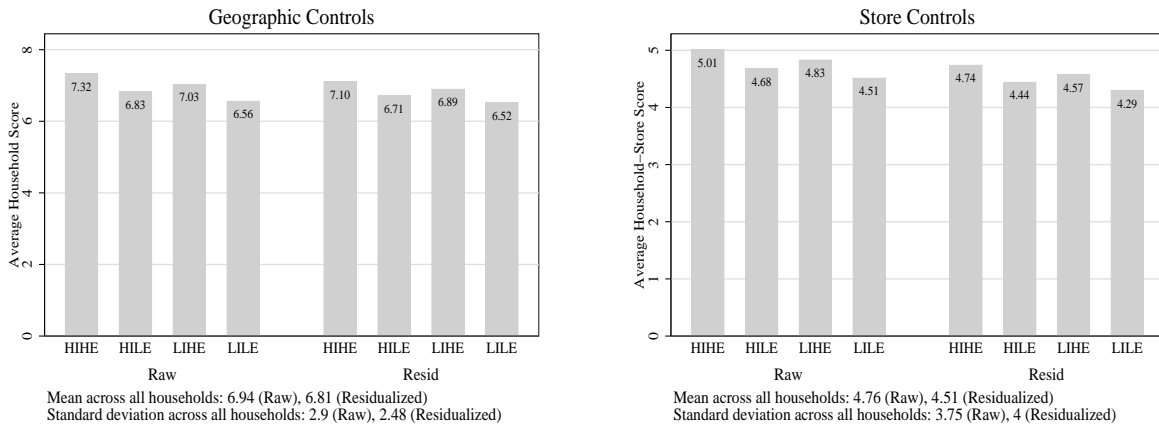
⁵³The store-level expenditure and nutrient scores are positively correlated (correlation coefficient of 0.49).

Figure A.15: Expenditure Scores Across Census Stores: Available versus Sold



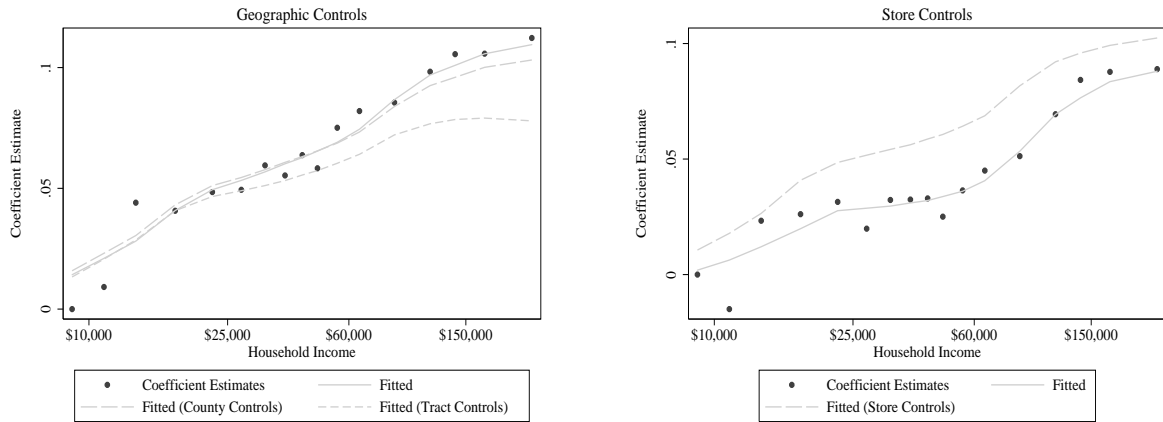
Notes: The above figure presents raw store-level expenditure score averages, computed using either national-sales weights (left) or store-sales weights (right), across census tracts with different socioeconomic compositions. See Footnote 14 for a description of how tracts are separated according to income and education levels. These results are for January 2010; they are representative of other months in the Scantrack sample.

Figure A.16: Residualized Expenditure Scores Across Households



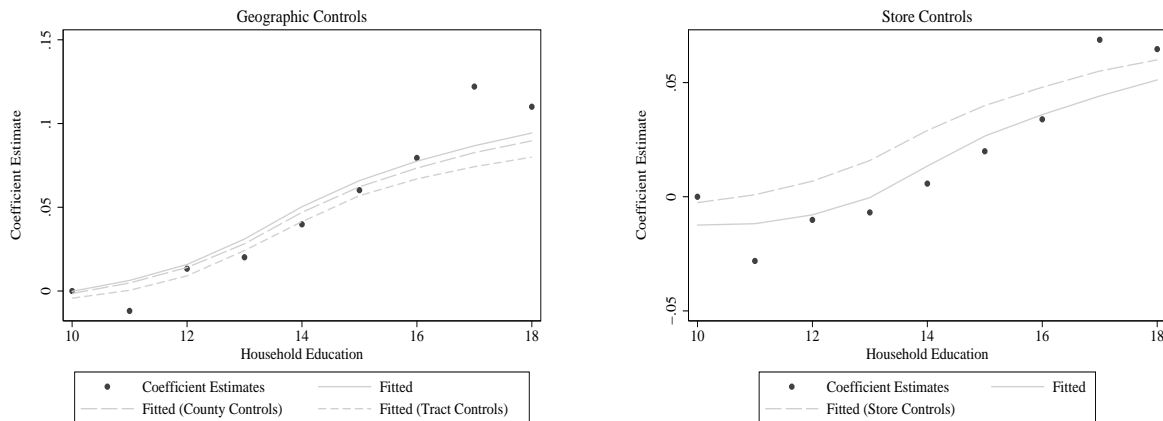
Notes: The subfigure on the left (right) presents average raw and residualized household-level (household-store-level) expenditure scores across households with different socioeconomic profiles. Residualized scores in the subplot on the left (right) are obtained by subtracting census tract (store) fixed effects estimated in regressions of the log scores against demographic controls, including interacted income and education group fixed effects, year-month fixed effects, and census tract (store) dummies. See Footnote 12 for a description of how households are separated according to income and education levels. These results are for January 2010; they are representative of the other months in the Homescan data.

Figure A.17: Income Effects with Geographic and Store Controls



Notes: The above plots depict how the association between income and the nutritional quality of household purchases changes when we control for access using either location or store controls. Observations in the left subplot are at the household-month level; observations in the right subplot are at the household-store-month level. The dots in each plot are the coefficient estimates on income dummies from an expenditure-weighted regression of log household expenditure scores on income dummies, education, other household demographics, and month-year fixed effects. The solid line depicts the smoothed kernel of these estimates. The dashed lines reflect the smoothed kernels of the coefficients on income dummies from the same regression with the addition of either geographic or store controls.

Figure A.18: Education Effects with Geographic and Store Controls



Notes: The above plots depict how the association between education and the nutritional quality of household purchases changes when we control for access using either location or store controls. Observations in the left subplot are at the household-month level; observations in the right subplot are at the household-store-month level. The dots in each plot are the coefficient estimates on education dummies from an expenditure-weighted regression of log household expenditure scores on education dummies, log income, other household demographics, and month-year fixed effects. The solid line depicts the smoothed kernel of these estimates. The dashed lines reflect the smoothed kernels of the coefficients on education dummies from the same regression with the addition of either geographic or store controls.

D Store Inventory and Pricing

D.1 Store-level Nutrient Scores

The nutrient score for store s in month t is given by

$$\begin{aligned} \text{Nutrient Score}_{st} = & \left[\sum_{j \in J_{\text{Healthful}}} \left(\frac{pc_{jst} - pc_j^{DGA}}{pc_j^{DGA}} \right)^2 \Big|_{pc_{jst} < pc_j^{DGA}} \right. \\ & \left. + \sum_{j \in J_{\text{Unhealthful}}} \left(\frac{pc_{jst} - pc_j^{DGA}}{pc_j^{DGA}} \right)^2 \Big|_{pc_{jst} > pc_j^{DGA}} \right]^{-1} \end{aligned}$$

where j again indexes nutrients, $J_{\text{Healthful}}$ and $J_{\text{Unhealthful}}$ are defined as in Section 3.1, and pc_j^{DGA} is the DGA's recommendation for the per calorie consumption of nutrient j . pc_{jst} is the per calorie amount of nutrient j that would be purchased by a representative household in store s in month t , calculated as

$$pc_{jst} = \sum_{u \in U_{st}} \left[\left(\frac{v_{ut}}{\sum_{u \in U_{st}} v_{ut}} \right) pc_{ju} \right]$$

where pc_{ju} is the per calorie amount of nutrient j in UPC u , U_{st} is the set of all UPCs with positive sales in store s in month t , and v_{ut} is the total value of sales of UPC u across all stores in the national Scantrack sample in month t .

D.2 Store Pricing

Since the Scantrack data includes prices, we can directly examine the role of differential pricing in driving consumption patterns across neighborhoods with different socioeconomic profiles. We first study whether stores in high-SES neighborhoods charge higher prices across all food products. We define the aggregate price index for store s in month t as

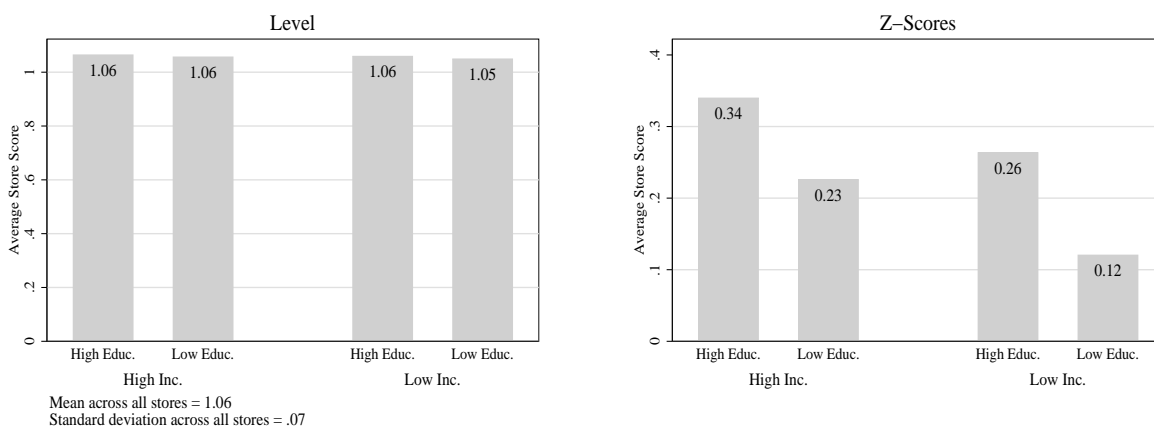
$$P_{st} = \prod_{u \in U_{st}} \left(\frac{p_{ust}}{p_{ut}} \right)^{\frac{v_{ut}}{\sum_{u \in U_{st}} v_{ut}}}$$

where p_{ust} is the sales-weighted average price of UPC u in store s in month t , p_{ut} is the sales-weighted average price of UPC u across all stores in the Scantrack sample in month t , and U_{st} denotes the full set of UPCs sold in store s in month t . This price index summarizes how the average price of each UPC that the store offers compares to the national average price for the UPC.

Figure A.19 shows how these aggregate price indexes vary with tract demographics from the

ACS. Not surprisingly, we see that prices are relatively higher in census tracts with higher levels of income and education. This suggests that low-income households facing tight budget constraints would be even more constrained in their purchases if they shopped in high-SES neighborhoods than they are shopping in low-SES neighborhoods.

Figure A.19: Aggregate Price Indexes Across Census Tracts



Notes: The figure above presents average store-level price indexes across tracts with different socioeconomic compositions. See Footnote 14 for a description of how tracts are separated according to income and education levels. These results are for January 2010; they are representative of the other months in the Scantrack sample.

Even if stores in low-SES neighborhoods offer lower prices in aggregate, they may still incentivize their customers to purchase more unhealthy foods than they would if they lived in a high-SES neighborhood by charging relatively higher prices over the national average for healthy food products than for unhealthy food products. To explore this hypothesis, we use store-level price indexes for healthful and unhealthy products to measure the spatial distribution of the cost of healthy and unhealthy eating. For each store, the healthful (unhealthful) price index summarizes how the average price of each healthful (unhealthful) UPC that the store offers compares to the national average price for that UPC. The price index of healthful products offered in store s in month t is defined as

$$P_{st}^{Healthful} = \prod_{u \in U_{st}^{Healthful}} \left(\frac{p_{ust}}{p_{ut}} \right)^{\frac{v_{ut}}{\sum_{u \in U_{st}^{Healthful}} v_{ut}}}$$

where $U_{st}^{Healthful}$ is the set of all UPCs sold in store s in month t that are classified in a healthful TFP food category. Analogously, the price index of unhealthy products offered in store s in month t is given by

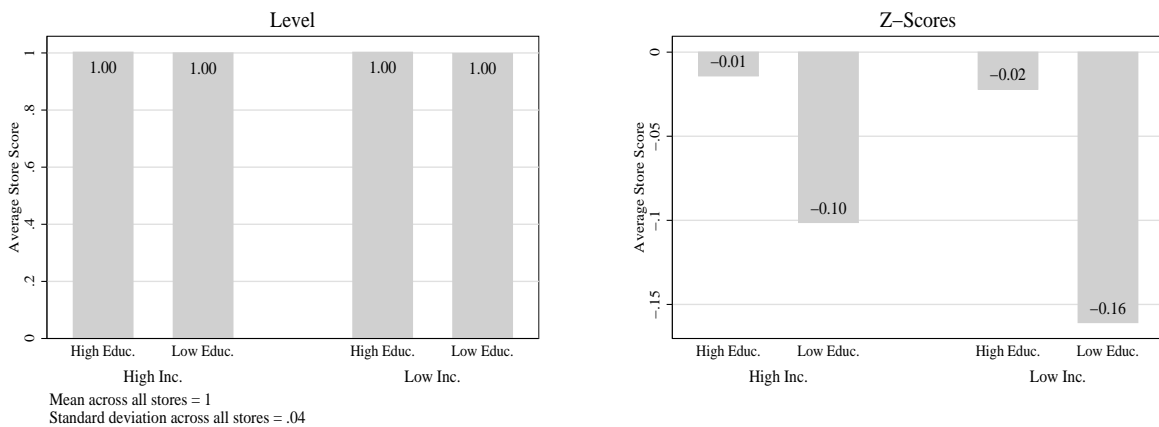
$$P_{st}^{Unhealthful} = \prod_{u \in U_{st}^{Unhealthful}} \left(\frac{p_{ust}}{p_{ut}} \right)^{\frac{v_{ut}}{\sum_{u \in U_{st}^{Unhealthful}} v_{ut}}}$$

where $U_{st}^{Unhealthful}$ is the set of all UPCs sold in store s in month t that are classified in an unhealthful TFP food category.

As our focus is on the accessibility of healthful versus unhealthful foods, we consider the ratio of a store’s healthful-to-unhealthful price indexes, i.e. $\frac{P_{st}^{Healthful}}{P_{st}^{Unhealthful}}$. This ratio, which we refer to as the “relative price index” and denote by $P_{st}^{Relative}$, compares a store’s average markup over national prices for the healthful products it offers to its average markup over national prices for the unhealthful products it offers. A store with a higher relative price index charges relatively more than average for its healthful versus its unhealthful products than a store with a lower relative price index. If differences in relative pricing are to blame for the consumption disparities that we observe, relative price indexes should be higher for stores in neighborhoods with lower levels of income and education.

Figure A.20 shows how relative price indexes vary with tract demographics from the ACS. Perhaps strikingly, we see very little variation in relative price indexes across neighborhoods. If anything, relative price indexes are higher in census tracts with higher levels of income and education. Based on these price patterns alone, we would expect the sales of stores in low-SES neighborhoods to be more, as opposed to less, healthful than the sales of stores in neighborhoods with wealthier and more educated residents.

Figure A.20: Relative Price Indexes Across Census Tracts



Notes: The figure above presents average store-level relative price indexes across tracts with different socioeconomic compositions. See Footnote 14 for a description of how tracts are separated according to income and education levels. These results are for January 2010; they are representative of the other months in the Scantrack sample.

Given the limited variation in price indexes across neighborhoods, we exclude price indexes from our time-series analysis in Section 5.2. However, using time-varying kernel densities of store prices, we can also measure how households in our sample responded to changing prices. Letting P_{slt} denote either the aggregate price index or the relative price index for store s in census tract l in time t , the weighted average price index for census tract l in time t is given by $\frac{1}{\sum_{s=1}^{S_t} w_{sl}} \sum_{s=1}^{S_t} w_{sl} P_{slt}$, where w_{sl} is defined as in Equation (1).

In Table A.16, we replicate our time-series analysis including these time-varying price measures. As before, in column (1) we begin by including these price measures and other continuous access controls with household demographics in place of household fixed effects. As expected, we see that households in neighborhoods with higher prices of healthy relative to unhealthy products purchase bundles with lower nutrient scores. Interestingly, when we include household fixed effects in columns (2) and (3), we see that households in neighborhoods with higher prices actually purchase healthier bundles. Since differences in price indexes are minimal (Figures A.19 and A.20), this responsiveness has little impact on our estimated effects of moving a low-SES household from the average low-SES to the average high-SES neighborhood. That being said, the positive coefficients on the price indexes suggest endogeneity: households with growing tastes for healthy foods live in neighborhoods where the access for such foods is improving. The possibility that changing tastes are driving our time-series results lends further support to our belief that we have identified an upper bound on the potential impact of access-improving policies that ignore local demand.

Table A.16: Response of Nutritional Quality of Household Purchases to Changes in Retail Access

| | Ln(Nutrient Score) | | |
|----------------------------|----------------------|---------------------|---------------------|
| | (1) | (2) | (3) |
| Ln(Income) | 0.043*** (0.006) | | |
| Education | 0.044*** (0.002) | | |
| Ln(Store Concentration) | 0.016*** (0.002) | 0.009 (0.009) | 0.010 (0.009) |
| * $\widetilde{Ln(Inc)}$ | | | -0.001 (0.002) |
| * \widetilde{Educ} | | | -0.001 (0.001) |
| Ln(Local Avg. Store Score) | 0.278*** (0.028) | 0.170*** (0.015) | 0.171*** (0.015) |
| * $\widetilde{Ln(Inc)}$ | | | -0.009 (0.016) |
| * \widetilde{Educ} | | | 0.015*** (0.006) |
| Ln(Local Avg. Price Index) | -0.142 (0.144) | 0.183* (0.102) | 0.193* (0.103) |
| * $\widetilde{Ln(Inc)}$ | | | 0.153 (0.127) |
| * \widetilde{Educ} | | | 0.057 (0.046) |
| Ln(Local Avg. Price Ratio) | -0.300*** (0.089) | 0.201*** (0.075) | 0.218*** (0.079) |
| * $\widetilde{Ln(Inc)}$ | | | -0.003 (0.114) |
| * \widetilde{Educ} | | | 0.041 (0.040) |
| Observations | 244,528 | 244,507 | 244,507 |
| R^2 | 0.038 | 0.001 | 0.001 |
| Demographic Controls | Yes | No | No |
| Household FEs | No | Yes | Yes |

Notes: Standard errors are in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Observations are at the household-year level. Column (1) reports two-way clustered standard errors by tract and year; columns (2) and (3) report Conley Spatial HAC standard errors. A tilde represents that the variable has been demeaned. All regressions include year fixed effects and demographic controls including household size dummies, average head of household age, a dummy for marital status of household heads, dummies for households with either a female or male household head only, a dummy for the presence of children, and dummies for whether the household reports being white, black, Asian, or Hispanic. The sample only includes households that resided in the same census tract throughout their entire participation in the Homescan panel.

E Theoretical Framework with Functional Form Assumptions

E.1 Set-up

There are M locations indexed by l . Each location, l , has an equal population normalized to equal one composed of heterogeneous individuals who differ in their income. We assume that the income distribution of households in each neighborhood is exogenously determined. We also assume that each household is immobile and can shop only at the retail stores in his or her location.

E.1.1 Demand

Household preferences are similar to those in Handbury (2013). Households have a two-tier utility where the upper-tier depends on utility from grocery shopping, U_G , and the consumption of an outside good, z :

$$U = U(U_G(z), z)$$

Outside good expenditure, z , is strictly increasing in income, both by assumption and in the Nielsen Homescan data. In what follows, we refer to z as indexing a households' income level.

Preferences for groceries are given by a nested-CES utility function over a continuum of varieties indexed by u . The nests are defined by the healthfulness of the product u , denoted by $q(u) \in \mathbb{Q}$. Let \mathbb{U}_q denote the set of products of the same healthfulness. A household in location l will select their grocery purchases, $x(u)$, to maximize utility over the products available in location l , \mathbb{U}_l , subject to a budget constraint. The budget constraint is defined by local grocery prices, $p(u, l)$, and the per-capita grocery expenditure, $y - z$, which we normalize to one. That is,

$$\max_{x(u)} U_G(z) = \left[\int_{q \in \mathbb{Q}} \alpha(q, z) \left(\int_{u \in \mathbb{U}_q} x(u)^{\rho_w} du \right)^{\frac{\rho_a}{\rho_w}} \right]^{\frac{1}{\rho_a}} \quad \text{subject to} \quad \sum_{u \in \mathbb{U}_l} p(u, l)x(u) \leq y - z = 1$$

where $\rho_a \in (0, 1)$ reflects the degree of perceived horizontal differentiation between varieties of different nutritional qualities and $\rho_w \in (0, 1)$ reflects the degree of perceived horizontal differentiation between varieties of the same healthfulness. The elasticity of substitution between varieties of different healthfulnesses and between varieties of the same healthfulness can be expressed as $\sigma_a = 1/(1 - \rho_a)$ and $\sigma_w = 1/(1 - \rho_w)$, respectively. We assume $\sigma_w > \sigma_a > 1$. We also assume that varieties are also differentiated vertically by their degree of healthfulness, so the amount of utility a consumer with SES h gets from a unit of consumption of a given variety is scaled up (or down) by their taste for healthfulness, denoted by $\alpha_h(q(u)) > 0$.

The grocery demand of a household with income level z in market l can be characterized by their expenditure share on product u :

$$x(u, l, z) = \left(\frac{p(u, l)}{P(q, l)} \right)^{-\sigma_w} \left(\frac{P(q, l)/\alpha(q(u), z)}{P(l, z)} \right)^{-\sigma_a}$$

where $P(q, l)$ denotes the price index for products of healthfulness q available in market l ($\mathbb{U}_{q,l} = \mathbb{U}_q \cap \mathbb{U}_l$), defined as

$$P(q, l) = \left[\int_{u \in \mathbb{U}_{q,l}} (p(u, l))^{1-\sigma_w} \right]^{\frac{1}{1-\sigma_w}}$$

and $P(l, z)$ denotes the aggregate taste-adjusted price index that consumers with income level z face in market l , defined as

$$P(l, z) = \left[\int_{q \in \mathbb{Q}} \left(\frac{P(q, l)}{\alpha(q, z)} \right)^{1-\sigma_a} \right]^{\frac{1}{1-\sigma_a}}$$

A household total expenditure on all varieties of quality q is given by

$$x(q, l, z) = \left(\frac{P(q, l)/\alpha(q, z)}{P(l, z)} \right)^{-\sigma_a}$$

Assume that there are two types of households, one with high SES and outside good consumption z_H and another with low SES and outside good consumption z_L . The relative expenditure of high-SES to low-SES households on products of the same healthfulness in the same location can be expressed as

$$\frac{\partial x(q, l, z_H)/x(q, l, z_L)}{\partial q} = \sigma_a \left(\frac{\alpha(q, z_H)}{\alpha(q, z_L)} \right)^{\sigma_a} \left(\frac{P(l, z_H)}{P(l, z_L)} \right)^{\sigma_a} \left(\frac{\alpha_1(q, z_H)}{\alpha(q, z_H)} - \frac{\alpha_1(q, z_L)}{\alpha(q, z_L)} \right) \quad (\text{A.1})$$

High-SES households will spend relatively more than low-SES households on healthful products when $\frac{\alpha_1(q, z_H)}{\alpha(q, z_H)} > \frac{\alpha_1(q, z_L)}{\alpha(q, z_L)}$ for all q . We assume that this inequality holds in all cases where tastes vary with SES.

Here we have assumed that preferences vary with SES due to variation in the exogenous taste-shifters. This can be thought of as a reduced-form way of capturing the variation in demand that arises endogenously from complementarities between non-food products and the quality of food products. For example, the results here carry through in a model that instead uses the nested-logit demand system from Fajgelbaum et al. (2011) and assumes that high- and low-SES households

earn different incomes. In that model, the differences in consumption arise endogenously due to a complementarity between the quality of the differentiated food product purchased and the quantity of a homogeneous outside good. We choose to use the nested-CES model above because it allows for us to turn off the non-homotheticity in demand, in order to demonstrate how the observed differences in demand across high- and low-SES households can be generated by supply-side mechanisms alone. The Fajgelbaum et al. (2011) nested-logit model is a variant of the vertical differentiation model from Shaked and Sutton (1982, 1983).

In the classic models of vertical differentiation, variation in the demand for quality is isomorphic with variation in households' price sensitivities, which would generate your more standard "income effect" (where households with lower incomes purchase lower quality products because they cost less). Here, however, the α parameters that govern demand for quality are different to the σ parameters that govern households' price elasticities. We could, therefore, allow for households' demand for quality and price sensitivities to vary with their income or SES as in Handbury (2013). The results below follow through in an extension of this model where the key substitution elasticity governing how prices influence how households allocate expenditure across healthy and unhealthy products, σ_a , varies with income. In this case, the derivative in equation (A.1) above becomes:

$$\frac{\partial x(q, l, z_H)/x(q, l, z_L)}{\partial q} = \left(\frac{x(q, l, z_H)}{x(q, l, z_L)} \right) \left\{ \underbrace{\left[(\sigma_a(z_L) - \sigma_a(z_H)) \frac{P_1(q, l)}{P(q, l)} \right]}_{\text{Price Sensitivity}} + \underbrace{\left[\sigma_a(z_H) \left(\frac{\alpha_1(q, z_H)}{\alpha(q, z_H)} \right) - \sigma_a(z_L) \left(\frac{\alpha_1(q, z_L)}{\alpha(q, z_L)} \right) \right]}_{\text{Tastes}} \right\}$$

where there is an extra term related to the difference in the price sensitivities of high- and low-SES households. When high-SES households are less price sensitive in switching across product quality groups, that is, $\sigma_a(z_L) > \sigma_a(z_H)$, and high quality products are relatively more expensive than low quality products, $P_1(q, l) > 0$, then this term will be positive, driving high-SES households to consume relatively more healthful products than low-SES households. The second term is similar to the derivative in equation (A.1), except that each quality elasticity has a z -specific price elasticity coefficient. This term will be positive, driving high-SES households to consume relatively more healthful products, when $\left(\frac{\alpha_1(q, z_H)}{\alpha(q, z_H)} \right) \left(\frac{\alpha_1(q, z_L)}{\alpha(q, z_L)} \right)^{-1} > \frac{\sigma_a(z_H)}{\sigma_a(z_L)}$; that is, when the relative quality elasticity across H and L households is greater than the relative substitution elasticity (which governs the relative degree of price sensitivity). We present the version of the model where only taste

parameters vary with income as this version of the model is more tractable and provides a clearer intuition for the main results.

E.1.2 Supply

In order to distribute x units of a food product of healthfulness q to a neighborhood with a λ_l share of high-SES residents, we assume that a firm must incur a fixed cost f ; a per unit wholesale cost that can vary with product healthfulness, $w(q)$; and a per unit shelf-space cost that can vary with the share of high-SES residents, $s(\lambda_l)$. To reflect higher rents in higher-SES neighborhoods, we assume that shelf-space costs are increasing in the share of high-SES individuals living in the location. We denote the total marginal cost of retail by $c(q, l) = w(q) + s(\lambda_l)$. We assume that there are no economies of scope, so each retailer sells only one variety in any one location l . Taking the behavior of competitors as given, the optimal price charged by a firm producing variety u of healthfulness q in location l is the price that maximizes profits. That is, the firm solves the following problem

$$\max_{p(u, l)} \pi(u, l) = (p(u, l) - c(q, l)) x(u, l) - f$$

where $x(u, l)$ denotes the demand for variety u in location l , with

$$x(u, l) = \lambda_l x(u, l, z_H) + (1 - \lambda_l) x(u, l, z_L)$$

where we have normalized the population in each location to one. For all varieties u of quality q sold in location l , the optimal pricing strategy is a proportional mark-up over marginal cost:

$$p(u, l) = \frac{c(q, l)}{\rho_w}$$

We can use this optimal price to rewrite the price index for quality q in location l as

$$P(q, l) = (N(q, l))^{\frac{1}{1-\sigma_w}} \left(\frac{c(q, l)}{\rho_w} \right) \quad (\text{A.2})$$

where $N(q, l)$ is the number of varieties of healthfulness q distributed to location l . The price index for a household with income level h in location l is

$$P(l, z) = \left[\int_{q \in \mathbb{Q}} \left(\frac{P(q, l)}{\alpha(q, z)} \right)^{1-\sigma_a} \right]^{\frac{1}{1-\sigma_a}} = \frac{1}{\rho_w} \left[\int_{q \in \mathbb{Q}} \left(\frac{(N(q, l))^{\frac{1}{1-\sigma_w}} c(q, l)}{\alpha(q, z)} \right)^{1-\sigma_a} \right]^{\frac{1}{1-\sigma_a}}$$

Therefore, the quantity of sales of any firm selling a variety of healthfulness q in location l is given by

$$x(q, l) = (N(q, l))^{\frac{\sigma_w - \sigma_a}{1 - \sigma_w}} \left(\frac{c(q, l)}{\rho_w} \right)^{-\sigma_a} [\lambda_l (\alpha(q, z_H) P(l, z_H))^{\sigma_a} + (1 - \lambda_l) (\alpha(q, z_L) P(l, z_L))^{\sigma_a}] \quad (\text{A.3})$$

E.1.3 Equilibrium

We assume that there is free entry into retailing, so active firms earn zero profits. This implies that the scale of firm sales in any given market is given by

$$x(q, l) = \frac{f}{c(q, l)} (\sigma_w - 1) \quad (\text{A.4})$$

E.2 Comparative Statics

E.2.1 Equilibrium Pattern of Product Availability and Consumption Across Locations

Taken together, the zero profit condition (Equation (A.4)), the aggregate demand condition (Equation (A.3)), and the healthfulness-location-specific price index (Equation (A.2)) implicitly define the number of varieties of healthfulness q in each location l as a function of the fixed and marginal costs of producing each variety, the local share of households in each socioeconomic class, and the model parameters:

$$N(q, l) = \Gamma \underbrace{(c(q, l))^K}_{\text{Cost}} \underbrace{[\lambda_l (\alpha(q, z_H) P(l, z_H))^{\sigma_a} + (1 - \lambda_l) (\alpha(q, z_L) P(l, z_L))^{\sigma_a}]^{\frac{\sigma_w - 1}{\sigma_w - \sigma_a}}}_{\text{Demand}} \quad (\text{A.5})$$

where $\Gamma = \left[f (\sigma_w - 1) \left(\frac{\sigma_w - 1}{\sigma_w} \right)^{-\sigma_a} \right]^{\frac{1 - \sigma_w}{\sigma_w - \sigma_a}} > 0$ and $K = \frac{(1 - \sigma_w)(\sigma_a - 1)}{\sigma_a} < 0$. Given the distribution of socioeconomic classes across locations and the retail technology, the pattern of product availability is determined by two forces, each reflected by an individual term in the above expression for product availability. The first, labeled *Cost*, reflects the role that costs play in determining the healthfulness distribution in different locations. The second, labeled *Demand*, reflects the role played by differences in tastes across socioeconomic groups combined with differences in the share of socioeconomic classes in each location's population.

We now demonstrate that each of these mechanisms could individually explain the qualitative patterns that we observe in product availability across neighborhoods and purchases across house-

holds. We are interested in showing that the number of healthful, relative to unhealthful, varieties available in a location is increasing in the share of high-SES households in the location (*i.e.*, that $\frac{N(q,l)}{N(q',l)} > \frac{N(q,l')}{N(q',l')}$ for $\lambda > \lambda'$). If tastes are weakly supermodular in quality and household SES, high-SES households will spend at least as much on high-quality food products as low-SES households in the same location.

Therefore, if the healthfulness of available products is increasing in the share of high-SES households in a neighborhood, it follows that high-SES households will spend more on healthful food products. Even if high-SES and low-SES households share the same tastes, all households will spend more on healthful foods in locations where more of these are available. Since high-SES households are, by definition, disproportionately located in high-SES locations, on average high-SES households will spend more on healthful food products.

We start by turning both mechanisms off. That is, we assume that **tastes are identical** across consumers, *i.e.* $\alpha(q, z) = \alpha(q)$ for all z and q , and that **wholesale costs are equal** across products of different healthfulnesses, *i.e.* $w(q) = w$ for all q . If wholesale costs are equal across products, then the healthfulness of the varieties available in each location will be determined by the taste shifter, $\alpha(q)$:

$$N(q, l) = \Gamma(c(l))^K (\alpha(q)P(l))^{\frac{\sigma_a(\sigma_w-1)}{\sigma_w-\sigma_a}} \quad (\text{A.6})$$

Since tastes are assumed to be identical across consumers, the distribution of healthfulness of available varieties will be identical across locations. To see this, note that the relative number of varieties of two healthfulness levels, q and q' , in location l can be written as the ratio of the common taste shifter for varieties of quality q relative to q' . That is,

$$\frac{N(q, l)}{N(q', l)} = \left(\frac{\alpha(q)}{\alpha(q')} \right)^{\frac{\sigma_a(\sigma_w-1)}{\sigma_w-\sigma_a}} \quad (\text{A.7})$$

Since tastes are identical across households and the distribution of healthful products available is identical across locations, Marshallian demand must be also identical across households, regardless of their SES or location.

If we assume that **tastes are identical** (and, for simplicity, do not vary with product quality), *i.e.* $\alpha(q, z) = \alpha(q)$ for all z and q , but allow **wholesale costs to vary** with healthfulness, then the zero profit condition reduces to

$$N(q, l) = \Gamma(c(q, l))^K (\alpha P(l))^{\frac{\sigma_a(\sigma_w-1)}{\sigma_w-\sigma_a}} \quad (\text{A.8})$$

Taking the derivative with respect to healthfulness q and location l and imposing that retail costs are equal to the sum of wholesale and shelf costs, *i.e.*, $c(q, l) = w(q) + s(\lambda_l)$, we see that as long as wholesale costs are increasing in quality and shelf-space costs are increasing in λ_l , the healthfulness- and location-specific variety counts are supermodular in quality, q , and the share of high-SES households, λ_l :

$$\frac{\partial N(q, l)}{\partial q \partial \lambda_l} = \Gamma K (\alpha P(l))^{\frac{\sigma_a(\sigma_w-1)}{\sigma_w-\sigma_a}} \frac{w'(q)s'(\lambda_l)}{(w(q) + s(\lambda_l))^{2-K}} > 0 \text{ for } w'(q), s'(\lambda_l) > 0.$$

This result implies that high-SES households are more likely to live in locations with a greater variety of healthful food products. The ratio of the price of healthful relative to unhealthy food products will be identical across locations, so households in locations with a greater variety of healthful food products available will purchase relatively more of these products. As a result, we expect to see high-SES households spending more on healthful food products, on average, even if they have the same preferences as low-SES households. That is, socioeconomic disparities in access to healthful and unhealthy food products alone can generate socioeconomic disparities in household purchases.

If we instead assume that **the cost functions are identical** across locations, *i.e.*, $c(q, l) = c(q)$ for all l , but allow for **tastes to vary** with SES, the zero profit condition becomes:

$$N(q, l) = \Gamma(c(q))^K [\lambda_l (\alpha(q, z_H)P(l, z_H))^{\sigma_a} + (1 - \lambda_l) (\alpha(q, z_L)P(l, z_L))^{\sigma_a}]^{\frac{\sigma_w-1}{\sigma_w-\sigma_a}} \quad (\text{A.9})$$

To characterize how the quality distribution is determined by demand, we start by considering the simplest case and compare two locations, l and l' , which are populated entirely by high-SES and low-SES consumers, respectively. The ratio of the product counts across the two locations at any given quality level q is given by

$$\frac{N(q, l)}{N(q, l')} = \left(\frac{\alpha(q, z_H)P(l, z_H)}{\alpha(q, z_L)P(l, z_L)} \right)^{\frac{\sigma_a(\sigma_w-1)}{\sigma_w-\sigma_a}} \quad (\text{A.10})$$

since $\lambda_l = 1$ and $\lambda_{l'} = 0$. Taking the derivative of this function with respect to healthfulness we

see that the ratio of varieties available for a given healthfulness level across the two locations will be increasing in healthfulness as long as $\frac{\alpha_1(q, z_H)}{\alpha(q, z_H)} > \frac{\alpha_1(q, z_L)}{\alpha(q, z_L)}$. This is the same condition required for the relative expenditure share of high-SES to low-SES households to be increasing in quality:

$$\frac{\partial \frac{N(q, l)}{N(q, l')}}{\partial q} = A \frac{N(q, l)}{N(q, l')} \left(\frac{\alpha_1(q, z_H)}{\alpha(q, z_H)} - \frac{\alpha_1(q, z_L)}{\alpha(q, z_L)} \right) > 0 \text{ for } \frac{\alpha_1(q, z_H)}{\alpha(q, z_H)} > \frac{\alpha_1(q, z_L)}{\alpha(q, z_L)} \quad (\text{A.11})$$

for $A = \left(\frac{\sigma_a(\sigma_w - 1)}{\sigma_w - \sigma_a} \right) > 0$.

Now, consider two locations with intermediate, but non-equal, shares of high-SES households. When costs are identical across locations, the zero profit condition implies that the scale of firms producing varieties of the same healthfulness is also identical across locations. The number of varieties available at each healthfulness level will be determined solely by demand for products at that healthfulness level. Since demand for healthful varieties is increasing in SES, and all households earn the same income, we must therefore have that locations with more high-SES households can support a greater variety of healthful food products.

E.2.2 Upper Bound for the Role of Access in Generating Cross-Sectional Disparities

We have demonstrated that two separate forces can each individually explain the distribution of product availability and consumption that we observe across locations. The correlation between access and household purchases demonstrated in the previous literature, however, is insufficient to determine the role that differences in access play in driving differences in consumer behavior (or vice versa). In what follows, we show that by comparing the differences in household purchases across locations to those within locations, we can identify an upper bound on the role that access plays in generating these differences. The critical result is that demand alone determines differences in purchases across households of different SES in the same location. From here, we can show that any sorting across locations based on unobservable tastes will imply that the observed differences in purchases across the selected households who live or shop in the same location are, on average, smaller than the differences in purchases that would persist if access was equalized for all households.

Both access and tastes could be at play in generating the socioeconomic disparities that we observe in purchases across households living in different locations. To see this, note that the expenditures of a household with income level z on products of a given healthfulness q are determined both by their taste for that healthfulness $\alpha(q, z)$, and by the price index of products of that

healthfulness in their location:

$$x(q, l, z) = (\alpha(q, z))^{\sigma_a} \left(\frac{P(q, l)}{P(l, z)} \right)^{1-\sigma_a} \quad (\text{A.12})$$

We saw above that high-SES households purchase more healthful food products either because there are more of these products available in the locations where they live and/or because they have a stronger taste for these products. To see this mathematically, note that the average expenditure share of healthfulness q varieties for high-SES relative to low-SES individuals living across two locations, l and l' , is given by

$$\begin{aligned} \frac{x(q, z_H)}{x(q, z_L)} &= \left(\frac{\lambda_l x(q, l, z_H) + \lambda_{l'} x(q, l', z_H)}{(1 - \lambda_l) x(q, l, z_L) + (1 - \lambda_{l'}) x(q, l', z_L)} \right) \left(\frac{2 - \lambda_l - \lambda_{l'}}{\lambda_l + \lambda_{l'}} \right) \\ &= \underbrace{\left(\frac{\alpha(q, z_H)}{\alpha(q, z_L)} \right)^{\sigma_a}}_{\text{Tastes}} \underbrace{\left(\frac{\lambda_l \left(\frac{P(q, l)}{P(l, z_H)} \right)^{1-\sigma_a} + \lambda_{l'} \left(\frac{P(q, l')}{P(l', z_H)} \right)^{1-\sigma_a}}{(1 - \lambda_l) \left(\frac{P(q, l)}{P(l, z_L)} \right)^{1-\sigma_a} + (1 - \lambda_{l'}) \left(\frac{P(q, l')}{P(l', z_L)} \right)^{1-\sigma_a}} \right)}_{\text{Availability}} \left(\frac{2 - \lambda_l - \lambda_{l'}}{\lambda_l + \lambda_{l'}} \right) \end{aligned} \quad (\text{A.13})$$

The first term reflects taste differences alone. The second term reflects differences in access that, as we outlined above, could be the result of either firms catering to local tastes or to supply-side factors, such as the complementarities between healthfulness and local distribution costs proposed above. These differences in local product availability are reflected through the local price indexes, with $P(q, l)$ decreasing in the number of healthfulness q varieties that are available in location l . There are relatively more healthful varieties available in a location l where there are more high-SES individuals, so the local healthfulness q price index will be lower, relative to the overall price index a household faces in a location ($P(l, z_H)$ or $P(l, z_L)$), in high- λ_l locations relative to locations with a lower share of high-SES residents. This correlation implies that the numerator of the availability term is increasing in quality (since $1 - \sigma_a < 0$), whereas the denominator is falling in quality.

This is easy to see in the case where tastes are identical across households:

$$\frac{x(q, z_H)}{x(q, z_L)} = \left(\frac{\lambda_l \left(\frac{P(q, l)}{P(l)} \right)^{1-\sigma_a} + \lambda_{l'} \left(\frac{P(q, l')}{P(l')} \right)^{1-\sigma_a}}{(1 - \lambda_l) \left(\frac{P(q, l)}{P(l)} \right)^{1-\sigma_a} + (1 - \lambda_{l'}) \left(\frac{P(q, l')}{P(l')} \right)^{1-\sigma_a}} \right) \left(\frac{2 - \lambda_l - \lambda_{l'}}{\lambda_l + \lambda_{l'}} \right) \quad (\text{A.14})$$

To the extent that healthful goods are relatively more abundant in locations with many high-SES

individuals, $P(q, l)$ will also be lower in these locations for healthful goods. Since, by definition, more high-SES individuals live in the locations with more abundant healthful goods, they will tend to consume more healthful goods on average across the two locations than low-SES individuals, who are more likely to live in locations with fewer healthful goods available.

If we instead look at the average expenditure share of healthfulness q varieties for high-SES relative to low-SES households in the same location, l , this availability term no longer varies with product quality:

$$\frac{x(q, l, z_H)}{x(q, l, z_L)} = \left(\frac{\alpha(q, z_H)}{\alpha(q, z_L)} \right)^{\sigma_a} \left(\frac{P(l, z_L)}{P(l, z_H)} \right)^{1-\sigma_a} \quad (\text{A.15})$$

Any systematic variation that we observe in the healthfulness consumed by high-SES relative to low-SES households living in the same location must be attributed to tastes alone.

Note that this within-location variation in healthfulness only provides a lower bound for the role of tastes in generating differences in the healthfulness of purchases across socioeconomic groups, because tastes could also explain part (or all) of the differences in the availability of products in locations where these households reside. Further, in the context of the model, the within-location variation in healthfulness also exactly identifies the disparity that would persist were availability to be equalized across all locations at the level observed in location l . This model is highly stylized, so there are various additional reasons why within-location socioeconomic disparities in healthfulness may reflect more than differences in tastes alone. Important factors that the model abstracts from include the mobility of both products and households between locations, unobserved heterogeneity in tastes across households within the same socioeconomic class, and differences in the mobility of households and the availability of products within locations. These biases will tend to lead us to further overestimate the role of product availability in explaining the overall socioeconomic disparities in purchases. Take, for example, unobserved heterogeneity in tastes. Suppose that households sort into retail locations based on tastes. We can reflect this heterogeneity and sorting by allowing the taste coefficients α , to vary with SES and location, such that the tastes for a product with healthfulness q for a household with SES h in location l is denoted $\alpha_l(q, z)$. Under this assumption, we now have that the relative expenditures of high-SES to low-SES households in the same location l can be written:

$$\frac{x(q, l, z_H)}{x(q, l, z_L)} = \left(\frac{\alpha_l(q, z_H)}{\alpha_l(q, z_L)} \right)^{\sigma_a} \left(\frac{P(l, z_L)}{P(l, z_H)} \right)^{1-\sigma_a} \quad (\text{A.16})$$

Under the new assumption that households are spatially sorted by heterogeneous tastes, this relative expenditure no longer exactly identifies the disparity that would persist were availability equalized

across all locations at the level observed in location l . In particular, since $Corr(\alpha_i(q, z_H), \alpha_i(q, z_L)) \geq Corr(\alpha_i(q, z_H), \alpha_{i'}(q, z_L))$ for any two locations l and l' , then $x(q, l, z_H)/x(q, l, z_L) \leq x(q, l, z_H)/x(q, l', z_L)$ for any two locations l and l' . The relative expenditures of high-SES and low-SES residents in the same location therefore provides a lower bound on the true amount of variation that will persist in the full cross-section of households if access were to be equalized across all locations.

E.2.3 Upper Bound for the Role of Changing Access on Consumption Disparities

If we recast locations as markets that are separated by time instead of by space, we can use the model presented above to interpret the changes that we observe in household purchases over time as their retail environments change. Our goal is to estimate the impact that policies to improve access in underserved areas will have on household purchases without any changes in tastes over time. This is unlikely to be the case in the data, however. The observed changes in access are likely to be correlated with unobserved changes in tastes since households sort into neighborhoods that offer consumption amenities that suit their tastes and stores select their product offerings to cater to local tastes. To see this, consider how the average expenditure share of healthfulness q varieties varies for a household of the same SES h between a market l and another market l' . When deriving this expenditure share for Equation (A.12) above, we assumed that tastes do not vary across markets. This is reasonable when thinking about how household expenditures vary across geographic markets in a single time period, but less reasonable when considering how expenditures vary for a given household over time. Extending Equation (A.12) to allow for tastes to vary over time, we can see that the relative expenditures in market l relative to market l' depend on the change in tastes across the two markets as well as the change in availability:

$$\frac{x(q, l, z)}{x(q, l', z)} = \underbrace{\left(\frac{\alpha_l(q, z)}{\alpha_{l'}(q, z)}\right)^{\sigma_a}}_{Tastes} \underbrace{\left(\frac{P(q, l) P(l', z)}{P(q, l') P(l, z)}\right)^{1-\sigma_a}}_{Availability} \quad (\text{A.17})$$

Given the fixed costs of differentiated good production, stores cater to the tastes in a market. Therefore, changes in availability across markets will be correlated with unobserved changes in the prevalent tastes of local residents. While the tastes of any one panelist household might not reflect the prevalent local tastes (a household's tastes may not change or may change in the opposite direction), we expect that the tastes of our sample households are, on average, correlated and covary with local tastes. As a result, we expect that our estimate of the elasticity of household purchases with respect to changes in their retail environment to be subject to an upward omitted variable bias. Therefore, we interpret these elasticities as an upper bound for the true elasticity that we

expect to govern the response of purchases to improved access that is driven by policy as opposed to endogenous firm responses to changes in market fundamentals.