
FRESH FRUIT AND VEGETABLE CONSUMPTION: THE IMPACT OF ACCESS AND VALUE

A PREPRINT

Retsef Levi

Sloan School of Management
MIT
Cambridge, MA 02141
retsef@mit.edu

Elisabeth Paulson

Operations Research Center
MIT
Cambridge, MA 02141
epaulson@mit.edu

Georgia Perakis

Sloan School of Management
MIT
Cambridge, MA 02141
georgiap@mit.edu

September 16, 2020

ABSTRACT

The goal of this paper is to leverage household-level data to improve food-related policies aimed at increasing the consumption of fruits and vegetables (FVs) among low-income households. Currently, several interventions target areas where residents have limited physical access to grocery stores, many focusing directly on increasing their access to healthy food. However, despite growing attention being placed on access at both federal and local levels, there is surprisingly little agreement about the efficacy of these interventions, due in part to the observational nature of most data, and lack of understanding of causal mechanisms and effect heterogeneity. The paper leverages the USDA's FoodAPS dataset, employing the technique of matching to estimate the effect of food stamp recipient households' access to grocery stores and value of nutrition—the extent to which a household values healthy eating—on FV spending. The analysis finds that access impacts a household's FV spending by affecting shopping frequency, but only among households with a low value of nutrition and at distances of less than 1 mile. Value of nutrition impacts FV spending through both changes in food choices and shopping frequency, primarily among households with poor access. These findings can be used to inform more targeted and effective policy interventions.

Keywords Food deserts · Food access · Food policy · Causal inference

1 Introduction

The link between fruit and vegetable (FV) consumption and life-threatening chronic diseases such as cardiovascular disease and obesity is well established^{1,2}. Inadequate consumption of FVs is estimated to account for over 2 million deaths per year globally³. In the U.S., low FV intake is a particularly significant issue among underserved populations⁴. Thus, there are currently multiple strategies being proposed and implemented with the goal of increasing consumption of healthy food in underserved communities.

Financial support strategies (such as the Supplemental Nutrition Assistance Program, or SNAP, formerly the food stamp program) and nutrition education programs are widespread and largely found to be successful at increasing FV consumption, although with limited and varied impact^{4–7}. Access-related strategies, which increase households' physical access to healthy food, have also gained attention in the last decade. Many of these strategies target low income *food deserts*—areas where the majority of the population does not have access to large grocery stores. The federal

government currently allocates hundreds of millions of dollars each year to interventions aimed at increasing households' physical access to healthy food (henceforth referred to as *access-related interventions*), for example subsidizing new grocery stores in food deserts^{8,9}. The Healthy Food Financing Initiative (HFFI), which became a federal program in 2014, provides resources to retailers offering fresh, healthy food in underserved areas. However, due to limited funding, only about 20% of requested HFFI funds are awarded each year¹⁰. At the local level, strategies for increasing access to healthy food have also been appearing. For example, New York City, Philadelphia, and Baltimore have ongoing projects including building new grocery stores in underserved communities and increasing healthy options at corner stores^{11,12}.

Unfortunately, there is surprisingly little agreement regarding the impact of access-related interventions on FV consumption and diet. Many studies that evaluate the impact of new grocery stores on food desert residents' diets have reached mixed conclusions, often finding no average impact on FV consumption^{13–19}. Larger scale observational studies have also reached mixed conclusions on the impact of access on diet^{1,2,20–22}.

One of the key gaps in this debate is a lack of understanding of heterogeneous effects and mechanisms. Because of the observational nature of most datasets and thus the potential for edogeneity, causality is an extremely difficult topic which only a small handful of studies have meaningfully addressed^{22,23}. Moreover, estimation of how food shopping decisions vary based on key household characteristics is limited (two studies consider variation based on socioeconomic covariates^{1,24}), which can result in overlooking important information.

The goal of this paper is to inform the implementation of more effective and targeted food policy interventions, as well as inform evaluation of existing interventions. This is accomplished by providing a detailed analysis of the impact of access and *value of nutrition*—defined as how much the household values eating healthy—on FV spending (a commonly used close proxy for FV consumption), including 1) highlighting the important interplay between value of nutrition and physical access to grocery stores, 2) estimating the *continuous* impact of access on FV spending, and 3) exploring how these interventions impact diet through the mechanisms of shopping frequency and food choices.

The analysis in the paper relies on methods from statistics and operations research to elucidate the impact of access on household FV spending by considering its relationship to a household's value of nutrition. The empirical analysis in the paper employs the technique of matching for observational studies²⁵ on the nationally representative Food Acquisition and Purchase Survey (FoodAPS) dataset, focusing on SNAP-recipient households residing in urban areas¹. The paper evaluates and confirms the following two hypotheses:

Hypothesis 1. Improving access to grocery stores increases FV spending for SNAP households with a low value of nutrition, especially among those without vehicles. This increase in FV spending is caused primarily by an increase in the household's food shopping frequency.

Hypothesis 2. Increasing a household's value of nutrition (which can be accomplished, for example, through nutrition education programs) increases FV spending for all SNAP households, but especially among those with low access to grocery stores. This increase in FV spending is primarily caused by changing the household's food choices at the store.

After confirming the hypotheses above, an additional analysis is performed to further elucidate the effect of access on FV spending. The continuous effect of access on FV spending is estimated in order to answer the following question: *To what extent does a household's access need to be improved in order to increase their FV spending by X amount?* We find that for households whose closest grocery store is between 0 and 1 mile, the relationship between access and FV spending is linear, implying that any small improvement in access helps to increase FV spending. However, for households who live farther than 1 mile from a grocery store, marginal changes in access have less impact on FV spending.

The confirmation of Hypotheses 1 and 2 as well as the follow-up analysis have implications both in policy personalization and evaluation. With regard to policy evaluation, the findings suggest that household-level access and value of nutrition should be considered and measured when evaluating the impact of interventions. These findings also suggest insights for policy personalization. For example, an access-related program, such as providing free trips to grocery stores or grocery delivery services, should specifically target households with a low value of nutrition if possible. Additionally, nutrition education programs should be specifically targeted to households in food desert areas. This would achieve cost savings and result in more effective allocation of government resources.

The remainder of this paper is organized as follows. Section 2 presents an overview of the related literature and how it relates to Hypotheses 1 and 2. Section 3 reviews the data and methods used in the empirical analysis. Section 4 presents the results of the empirical analysis. Section 5 discusses the main results and proposes underlying causal mechanisms that explain the empirical findings. Section 6 shows how the empirical findings can be used to inform more personalized policy interventions. Finally, Section 7 discusses the limitations of this study, and Section 8 concludes and presents future directions.

¹According to the USDA, urban food deserts comprise between 86-93% of all food deserts

2 Literature Review

Existing literature provides evidence both for and against Hypothesis 1. Broadly speaking, the notion that distance to retail locations influences consumer decision-making is not new. In the marketing, operations management, and economics literature, the impact of distance on consumer choice has been studied in a variety of contexts. For example, Allon et al.²⁶ and Thomadsen²⁷ study consumer choice regarding fast-food outlets, and find, in part, that distance significantly impacts decision-making regarding which establishment to patronize. Davis²⁸ considers a consumer's choice of movie theater, and finds that the impact of distance is nonlinear and exhibits diminishing marginal returns. Rao and Syam²⁹ consider a consumer's choice between two grocery stores and how it is influenced in part by distance and by the goods offered by each store. Therefore, a priori it may be natural to assume that access to grocery stores will impact consumers' shopping and food choices.

Indeed, specifically with regard to food shopping, many studies have reported an association between access to grocery stores and FV spending. Gustafson et al.³⁰ find that being closer to supermarkets is correlated with increased odds of purchasing fruits and vegetables, and Liese et al.³¹ find that the availability of stores has a relationship to FV spending, but only through its relationship to the frequency of store visits. Rose and Richards²⁰ use a regression-based approach to examine the relationship between grocery store access and FV spending among SNAP households, and find a significant relationship between access and fruit consumption, and between nutritional attitude and FV consumption.

However, more recent work argues that the associations above are largely correlational in nature, and do not reflect a causal relationship between access and FV spending. Ver Ploeg and Rahkovsky²¹ note that many low income households in food deserts are willing to travel to shop at grocery stores, and thus argues that access has limited impact. In a recent study, Allcott et al.²² find that differences in access explain only a small portion of the disparities in diet found across income groups using an instrumental variables approach combined with structural model estimation. The authors conclude that demand-side factors play a much larger role, and that the impact of access (defined as the number of grocery stores within a zipcode) is marginal. Additionally, case studies on the impact of new grocery stores built in food deserts have seen mixed results, often finding no impact of the new store on diet¹³⁻¹⁹. Among the studies that reported no significant impact, the lack of effect of the new store is primarily attributed to either a low adoption rate, or stubborn food preferences/habits^{13,14,16,18}.

These studies provide evidence counter to Hypothesis 1. The arguments made in Ver Ploeg and Rahkovsky²¹ and Allcott et al.²² provide possible reasons that Hypothesis 1 may not be confirmed. It is possible that the impacts of access are simply too small to be detected, and that access should not be a first-order concern among policymakers. However, it is also possible that access *is* important, but only among certain households. Furthermore, it could be that the impact of access is highly sensitive to the distances considered, and that the definition of access used in Allcott et al.²² is not granular enough to capture potential impacts of small distances. These reasons may also explain the mixed results found in case studies. Most of these studies did not measure household covariates nor their individual distances to the new store, and therefore more detailed analyses cannot be conducted. One case study by Wrigley et al.¹⁵ that examined the impact of a new store built in a food deserts in Leeds, however, did measure certain attributes of the households in the study, and notably found that households with the lowest initial levels of FV consumption were the ones most impacted by the new store. This study indicates the importance of considering effect heterogeneity and provides some evidence in support of Hypothesis 1.

Finally, the point made by Ver Ploeg and Rahkovsky²¹—that low-income households residing in food deserts are willing to travel to grocery stores—ignores the importance of shopping frequency and perishability, which was highlighted in Liese et al.³¹. Even though many low income households appear willing to travel to grocery stores, these households likely shop with lower frequency than they would if the store were closer. Therefore, they would be less likely to purchase perishable foods. The analysis in this paper addresses all three of these nuances—heterogeneous effects, varying definitions of access, and the notion of shopping frequency—in order to reveal new insights that can be used to directly inform policy.

Hypothesis 2 is more accepted in the literature. Many studies have found a positive relationship between nutrition education and healthy eating. Beydoun et al.³² and Wardle et al.⁶ find that nutritional attitude/knowledge has a positive relationship with FV consumption, and Gibson et al.³³ find that a mother's nutritional knowledge has a significant relationship with her children's vegetable intake. Axelson et al.⁷ and Spronk et al.³⁴ perform a meta-analysis of studies on the relationship between nutrition education (or nutritional attitudes, values, beliefs, etc.) and fruit and vegetable consumption, and conclude that although the relationship is generally positive, the association is typically weak, and the effect can vary widely depending on the population of interest and the chosen proxy to measure nutritional knowledge. The interaction between nutrition education and other attributes of the household has been studied to some extent (for example, Beydoun and Wang³⁵ consider variations in value of nutrition with respect to certain demographics), however, to the authors' knowledge, the interaction between value of nutrition and access has not been studied.

3 Study Data and Methods

3.1 Data

This paper employs the USDA’s Food Acquisition and Purchase Survey (FoodAPS) dataset, which was made available for public use in 2016. Between April 2012 and January 2013, 1,581 SNAP-receiving households were surveyed for a one-week period (which is referred to as their “audit week”) and asked to keep detailed food purchasing records, including store-bought and restaurant food. This dataset is novel in that it is the first nationally representative dataset with detailed information on dietary habits, household-level demographic covariates, and household-level covariates related to the food retail environment and access to various food outlets. This paper focuses on the analysis of the set of SNAP households residing in urban areas contained within this dataset.

3.2 Outcome variables

The primary outcome of interest is a household’s total FV spending over their audit week. Although FV consumption is of primary interest because of its known association to health outcomes, FV spending is a commonly used proxy for FV consumption because it can be measured and verified. Furthermore, among low income households who are unlikely to waste large amounts of food, it is likely a close approximation of FV consumption.

FV spending includes purchases made at grocery stores, convenience stores, and other stores, but does not include expenditures at restaurants (see Appendix A.1 for a discussion). Total FV spending is standardized to account for different household compositions in terms of household size, family members’ ages, and family members’ sexes, and is thus in units of “spending per standardized person per week.” Total FV spending considers all items purchased within the USDA food category Fruits and Vegetables, which includes fresh, frozen, and canned fruits and vegetables. However, in the FoodAPS dataset, 96% of items purchased by SNAP households in this category are fresh fruits and vegetables.

A secondary outcome of interest is whether the household visited a large grocery store (defined below in Section 3.3.1) or not during their audit week. Across all urban SNAP households in the dataset, 83% visited a large store during their audit week.

3.3 Treatment variables and covariates

3.3.1 Access

This paper defines access in accordance with USDA definitions as distance to the nearest supermarket, supercenter, or large grocery store, all of which are considered to be “large grocery stores”³⁶. The household’s distance to the nearest large grocery store is referred to as its “store distance”. Because there is some evidence that superstores may have a negative impact on diet^{37,38}, an additional analysis where supercenters are removed from the definition of “large grocery store” is included in Appendix G.3.

Straight-line distances between each household and the nearest food outlet of various types is reported in the FoodAPS dataset. Through these metrics, store distance is constructed by taking the minimum of a household’s straight-line distance across all store types included in the definition of “large grocery store”.

Standard USDA definitions of urban food deserts arbitrarily define “low access” as having a store distance farther than either 0.5 or 1 mile, depending on the definition. The definition that employs a threshold of 1 mile is the most commonly adopted definition in the literature. This paper takes a data-driven approach for choosing the threshold, and allows the threshold to vary for households with and without vehicles. The thresholds used in this analysis are 0.5 miles and 0.4 miles for households in urban areas with and without vehicles, respectively. These thresholds are the median store distances among all households in the specified population. The median store distances were chosen in order to increase the statistical power of the tests performed. A post-hoc analysis that considers choosing different thresholds is shown in Appendix G.1.

An additional analysis is performed in which access is treated as continuous, rather than binary. Instead of imposing a threshold on store distance, in this analysis the raw store distance of each household is considered to be a continuous treatment. This allows for a more detailed understanding of the relationship between access and FV spending.

3.3.2 Value of nutrition

A household’s *value of nutrition* is a measure of how much the household cares about eating nutritiously, and is estimated using certain behaviors which are commonly inquired about in USDA surveys; for example, how often they

search for nutrition information online, look at nutrition labels, or try to follow the Food Pyramid guidelines. Using the answers to these questions, each household is classified as having either a “high value of nutrition” or “low value of nutrition.” More details on this construction can be found in Appendix B.2, and a detailed post-hoc analysis of the relative impact of each covariate used in the definition of *value of nutrition* can be found in Appendix G.4.

3.3.3 Covariates

Relevant covariates—those that are likely to either impact FV spending, or to be correlated with access or value of nutrition—are shown in Table 1, aggregated separately for households with low and high access. Most of these covariates are controlled for in the analysis by their inclusion in the matching step, described in Section 3.4.2.

3.4 Statistical analysis

3.4.1 Hypotheses

The analysis tests the following hypotheses:

Hypothesis 1. Improving access to grocery stores increases FV spending for SNAP households with a low value of nutrition, especially among those without vehicles. This increase in FV spending is caused primarily by an increase in the household’s food shopping frequency.

Hypothesis 2. Increasing a household’s value of nutrition (which can be accomplished, for example, through nutrition education programs) increases FV spending for all SNAP households, but especially among those with low access to grocery stores. This increase in FV spending is primarily caused by changing the household’s food choices at the store.

Both hypotheses are tested by employing the technique of matching in observational studies combined with randomized inference. This technique is described in more detail below.

3.4.2 Matching in observational studies

Matching techniques comprise a class of intuitive and transparent methods for estimating the causal effects of a treatment using observational data, and are an alternative to methods such as instrumental variables^{25,39}. The idea behind matching techniques is to group together households who are similar based on their observed covariates, but differ in their treatment assignment. Matching techniques, like most techniques for observational studies, rely on the assumption of unconfoundedness, which says that all covariates that either impact the treatment assignment or the outcome are accounted for. Depending on the quality of the match and the inclusion of relevant covariates, valid randomized inference can be performed within each group. In other words, each group can be treated as if it resulted from a randomized study. Thus, causal effects can be estimated.

Propensity score matching (PSM) is a specific type of matching, and Ho et al.⁴⁰ cites PSM as a technique that is under-utilized in the empirical OR/MS literature. Propensity score matching groups households together based on their *propensity score*, which is defined as their probability of receiving treatment given their observed covariates. A household’s true propensity score is never known, but is estimated using the household’s observed covariates and treatment assignment.

All covariates that can either affect the outcome or treatment assignment should be included in the propensity score estimation in order to minimize the possibility of confounding³⁹. On the other hand, covariates that are affected by the treatment should *not* be included in the propensity score estimation³⁹. Table 1 lists relevant covariates, indicating which were chosen for inclusion in the propensity score estimation. Those excluded were deemed likely to be impacted by the treatment (in this case, access to grocery stores or value of nutrition).

Propensity-score matching has many desirable statistical properties. First, after matching on propensity score, randomized inference within the matched groups is valid⁴¹. In other words, groups of households matched together based on true propensity score can be treated as if they were the result of a randomized study. Second, the underlying covariates of the households in the matched treated and control groups have the same distribution⁴¹. Therefore, matching on propensity score is intuitive in that it balances the households’ underlying covariates in each group.

This second theoretical property of PSM—its ability to balance the underlying covariates—provides a diagnostic tool for checking the quality of the matches obtained when PSM is performed in practice. If the underlying covariates are imbalanced after performing PSM, this suggests a problem in either the propensity estimation or the matching technique. In this study, the balance of all matches were checked (shown in Appendix D.1), and in some cases if the balance produced by PSM was inadequate, Mahalanobis matching with propensity score calipers was used to better balance the imbalanced covariates³⁹.

Since the balance of each match was carefully examined and only accepted if it met certain balance requirements, it is unlikely that other matching methods (such as coarsened exact matching, which matches directly on the underlying covariates) would produce significantly different results. Matching methods were chosen over regression-based methods for a few reasons: i) Appropriate instrumental variables could not be identified for this dataset, ii) Matching explicitly considers covariate balance and overlap, and allows for straightforward sensitivity analyses, and iii) Matching methods typically rely less on extrapolation and model-based assumptions³⁹.

Observational studies typically proceed under the assumption of strong ignorability, which includes an assumption of no hidden confounding (or unconfoundedness). This assumes that all covariates that either impact the outcome or treatment assignment have been accounted for. If this assumption does not hold, there could exist a hidden confounder that impacts both the treatment assignment and the outcome, thus rendering the propensity score estimation and randomized inference invalid. The assumption of strong ignorability can never be verified because of the immutable potential for hidden biases in observational studies. However, this study performs a sensitivity analysis in order to assess the robustness of the results to potential hidden confounders⁴². The unconfoundedness assumption is further discussed in Section 7 and Appendix F and E.

3.4.3 Treatment effect

To estimate treatment effects, the Neyman-Rubin potential outcome framework and stable-unit treatment value assumption are employed⁴³. Under this framework, each unit has two *potential outcomes*—one if they were to receive treatment, and one if they were to receive control. The stable-unit treatment value assumption says that one unit’s response is not affected by another unit’s treatment assignment. The treatment effect is the difference in potential outcomes for an individual, or the average difference over a population. The fundamental problem of causal inference is that one cannot observe more than one potential outcome for a given unit at a given time.

This paper considers two treatments: access to grocery stores and value of nutrition. Under the Rubin causal model, these are valid treatments because they can be manipulated⁴⁴ (i.e., potential outcomes can be fathomed). In particular, the goal of nutrition education programs is to manipulate one’s value of nutrition⁴⁵.

In order to estimate the impact of access (as a binary treatment), and value of nutrition on the binary outcome of whether or not a household visited a large store during their audit week, this paper focuses on estimation of the risk difference, and tests the null hypothesis that the risk difference is equal to zero⁴⁶. The risk difference, denoted δ , is interpreted as the difference in the proportion of households who would have gone to a store if *everyone* had received treatment and the proportion of households who would have gone to a store if *no one* had received treatment. At the household level, this can be thought of as the increase in likelihood of a store visit if the household were to receive the treatment (either better access or higher value of nutrition). Notice that this relates to a household’s shopping frequency. An intervention that increases a household’s likelihood of a store visit in a given week must also increase the household’s average shopping frequency.

To assess the impact of access (as a binary treatment) and value of nutrition on FV spending, a stratified Wilcoxon test is performed in order to test for a significant difference in FV spending between the treatment and control groups after matching. In addition to reporting the p-values obtained by this test, the treatment effect sizes are estimated by assuming a Tobit effects model. A Tobit effects model assumes that FV spending, denoted Q_i for household i , can be written as $Q_i = \max\{\alpha_i + \tau z_i, 0\}$ where $\alpha_i \geq 0$ is a latent variable that is specific to household i , z_i indicates if household i receives the treatment ($z_i = 1$) or control ($z_i = 0$), and τ is the average treatment effect⁴⁷. The potential outcome for household i if they were to be in the control group is denoted Q_i^c . For households in the control group, $Q_i^c = Q_i$, and for households in the treated group, $Q_i^c = \max\{Q_i - \tau, 0\}$. The estimate of τ is given by the Hodges-Lehman point estimate, which is the value of τ for which the calculated test statistic for the stratified Wilcoxon test, using the value Q_i^c for household i , is closest to its expectation⁴⁷.

In order to better understand the mechanisms linking the treatments to changes in FV spending, the impact of access and value of nutrition on FV spending *per store visit* is also estimated (conditional on having gone to the store at least once). This treatment effect estimate is denoted τ_{pv} . A positive estimate of τ_{pv} means that the given treatment is estimated to increase the amount of FVs bought per shopping trip.

A positive estimate of τ_Q —the overall increase (interpreted in the context of the Tobit model) in FV spending due to either access or value of nutrition—should correspond to either a positive estimate of δ or τ_{pv} , or both. Intuitively, this means that an increase in FV spending must either be caused by an increase in shopping frequency, or an increase in the amount of FVs bought each shopping trip, or both.

In addition to considering access in binary terms, the *continuous* effect of access (i.e., store distance) on FV spending is estimated. In general settings, the function that estimates the impact of a continuous treatment on an outcome is referred to as the *dose-response function* (DRF). In this study, the DRF is estimated as a smooth coefficient model

of store distance and the propensity function, which is a generalized version of the propensity score for non-binary treatments⁴⁸. The DRF estimates changes that occur in FV spending due *solely* to changes in store distance, again under the assumption of no hidden confounding. Details on this method can be found in Appendix F.1.

4 Study Results

Table 2 shows the results of the empirical analysis. Columns d.1-d.3 pertain to Hypothesis 1 and show the estimated effect of access on overall FV spending, the likelihood of a store visit in a given week, and amount of FVs bought per visit in the populations listed in Column 1. Columns v.1-v.3 pertain to Hypothesis 2 and show the estimated effects of value of nutrition. For example, Row 2 Column d.1 (or Column v.1) shows the estimated effect of access (or value of nutrition) on total FV spending across all households without a vehicle.

Hypothesis 1. The results in Table 2 confirm Hypothesis 1. Among the overall population, access is estimated to have a positive but insignificant impact on total FV spending ($\tau = 0.24, p > 0.1$). However, when the population is broken down by either vehicle ownership or value of nutrition, the impact of access is elucidated. Although both the populations without a vehicle and with a low value of nutrition are impacted by access, Rows 8-11 of Table 2 suggest that value of nutrition is a more important characteristic than vehicle ownership for understanding and predicting the effect of access on total FV spending. Furthermore, among most subpopulations it appears that access impacts the likelihood that a household visits a large grocery store in a given week (column d.2), and does not have a significant impact on the amount of FVs bought per visit (column d.3).

Noting that the average FV spending among households with low access is \$3.83 per person per week (Table 1), the estimated effects of access on the population without a vehicle and the population with low value of nutrition (\$1.25 and \$0.76, respectively) are substantial. Interpreted within the context of the Tobit effects model, it is estimated that a household with low value of nutrition and high access would buy \$0.76 less FVs per person per week if they had low access. This translates into an average decrease of 27% in total FV spending among this population, based on the FoodAPS dataset. Similarly, those with a high access but without a vehicle would purchase \$1.25 less FVs per person per week if they had low access, translating into an average decrease of 33% based on the FoodAPS dataset.

Hypothesis 2. The results in Table 2 confirm Hypothesis 2. Although value of nutrition appears to have a positive and generally large estimated impact on total FV spending among all populations considered, it is especially pronounced among those with low access (Rows 6, 14-15 of Column v.1, Table 2). Surprisingly, value of nutrition impacts FV spending by impacting *both* the amount of FVs bought per visit as well as the likelihood of a store visit in a given week. This is discussed in the next section.

After confirming Hypotheses 1 and 2, an additional analysis of the impact of access on FV spending is performed in order to understand their relationship in more detail. Access—specifically, store distance—is of course a continuous covariate. Typically in the literature, and as mentioned in Section 3.3.1, a threshold is chosen, and households with store distances less than the threshold are classified as having “low access.” To the authors’ knowledge, the continuous impact of store distance on FV spending has not been studied. However, it is extremely important for determining effective intervention strategies.

The effect of store distance, as a continuous treatment, on FV spending is visualized through the estimated dose-response function (DRF). Figure 1 shows the estimated DRF among households with a low value of nutrition. The DRF is the estimated average change in FV spending that is due *solely* to changes in store distance (i.e., with all other covariates held constant). Based on Figure 1, it appears that store distance has a linear effect on FV spending for distances between about 0-1 mile. For distances larger than 1 mile, the marginal impact of store distance on FV spending diminishes. This implies that, for distances between 0-1 mile, any decrease in store distance has the same marginal effect on FV spending. Changes that occur at larger distances, however, are less effective.

5 Discussion and Behavioral Mechanisms

First, the analysis suggests that access does have an impact on FV consumption, but only among certain subpopulations. In particular, access primarily impacts households with a low value of nutrition, and at small distances. The estimated DRF in Figure 1 illustrates the importance of considering small distances, indicating that building a new grocery store will likely only impact households who are very close to the new store (within about 1 mile). This implies that the definition of access is extremely important. It is likely that commonly used region-based definitions (for example, how many grocery stores are in a certain zipcode or census tract^{22,31}) are not precise enough or sufficiently aggressive to elucidate the impact of access. Instead, household-specific measures of access and small distance thresholds should be utilized in order to capture the true effect of access.

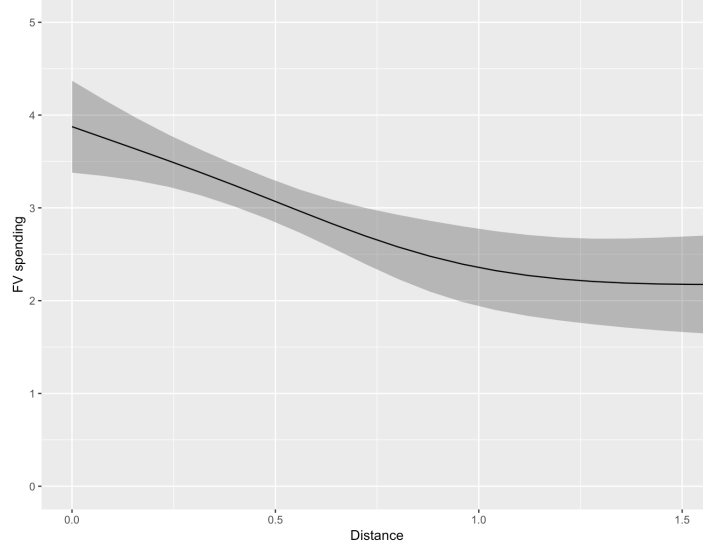


Figure 1: The estimated dose-response function (DRF) for the effect of distance on FV spending among households with a low value of nutrition. Grey represents standard errors estimated from 1,000 bootstrap samples.

Furthermore, considering a household’s value of nutrition in conjunction with access is crucial. The analysis suggests that value-based interventions (for example, nutrition education programs that change a household’s value of nutrition) are likely to be most effective among households with low access, and that access-related interventions are most effective among households with a low value of nutrition. This finding could partly explain why case studies have reached different conclusions regarding the effectiveness of access-based interventions—specifically, not considering the appropriate heterogeneous effects. One case study on a new grocery store built in a food desert that did consider heterogeneous effects concluded that the new grocery store primarily impacted households who were very close to the new store, or had the lowest amounts of FV consumption to begin with¹⁵. These results are in line with the predictions of this paper, since households with low value of nutrition typically consume less FVs to begin with.

The above insights suggest certain underlying behavioral mechanisms that drive households’ shopping decisions. Changes in total FV spending are explained through two mechanisms: (i) Changes in shopping frequency, and (ii) Changes in the amount of FVs bought each grocery store visit.

Column d.2 of Table 2 indicates that access primarily impacts grocery shopping frequency, which translates into changes in total FV spending for certain subpopulations. In the overall population (Row 1), decreasing access is estimated to increase shopping frequency but not significantly impact total FV spending on average (a finding which is similar to the conclusions of Liese et al.³¹). However, subsetting the population based on value of nutrition provides new insights to better understand which households’ FV spending, not just shopping frequency, will be impacted by access.

Value-based interventions, on the other hand, appear to primarily impact the households’ food choices at the store. However, value-based interventions can also increase grocery shopping frequency for households with low access. To our knowledge, this mechanism has not been discussed in the literature. In what follows, a basic model of grocery shopping dynamics is discussed that is able to explain this finding as well as many of the heterogeneous effects observed in Table 2.

One approach to model how access and value of nutrition impact shopping frequency is through what is referred to as *shopping utility* and *shopping disutility*. At any point in time, each household has a *shopping disutility*—the degree to which grocery shopping is a burden—that depends on the household’s characteristics (access to the store, opportunity cost, etc.) and exogenous temporal variables (day of week, time of day, etc.). Furthermore, at any point in time, each household has a *shopping utility* for going to the grocery store based on how much they would be benefited by the bundle of food they would buy. Shopping utility is relatively low for households who recently went shopping, and is positive and constantly increasing once food begins to run out. A shopping trip occurs once the household’s shopping utility exceeds its shopping disutility. Shopping frequency is thus impacted by access, value of nutrition, and other characteristics through their impact on shopping utility and disutility.

Furthermore, each time a household visits the grocery store, they choose a bundle of food to purchase based on their preferences, the prices, their value of nutrition, and the foods’ perishability, among other factors. The bundle’s size

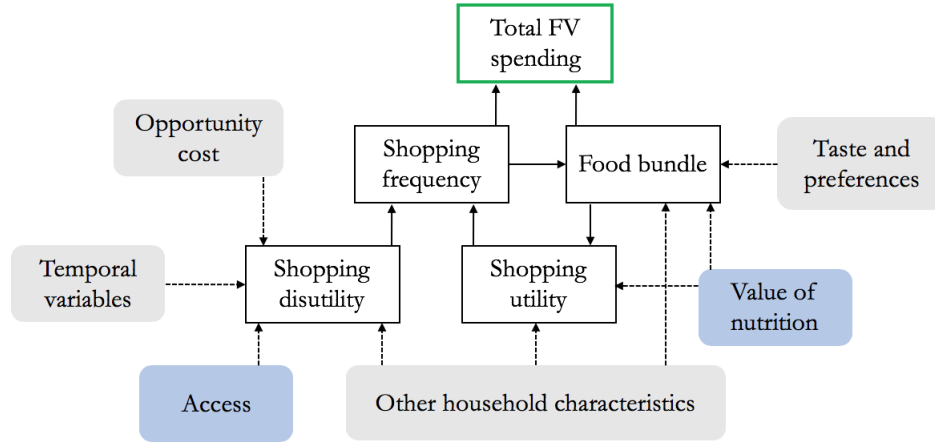


Figure 2: Proposed causal pathways and mechanisms linking household characteristics, access, and value of nutrition to mechanisms, and mechanisms to total FV spending.

and perishability will also affect how shopping utility evolves over time, thus potentially affecting shopping frequency. Total FV spending is a combination of both shopping frequency and bundle of foods bought each visit (specifically, the amount of FVs in this bundle). In order to determine the net effect of value of nutrition on total FV spending, it is important to understand how it changes the amount of FVs in the food bundle and subsequently how this change impacts shopping frequency. A diagram explaining the proposed causal pathways is shown in Figure 2. Note in particular the feedback loop between Shopping frequency, Food bundle, and Shopping Utility.

First, this model explains why value of nutrition impacts shopping frequency in addition to the amount of FVs bought each visit. Since households with a high value of nutrition purchase more FVs per grocery trip, they must also shop more frequently due to the perishability of the food. This mechanism becomes clear after considering the interaction between the bundle of food and shopping frequency.

This model can also be used to explain many of the complex heterogeneous effects found in the empirical analysis; for example, the finding that the shopping frequency of households with low value of nutrition is more impacted by access than the shopping frequency of those with a high value of nutrition. In terms of the proposed causal pathways, the explanation for this phenomenon is that the shopping utility of households with a low value of nutrition increases more slowly than that of households with a high value of nutrition. Therefore, an increase in access, resulting in a decrease in shopping disutility, is more impactful for households with a low value of nutrition. In other words, households with a high value of nutrition but low access are likely to return to the grocery store soon after their grocery food depletes, even if this requires significant effort, because their urgency to return to the store (their *shopping utility*) is increasing very quickly. Therefore, a new grocery store will not impact their shopping frequency as much. On the other hand, households with a low value of nutrition and low access will likely take more time to return to the grocery store after their food is depleted, since they are more willing to rely on food from other outlets such as fast food restaurants or convenience stores until their urgency to return to the store (their *shopping utility*) has grown large enough. A closer grocery store therefore impacts their shopping frequency more significantly.

6 Policy implications

From a policy perspective, these findings imply that the deployment of access-related interventions should be based not only on the spacial composition of a neighborhood, but also on its residents' values of nutrition. Access-related interventions should not just target low income food deserts, but should target *low value food deserts*—areas where, in addition to low access and low income, most households have a low value of nutrition.

As a concrete illustrative example, consider a synthetic population of low income households residing in two different food desert neighborhoods, shown in Figure 3. For each household, it is assumed that their location (plotted as points in Euclidean space) and value of nutrition are given. The policymaker must decide in which neighborhood to build a new grocery store. This decision should be based on the predicted impact of the new store, which depends in part on the spacial composition of the neighborhood surrounding the store and the households' values of nutrition. In Figure 3, the

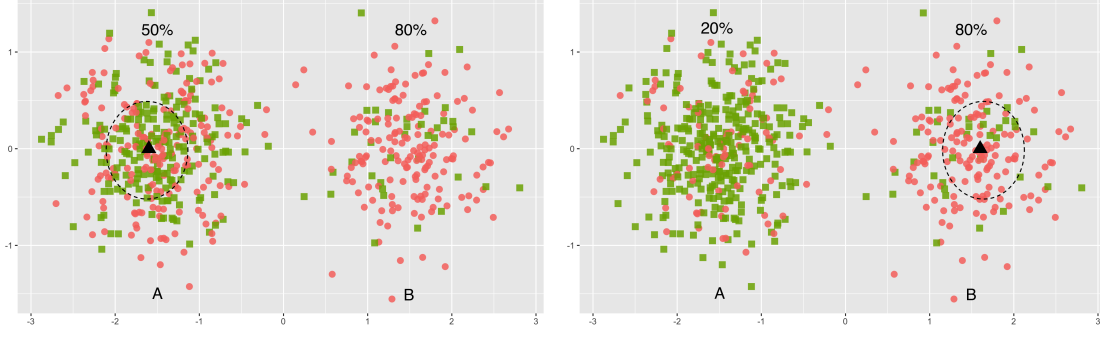


Figure 3: Illustrative example of the importance of considering value of nutrition in conjunction with spatial factors for grocery store placement. Households with a low value of nutrition are pink and households with a high value of nutrition are green.

Each dot represents a household’s location, and value of nutrition is shown in color. A new grocery store is represented as a black triangle, with a circle drawn around to emphasize the households that are “near” the new store. Percentages shown above each neighborhood show the portion of households with a low value of nutrition.

households are colored according to their value of nutrition, and a “radius of influence” is drawn around the new stores. This radius represents a simplified characterization of the impact of a new store. For this example it is assumed that households inside the radius are impacted uniformly by the introduction of the new store (based on the results of the analysis and Figure 1, the radius should be somewhere between 0.5-1 mile, although the exact radius is not important for this example). The proportion of households in each neighborhood with a low value of nutrition is also given. The impact of a new store is likely to be greatest when placed in an area with the largest number of households with a low value of nutrition.

In Figure 3 (left), although neighborhood B has a greater portion of households with a low value of nutrition (80%), neighborhood A has a higher population density and enough households with a low value of nutrition (50%) such that it is the preferable neighborhood to build a new store. In Figure 3 (right), neighborhood A has fewer households with a low value of nutrition (only 20%). In this scenario, even though neighborhood A has a high population density, it may be more beneficial to build a new store in neighborhood B in order to impact more households with a low value of nutrition. This small illustrative example demonstrates the importance of considering spatial composition and value of nutrition in conjunction with one another. The cost-effectiveness of a new grocery store depends both on how far neighboring households will live from the new store, and their value of nutrition.

This discussion highlights two important lingering questions: (1) How can value of nutrition be estimated at the household or neighborhood level, and (2) Do *low value food deserts* exist (in other words, does value of nutrition vary a significant amount across food deserts)?

There are a number of methods that could be used to evaluate neighborhood-level average value of nutrition. For example, policymakers could deploy surveys specifically to food desert residents and ask questions related to value of nutrition. Table 3 shows the questions that were contained in the FoodAPS dataset and used as proxies for value of nutrition in this study. In particular, one of the questions listed in Table 3 asks whether the household has searched the internet for nutrition-related information recently. This question motivates a second approach for estimating neighborhood-level value of nutrition based on Google search data. Namely, if certain nutrition-related neighborhood-level Google search trends were made available to policymakers, *low value food deserts* could be identified. For example, if one neighborhood has a smaller proportion of searches related to nutrition, calories, etc. than another, it might be reasonable to assume that this neighborhood has a lower average value of nutrition.

To illustrate this approach, publicly available Google search data (whose lowest level of granularity is the city-level) is used to estimate city-level value of nutrition for cities within Los Angeles County. Figure 4 shows a map of Los Angeles County with cities colored according to this proxy for value of nutrition. Namely, in Figure 4, the darker red regions correspond to higher frequencies of search terms related to “calories” (see Appendix C for more details). Low income food deserts—defined according to the USDA definition as census tracts where the majority of the population has a store distance greater than 0.5 miles and are low income, based on 2015 USDA data³⁶—are overlayed and outlined in black.

Because neighborhood-level Google search data is not publicly available, this illustrative example makes the strong assumption that food deserts lying within a given city have the same value of nutrition as the city-wide average. Under

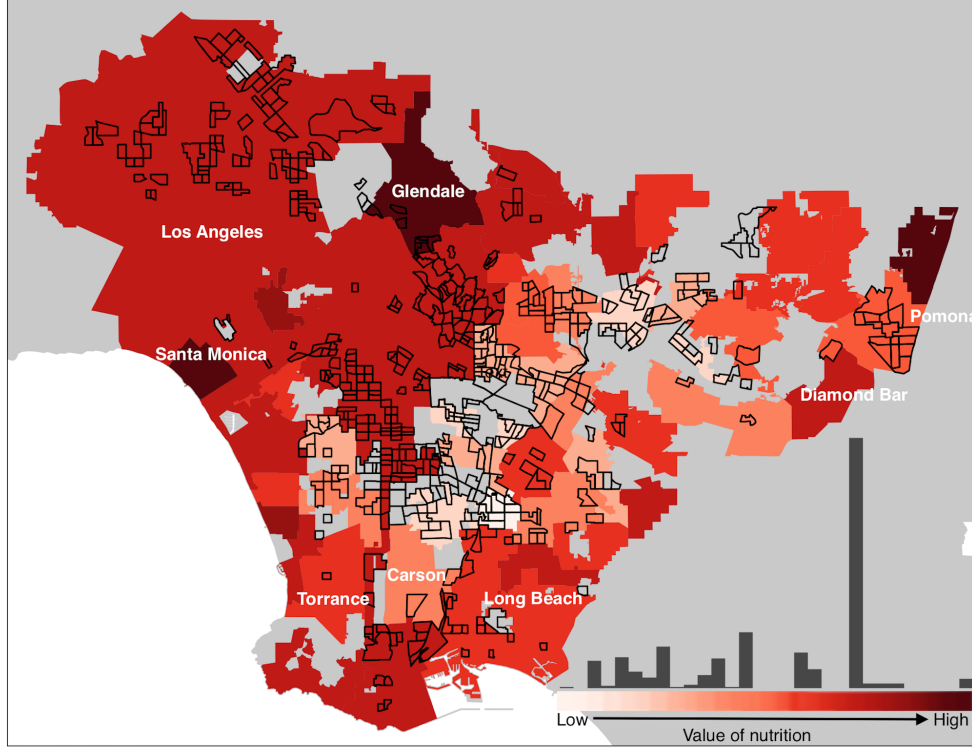


Figure 4: Map of part of Los Angeles County (with a few cities labeled for reference) with low income food deserts[†] shaded[‡] by a proxy for value of nutrition based on Google search data.

[†] Food deserts are defined as census tracts where the majority of the population resides over 0.5 miles from the closest large grocery store.

[‡] Light grey regions are either outside of L.A. County or do not have Google search data available.

Histogram shows the number of L.A. County food deserts with the given estimated value of nutrition. Areas with the lowest estimated value of nutrition search for calorie-related information half as frequently as those with the highest estimated value of nutrition.

this assumption, the histogram in the bottom right corner of Figure 4 shows the number of low income food deserts with estimated value of nutrition at the levels shown. Those with the lowest estimated value of nutrition search for calorie information half as frequently as those with the highest estimated value of nutrition. As illustrated by the histogram in Figure 4, there is significant variation in estimated value of nutrition across low income food deserts. This example provides initial evidence that *low value food deserts* do exist; namely, that there exist food deserts with high, medium, and low average value of nutrition.

Based on the analysis in this paper, food deserts with a low value of nutrition—lighter red in Figure 4—are more likely to be impacted by access-related interventions than food deserts with a high value of nutrition. Since there are limited funds for access-related interventions (only 20% of requested HFFI funds are awarded¹⁰), targeting specific food deserts based on value of nutrition could be extremely useful for deciding which interventions to fund and improving the cost-effectiveness of these strategies.

The method for estimating value of nutrition in this example has many limitations. The lowest level of granularity for publicly available Google search data is the city level, which may not accurately capture search trends at the neighborhood level. It is possible that more granular data could be obtained through a public-private partnership. Furthermore, searches for calorie-related information may not be indicative of true value of nutrition. In practice, policymakers should utilize multiple sources of data—including neighborhood-level survey data and internet search data—to robustly identify neighborhoods with a low average value of nutrition.

7 Limitations

There are several limitations to this study, largely due to the observational nature of the data. With regard to FV spending, although receipts were gathered to verify food purchases, it is possible that some purchases are misreported. Furthermore, the outcome variable—total FV spending—is not a perfect indication of FVs consumed, for two main reasons: 1) Prices are not equivalent to quantities, and 2) The household may not consume everything that is bought. In order to alleviate the first issue, the price of certain food items across households are compared and no systemic differences that could bias the results are found (Appendix A.3). The second point is mitigated by the fact that low income households are unlikely to waste large amounts of food, although this cannot be verified using the data.

With regard to value of nutrition, it is possible that some households exaggerated their nutrition-related behaviors in the survey. However, it is likely that misclassification of households' value of nutrition would result in *underestimating* the observed effects of value of nutrition. This point is discussed in more detail in Appendix B.2.

The dataset also contains only one week's worth of data for each household. Although FV spending over a one week period may not be a good indication of typical FV spending for a single household, this problem is likely partially mitigated by considering average FV spending over populations.

As with any observational study, the possibility of hidden confounding can never be ruled out. One main concern is the potential for hidden biases in the estimated effect of access, since this relationship is highly debated in the literature. In general, the elucidation of an intuitive dose-response relationship provides additional evidence in support of a causal association⁴⁹. This paper elucidates a dose-response relationship in two ways: 1) By estimating the DRF, and 2) by considering “extreme distances” in Appendix F.2. Estimation of the dose-response function in Figure 1 reveals a monotonic association between store distance and FV spending, with diminishing marginal returns. This trend is likely what one would expect to observe regarding the true causal relationship between store distance and FV spending. Furthermore, the analysis of store distance as a binary treatment is replicated for “extreme distances” in Appendix F.2: The effect, on FV spending, of having a “very short” versus a “very long” store distance is estimated. This results in a larger estimated effect of store distance, again in line with the hypothesized causal effect. These analyses provide additional evidence that the estimated effects of access on FV spending are indeed causal. In order for a hidden confounder to explain these results, it would need to be correlated with store distance as a continuous covariate.

Additionally, a sensitivity analysis is performed following the methodology of Rosenbaum²⁵. The sensitivity analysis characterizes the degree of departure from group-wise randomness that would be necessary to materially alter the conclusions of the study, in terms of the households' propensity scores. For this study, it is found that an unmeasured confounder would need to systematically affect the odds of living near a grocery store by about 26% in order to overturn the results related to the effect of access on FV spending. Similarly, an unmeasured confounder would need to impact the odds of having a high value of nutrition by 29% in order to overturn the results related to the impact of value of nutrition on FV spending. The specific results and methodology can be found in Appendix E. Although such confounders can never be ruled out, they would need to be strongly correlated with either value of nutrition or store distance, not already controlled for in the matching, and need to systematically affect the entire population under examination.

It is impossible to verify with certainty whether the relationships elucidated in this study are indeed causal. Given their alignment with findings in the literature and their largely intuitive nature, the authors believe it is unlikely that they are caused by hidden confounders. The evaluation of future access-related interventions could help support or reject the hypotheses proposed in this paper by measuring households' value of nutrition and access before and after the intervention.

8 Conclusions and Future Directions

Currently, strategies aimed at assisting low-income food desert residents follow an implement-then-evaluate framework, which is costly and time consuming. This study is a step towards a different approach that leverages an understanding of the underlying mechanisms and household-level characteristics in order to predict the impact of interventions and optimize them. The key insights of this paper can help to increase the efficacy of access-based interventions through better targeting and evaluation.

First, this paper finds evidence that access-related interventions are effective at increasing FV spending but only among households with a low value of nutrition. Furthermore, access-related interventions primarily impact FV spending by increasing the households' grocery shopping frequency. Additionally, in order to better understand the relationship between access and FV spending, this study estimates the continuous effect of access on FV spending through estimation

of the dose-response function. This analysis reveals that store distance linearly impacts FV spending for distances approximately between 0-1 mile. At larger distances, the marginal impact of store distance on FV spending diminishes. This suggests that access-related interventions need to dramatically improve access in order to be effective. For example, a new grocery store that decreases the average household's store distance from 2 miles to 1.5 is likely to be much less effective than a store which decreases the average store distance from 1 mile to 0.5 miles.

Next, this paper finds that value-related interventions, such as nutrition education programs, are most effective among households with poor access. Furthermore, value-related interventions appear to primarily impact households' food choices, but may also increase shopping frequency. The increase in frequency is likely related to the increase in perishable foods purchased, which require more frequent shopping trips. It remains an area for future research to determine which types of nutrition education programs have the greatest impact on value of nutrition.

The idea of personalization is permeating and revolutionizing fields such as medicine and advertising, and there is no reason that it could not provide the same benefit to public policy. However, in order to effectively prescribe personalized interventions, a deep understanding of consumer decision-making is required. As future work, the insights of this paper and findings from the literature could be used to develop a detailed consumer-level behavioral model of grocery shopping and food choice. Such a model could be used to fully optimize government investment into various interventions. Furthermore, it is likely that the impact of access is also affected by other household covariates besides value of nutrition. Exploring additional effect heterogeneities could reveal new insights and increase the accuracy of such a consumer-level model.

An accurate consumer-level model of food shopping has the potential to improve the efficacy of food policy interventions while decreasing costs. Not only could such a model capture the effects of a single intervention, such as access, but could even capture the joint effects of multiple types of interventions (such as combinations of access, value of nutrition, and price interventions).

Future work should also focus on randomized access interventions in order to verify the results of this study and others. With the government's strong focus on food deserts and access, a randomized study should be a priority. It is impractical to randomize the building of new grocery stores because of their huge fixed costs, however, other access interventions could be randomized. For example, a cohort of low income food desert residents could be given free grocery delivery memberships, thus removing the fixed cost of grocery shopping. This is an example of an access-related intervention that would largely reduce the disutility associated with grocery shopping. It is only through studies like this that the causal impact of access on FV spending can be verified without any potential for confounding.

Acknowledgment

The authors would like to thank their colleague Colin Fogarty of MIT Sloan for his methodological expertise and advice. In addition, the authors thank the anonymous referees and Associate Editor for their insightful comments that greatly improved the paper. This research is supported in part by the National Science Foundation Graduate Research Fellowship.

References

- [1] Kimberly Morland, Steve Wing, and Ana Diex Roux. Supermarkets, other food stores, and obesity: the atherosclerosis risk in communities study. *Am J Prev Med*, 2006.
- [2] Nicole Larson, Mary Story, and Melissa Nelson. Neighborhood environments: Disparities in access to healthy foods in the U.S. *American Journal of Preventative Medicine*, 2009.
- [3] Victoria Miller, Frederick Cudhea, Gitanjali Singh, Renata Micha, Peilin Shi, Jianyi Zhang, Jennifer Onopa, Dimitra Karageorgou, Patrick Webb, and Dariush Mozaffarian. Estimated global, regional, and national cardiovascular disease burdens related to fruit and vegetable consumption: An analysis from the global dietary database. In *Proc. Nutrition 2019*, 2019.
- [4] Diansheng Dong and Biing-Hwang Lin. Fruit and vegetable consumption by lowincome americans: Would a price reduction make a difference? Economic Research Report 70, U.S. Department of Agriculture, Economic Research Service, 2009.
- [5] Lauren EW Olsho, Jacob A Klerman, Parke E Wilde, and Susan Bartlett. Financial incentives increase fruit and vegetable intake among Supplemental Nutrition Assistance Program participants: a randomized controlled trial of the USDA Healthy Incentives Pilot. *The American Journal of Clinical Nutrition*, 104(2):423–435, 2016.
- [6] J Wardle, K Parmenter, and J Waller. Nutrition knowledge and food intake. *Appetite*, 34(269-275), 2000.

- [7] Marta L. Axelson, Terri L. Federline, and David Brinberg. A meta-analysis of food- and nutrition-related research. *Journal of Nutrition Education*, 17(2):51 – 54, 1985. ISSN 0022-3182. URL <http://www.sciencedirect.com/science/article/pii/S0022318285801758>.
- [8] The Reinvestment Fund. Online, 2019. URL <https://www.reinvestment.com/initiatives/hffi/>.
- [9] Randy Alison Aussenberg. SNAP and Related Nutrition Provisions of the 2014 Farm Bill (P.L. 113-79). Technical report, Congressional Research Service Report, 2014.
- [10] Illinois Facilities Fund. Cdfi fund awards \$4 million to iff. Online, September 2017. URL <https://iff.org/cdfi-fund-awards-4-million-iff/>.
- [11] Rachel Sacks, Stella S Yi, and Cathy Nonas. Increasing access to fruits and vegetables: perspectives from the New York City experience. *American Journal of Public Health*, 105(5):e29–e37, 2015.
- [12] Laura K Cobb, Cheryl AM Anderson, Lawrence Appel, Jessica Jones-Smith, Usama Bilal, Joel Gittelsohn, and Manuel Franco. Baltimore city stores increased the availability of healthy food after WIC policy change. *Health Affairs*, 34(11):1849–1857, 2015.
- [13] S Cummins, E Flint, and S A Matthews. New neighborhood grocery store increasing awareness of food access but did not alter dietary habits or obesity. *Health Affairs*, 33(2):283–291, 2014.
- [14] S Cummins, M Petticrew, C Higgings, A Findlay, and L Sparks. Large scale food retailing as an intervention for diet and health: quasi-experimental evaluation of a natural experiment. *J Epidemiol Community Healthy*, 59 (1035-1040), 2005.
- [15] N Wrigley, D Warm, and B Margetts. Deprivation, diet, and food-retail access: findings from the Leeds ‘food deserts’ study. *Environmental and Planning A*, 35:151–181, 2003.
- [16] Brian Elbel, Alyssa Moran, L Beth Dixon, Kamila Kiszko, Jonathan Cantor, Courtney Abrams, and Tod Mijanovich. Assessment of a government-subsidized supermarket in a high-need area on household food availability and children’s dietary intakes. *Public Health Nutrition*, 18(15):2881–2890, 2015.
- [17] Hee-Jung Song, Joel Gittelsohn, Miyong Kim, Sonali Suratkar, Sangita Sharma, and Jean 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*, 12(11):2060–2067, 2009.
- [18] Tamara Dubowitz, Madhumita Ghosh-Dastidar, Deborah A Cohen, Robin Beckman, Elizabeth D Steiner, Gerald P Hunter, Karen R Flórez, Christina Huang, Christine A Vaughan, Jennifer C Sloan, et al. Changes in diet after introduction of a full service supermarket in a food desert. *Health Affairs*, 34(11):1858, 2015.
- [19] Madhumita Ghosh-Dastidar, Gerald Hunter, Rebecca L Collins, Shannon N Zenk, Steven Cummins, Robin Beckman, Alvin K Nugroho, Jennifer C Sloan, Tamara Dubowitz, et al. Does opening a supermarket in a food desert change the food environment? *Health & Place*, 46:249–256, 2017.
- [20] Donald Rose and Rickelle Richards. Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. *Public Health Nutrition*, 2004.
- [21] Michele Ver Ploeg and Ilya Rahkovsky. Recent evidence on the effects of food store access on food choice and diet quality. *USDA Economic Research Service*, 2016.
- [22] Hunt Allcott, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell. Food deserts and the causes of nutritional inequality. *The Quarterly Journal of Economics*, 134(4):1793–1844, 2019.
- [23] Steven Cummins. The local food environment and health: some reflections from the United Kingdom. *American Journal of Public Health*, 93(4):521–521, 2003.
- [24] Dalia Stern, Whitney R Robinson, Shu Wen Ng, Penny Gordon-Larsen, and Barry M Popkin. U.S. household food shopping patterns: dynamic shifts since 2000 and socioeconomic predictors. *Health Affairs*, 34(11):1840–1848, 2015.
- [25] Paul Rosenbaum. *Observational Studies*. Springer, 2002.
- [26] Gad Allon, Awi Federgruen, and Margaret Pierson. How much is a reduction of your customers’ wait worth? an empirical study of the fast-food drive-thru industry based on structural estimation methods. *Manufacturing & Service Operations Management*, 13(4):489–507, 2011.
- [27] Raphael Thomadsen. Product positioning and competition: The role of location in the fast food industry. *Marketing Science*, 26(6):792–804, 2007.
- [28] Peter Davis. Spatial competition in retail markets: movie theaters. *The RAND Journal of Economics*, 37(4): 964–982, 2006.

- [29] Ram C Rao and Niladri Syam. Equilibrium price communication and unadvertised specials by competing supermarkets. *Marketing Science*, 20(1):61–81, 2001.
- [30] Alison Gustafson, James Allen, Nancy Schoenberg, and Mark Swanson. The relationship between neighborhood food environment and food store choice on purchasing habits among SNAP and lower income households, USDA FoodAPS data. In *University of Kentucky Center for Poverty Research Discussion Paper Series*. University of Kentucky, 2016.
- [31] Angela D Liese, Bethany A Bell, Timothy L Barnes, Natalie Colabianchi, James D Hibbert, Christine E Blake, and Darcy A Freedman. Environmental influences on fruit and vegetable intake: results from a path analytic model. *Public Health Nutrition*, 17(11):2595–2604, 2014.
- [32] M Beydoun, L Powell, and Y Wang. The association of fast food, fruit and vegetable prices with dietary intakes among us adults: Is there modification by family income? *Social Science and Medicine*, 66(11):2218–2229, 2008.
- [33] E.L. Gibson, J. Wardle, and C.J. Watts. Fruit and vegetable consumption, nutritional knowledge and beliefs in mothers and children. *Appetite*, 31(2):205 – 228, 1998. ISSN 0195-6663. doi: <https://doi.org/10.1006/appe.1998.0180>. URL <http://www.sciencedirect.com/science/article/pii/S0195666398901805>.
- [34] Inge Spronk, Charina Kullen, Catriona Burdon, and Helen O’Connor. Relationship between nutrition knowledge and dietary intake. *British Journal of Nutrition*, 111(10):1713–1726, 2014.
- [35] May A Beydoun and Youfa Wang. How do socio-economic status, perceived economic barriers and nutritional benefits affect quality of dietary intake among us adults? *European Journal of Clinical Nutrition*, 62(3):303, 2008.
- [36] Michele Ver Ploeg and Alana Rhone. Food access research atlas documentation. Technical report, USDA Economics Research Service, May 2017. URL <https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/#definitions>.
- [37] Charles Courtemanche and Art Carden. Supersizing supercenters? the impact of walmart supercenters on body mass index and obesity. *Journal of Urban Economics*, 69(2):165–181, 2011.
- [38] Richard Volpe, Abigail Okrent, and Ephraim Leibtag. The effect of supercenter-format stores on the healthfulness of consumers’ grocery purchases. *American Journal of Agricultural Economics*, 95(3):568–589, 2013.
- [39] Elizabeth Stuart and Donald Rubin. *Best Practices in Quasi-Experimental Designs: Matching methods for causal inference*, chapter 11. Sage Publications, 2007.
- [40] Teck-Hua Ho, Noah Lim, Sadat Reza, and Xiaoyu Xia. OM forum: Causal inference models in operations management. *Manufacturing & Service Operations Management*, 19(4):509–525, 2017.
- [41] P. Rosenbaum and D. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 1983.
- [42] Paul Rosenbaum. *Design of Observational Studies*. Springer, 2010.
- [43] D Rubin. Statistics and causal inference: Comment: Which ifs have causal answers. *Journal of the American Statistical Association*, 81:961, 1986.
- [44] Paul W Holland. Statistics and causal inference. *Journal of the American statistical Association*, 81(396):945–960, 1986.
- [45] Isobel R Contento. Nutrition education: linking research, theory, and practice. *Asia Pacific Journal of Clinical Nutrition*, 17(S1):176–179, 2008.
- [46] P Austin and E Stuart. Estimating the effect of treatment on binary outcomes using full matching on the propensity score. *Statistical Methods in Medical Research*, 2015.
- [47] Luke Keele and Luke Miratrix. Randomization inference for outcomes with clumping at zero. *The American Statistician*, pages 1–10, 2018.
- [48] Kosuke Imai and David A Van Dyk. Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association*, 99(467):854–866, 2004.
- [49] Austin Bradford Hill. The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*, 58(5), May 1965.
- [50] Ben B Hansen. Full matching in an observational study of coaching for the sat. *Journal of the American Statistical Association*, 99(467):609–618, 2004.
- [51] Colin B Fogarty, Pixu Shi, Mark E Mikkelsen, and Dylan S Small. Randomization inference and sensitivity analysis for composite null hypotheses with binary outcomes in matched observational studies. *Journal of the American Statistical Association*, 112(517):321–331, 2017.

- [52] Shandong Zhao, David van Dyk, and Kosuke Imai. Propensity-score based methods for causal inference in observational studies with fixed non-binary treatments. 2013.

A Outcome variable

A.1 Fruit and vegetable expenditures at restaurants

In this study we specifically focus on fruit and vegetable expenditures at all stores (including grocery stores, convenience stores, and other types of stores). However, we do not include FV spending at restaurants. This decision was made for a few reasons. First and foremost, it is difficult to precisely measure spending specifically on fruits and vegetables at restaurants. Second, the amount of fruits and vegetables bought at restaurants is very small compared the amount bought at grocery stores.

A.2 Missing expenditure data

In the FoodAPS dataset, across all FV items, the expenditure amounts are missing for 672 items (9% of all items). However, for all of these items a detailed USDA food category is present, for 575 of these items a detailed food description is present, and for 333 of these items the weight is present. The missing expenditures are thus imputed using the food item descriptions/categories and total grams (when present). An indicator variable is retained for the missing data to use in the matching.

A.3 Price of fruits and vegetables

Although our outcome of interest is household fruit and vegetable spending, we do not specifically consider the price of fruits and vegetables in our analysis. Variations in prices could bias the results if one were to believe that households with high value of nutrition or high access systematically buy more expensive fruits and vegetables. This would result in the appearance of these groups “consuming” more FVs when in fact they just spend more on the same FVs. If anything, we find the opposite to be true. By comparing the amount of money spent on equal grams of both bananas and tomatoes, we find that high value households actually tend to spend a bit less (about 4% less on bananas and 3% less on tomatoes). Households with high access spend about 9% less on tomatoes and 3% more on bananas. These numbers suggest that variations in the prices of fruits and vegetables are likely not biasing the results.

B Covariates

B.1 Measuring vehicle ownership

In this paper, the covariate `anyvehicle` is used to define vehicle ownership, which is the household’s response to the yes/no question “Does anyone in your household own or rent a vehicle?” However, in addition to the covariate `anyvehicle`, there is another covariate `primstoretravelmode` which is the primary respondent’s response to the question “How do you usually get to the store where you do most of your food shopping?” among the choices: drive own car, use someone else’s car, someone else drives me, walk, bus, taxi, bicycle. An anomaly in the survey is that there are 114 SNAP households (7.2% of the SNAP population) who answered “no” to `anyvehicle`, yet selected “drive own car” as the response to `primstoretravelmode`. In order to better understand whether these households likely own a vehicle or not, we impute their vehicle ownership status using their other covariates. Using this methodology, the imputation classifies 110 of these unknown households as not owning a vehicle, and 4 as owning a vehicle. This suggests that the covariates of this set of households is more similar to the set of households without vehicles than the set with vehicles.

As a robustness check, we replicate our analysis and count a household as owning a vehicle if they responded “yes” to `anyvehicle` *or* selected “drive own car” as the response to `primstoretravelmode`. In general the results do not change very much. The p-values for test statistics on subpopulations in which all households own a vehicle stay approximately the same, and the p-values for subpopulations which do not own a vehicle seem to uniformly increase very slightly. This slight increase could be due to the decrease in sample size for these subpopulations since we now classify 114 more households as owning a vehicle.

B.2 Measuring value of nutrition

We estimate value of nutrition based on certain behaviors that are asked about in the survey. In this study, we use five proxy variables which are, we believe, as close as we can come to measuring true value of nutrition given the data. The individual survey questions that are utilized are `nutritionsearch`, `nutritioneduc`, `nutritionfacts`, `mypyramidfollow`, and `myplatefollow`, which are described in Table 3. We classify a household as having a high value of nutrition if they fall into any one of the following five categories: 1) attended a nutrition education event in the

last two months, 2) searched the internet for nutrition information in the last two months, 3) try to follow MyPyramid guidelines, 4) try to follow MyPlate guidelines, or 5) report looking at nutrition labels “always” or “most of the time.” If a household did not fall into any of these categories, it was classified as having a low value of nutrition. This proxy results in 50.4% of all urban SNAP household being classified as having a high value of nutrition.

Because we use a proxy to estimate value of nutrition, it is necessary to understand issues that may arise from an inaccurate proxy. We believe that as long as the true effect of value of nutrition is non-negative (i.e., having a high value of nutrition will not decrease FV consumption), an inaccurate proxy will likely underestimate the effect of value of nutrition. This is because we believe it is more likely for households with a low value of nutrition to misreport their answers to the relevant survey questions in order to seem like they have a high value of nutrition, than vice versa. Therefore, it is likely that our proxy classifies some households as having a high value of nutrition when they really have a low value of nutrition. This type of misclassification would diminish the observed effect of value of nutrition on FV spending, since certain households in the “high value of nutrition” group actually have a low value of nutrition and thus on average purchase less FVs.

To validate this hypothesis, we seek evidence that those with a high value of nutrition are indeed estimated to have a high value of nutrition. One survey question asks the primary respondent whether they agree or disagree with the statement that “Healthy foods don’t taste good,” and another questions asks whether they agree or disagree that “People in my household don’t think that healthy food tastes good.” These questions are about healthy food in general, not just fruits and vegetables. However, if a household answered “agree” to both of these questions, then it may be reasonable to assume that if they *do* buy produce, they do so because they value eating healthy. We therefore look at FV spending among households that answered “agree” to both of these survey questions. In our proxy for value of nutrition is accurate, we expect households with a low estimated value of nutrition and answered “agree” to both questions (Group A) to buy very little produce because they 1) don’t like the taste of fresh food, and 2) don’t value nutrition. Furthermore, we would expect households estimated to have a high value of nutrition but also answered “agree” to both questions (Group B) to still purchase produce because they value nutrition. The median FV spending among households in Group A is \$0.833, and the median in Group B is \$2.223. Although this is not a rigorous analysis, it provides intuition that our proxy for value of nutrition is seemingly capturing true value of nutrition.

A post-hoc analysis on the impact of the individual covariates listed in Table 3 is performed in Appendix G.4.

C Los Angeles County

Figure 4 shows all low income food desert census tracts in Los Angeles County, overlayed with a proxy for value of nutrition based on city-level Google trend data. We use Google trend data for the topic “Calorie” for all available cities near Los Angeles (city is the most granular geographic level that Google trend makes public) over the last five years. The data consists of relative frequencies of searches that fall under the “Calorie” topic for several cities. The city with the most frequent calorie-related searches is scored 100, and all other cities are scored relative to 100. For example, a city with a score of 50 indicates that searches for calorie-related information occur half as frequently as the city which is scored 100. Not all cities are present in the data, which is why certain areas in Figure 4 are grey. These city-level relative search frequencies are used as a proxy for value of nutrition.

In order to conduct the analysis, we assume that census tracts have the same average value of nutrition as the city in which they lie. For example, a food desert that falls within a city with a score of 50 would receive a value of nutrition score of 50.

This method has many limitations. First, Google search data more granular than city-level is not available to the public. Therefore, this method may not be accurate for individual census tracts or neighborhoods. Second, search frequency of the “Calorie” topic may not be indicative of true value of nutrition. Therefore, policymakers should utilize multiple sources of data—include survey data—to create robust estimates of neighborhood-level value of nutrition.

D Statistical Analysis

D.1 Matching and covariates

Propensity score matching was performed using the R packages `MatchIt` and `OptMatch`. For each test, we began by performing a full match, limited to common support, and evaluated the resulting covariate balance. A full matching subsets households into matched sets in an optimal way while ensuring that each matched set contains at least one treated and one control unit. “Optimal” in this case means that the algorithm minimizes the total sum of propensity score distances within matched sets, as opposed to greedy nearest neighbor algorithms which do not have this guarantee^{25,50}.

After matching, if any covariate was largely unbalanced (in our case, if any covariate had an absolute standardized mean difference greater than 0.2 between the treatment and control group), propensity score calipers (caliper=.25) with Mahalanobis matching was used to better balance the imbalanced covariates. The balances for some of the matches are shown in Figures 5 and 6 (not all are shown for the sake of brevity).

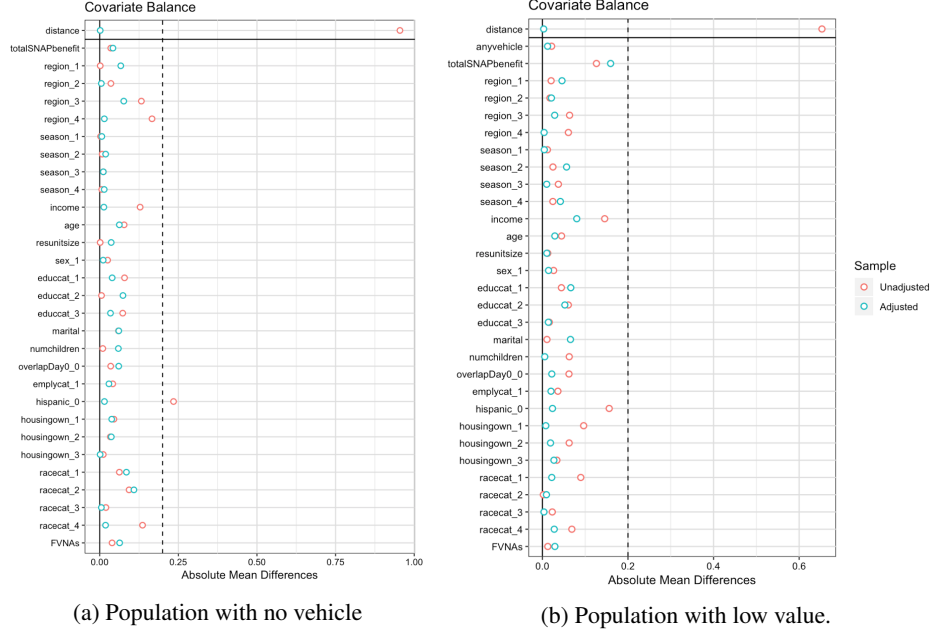


Figure 5: Plots of absolute standardized mean differences, where access is the treatment variable. In each plot, the pink dots show the standardized mean difference between the treated and control group without any adjustment. The blue dots show the standardized mean differences after matching. Note that the top row in each plot—labeled “distance”—is the propensity score, *not* store distance. Since we performed a full matching, these standardized mean differences were computed using weights for each household in accordance with their subclass.

Table 7 lists the covariates used in the matching. Most covariates listed in Table 1 were included, except those identified as potential “post-treatment” covariates. Post-treatment covariates are those that could be affected *by* the treatment, and including them in the matching can bias the results³⁹. For example, when examining the impact of access on FV spending, any covariates which could be impacted by a household’s access to grocery stores should not be included in the matching. This includes covariates such as the number of meals eaten out per week. Although covariates such as these were not included in the matching, their balance and potential for confounding can still be assessed in a post-hoc analysis.

First consider the number of meals eaten out per week. Among households with high access, the average is 1.09 meals per week, and among households with low access, the average is 1.00 meals per week (Table 1). After performing the matching, we again check the balance of this covariate and find that in all cases, there were no significant differences in this covariate between the treated and control groups.

Now consider the number of fast-food and other restaurants. It is possible that this covariate could explain the observed relationship between access and FV spending if one were to believe the following logic: households living far from grocery stores may have more fast food/restaurant options and therefore consume less FVs simply because of the availability of other options. Based on Table 1, this is likely not the case. First of all, as discussed above, there are no substantial differences between the number of meals eaten out per week for households with high versus low access. Furthermore, in the dataset, households with high access actually have far *more* fast food and restaurant options (Table 1). Therefore, if the above logic were true, then omitting this covariate would actually result in an *underestimation* of the impact of access on FV spending. Thus, we do not believe the availability of restaurants could explain the observed results.

D.2 P-values

Table 4 shows the pvals obtained for every test performed in the paper.

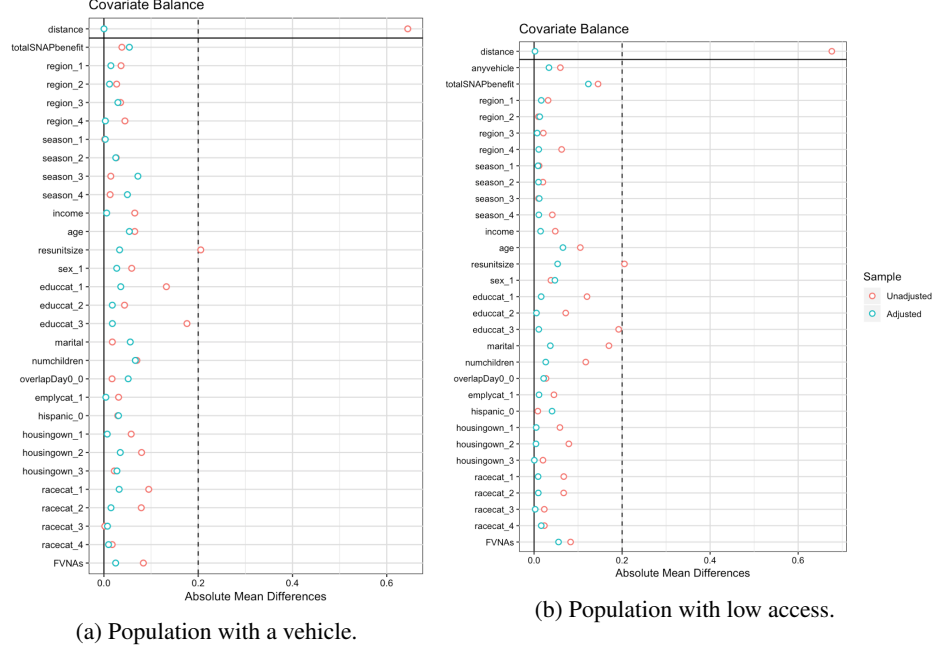


Figure 6: Plots of absolute standardized mean differences, where value of nutrition is the treatment variable. In each plot, the pink dots show the standardized mean difference between the treated and control group without any adjustment. The blue dots show the standardized mean differences after matching. Note that the top row in each plot—labeled “distance”—is the propensity score, *not* store distance. Since we performed a full matching, these standardized mean differences were computed using weights for each household in accordance with their subclass.

Figure 7: Covariates used in matchings

Household size	Number of children	Season	Region
Income	SNAP benefit amount	Benefit overlap	Race
Hispanic	Housing ownership	Employment	Education
Sex	Age	Marital status	# of missing prices of FV items
Vehicle ownership			

E Sensitivity analysis

For the sensitivity analysis we follow the methodology presented in Rosenbaum²⁵ which can be applied to a vast number of statistical tests when measuring the effect of a binary treatment (i.e. low versus high access, or high value of nutrition versus low value of nutrition). A sensitivity analysis asks the question: *How much would a hidden confounder need to impact treatment assignment in order to substantively alter the conclusions of our analysis?*

The magnitude of the impact of a hidden confounder on treatment assignment is parametrized by a single parameter Γ , which bounds the odds ratio of receiving treatment for two individuals with the same observed covariates. That is, if π_i is the probability that household i receives treatment, and \mathbf{x}_i are the observed covariates, then we enforce

$$\frac{1}{\Gamma} \leq \frac{\pi_k/(1 - \pi_k)}{\pi_l/(1 - \pi_l)} \leq \Gamma \text{ for } \mathbf{x}_l = \mathbf{x}_k \quad (1)$$

When $\Gamma = 1$, the odds ratio is 1 and both households have an equal chance of receiving treatment, meaning that the hidden confounder does not impact treatment assignment, given the observed covariates. Larger Γ allows for larger impacts of a hidden confounder on treatment assignment. For every study there is a Γ large enough such that the believed causal effects of the treatment are rendered insignificant.

For every Γ , there is a best- and worst-case treatment assignment allocation that results in lower and upper bounds for p-values, point estimates of treatment effects, and confidence intervals for the treatment effect (see Rosenbaum²⁵

for more details). A sensitivity analysis typically reports the largest Γ for which the upper p-value remains below the chosen significance level, which we call the “critical Γ ”. A critical Γ of 1.3 implies that it would take an unmeasured confounder that could alter the odds of receiving treatment in every stratum by at least 30% in order for the treatment effect to result in a p-value greater than 0.05. Larger critical Γ ’s indicate less potential for confounding, since a hidden confounder would need to have a larger impact in order to materially affect the results of the study. However, the potential for confounding also depends on whether such a confounder can be reasonably identified. That is, one hopes that the more covariates that are controlled for in the matching, the harder it is to identify a potential hidden confounder. These type of results leave the reader with a thought experiment: could there exist a confounder with an impact on treatment assignment of the specified magnitude?

For the tests with a continuous response (FV spending), we follow the classic methods presented in Rosenbaum²⁵, and for the tests with a binary response (whether or not the household visited a grocery store) we follow the methodology in Fogarty et al.⁵¹. The sensitivity analyses are performed in R using the script `compositeBinary.R` available at <http://www.mit.edu/~cfogarty/#software> for the sensitivity analysis with binary responses and `sensitivitySimple.R` for the sensitivity analysis of the continuous response.

The critical Γ ’s reported in Table 5 are the largest Γ such that the *largest possible* p-value is below 0.05. Looking at Table 5, certain results are clearly more robust than others according to the sensitivity analysis. For example, the impact of access on the population with a low value of nutrition and vehicle has a critical Γ of 1.258, indicating that a hidden confounder need to systematically alter the odds of living far from a grocery store by 25.8% in order to alter this result. However, the critical gamma for the impact of access on the population with a low value of nutrition is only 0.023. The sensitivity analysis is closely related to statistical power. Smaller sample sizes, or smaller effect sizes, will result in smaller critical Γ ’s²⁵. Therefore, we hypothesize that the small critical Γ of 1.023 is likely due to a small effect size.

Of course, in any observational study the possibility of a hidden confounder can never be ruled out. If a potential hidden confounder is identified, future studies should consider including such a confounder in the analysis. Because of the sensitivity of some of the results to hidden confounding, we also estimate the dose-response relationship between store distance and FV spending to help either confirm or reject the existence of a causal relationship. This is discussed in Appendix D.2.

F Dose-response relationship between access and FV spending

Identification of a dose-response relationship between store distance and FV spending is beneficial both to further elucidate the impact of access, as well as provide additional evidence in support of a causal relationship. We look for evidence of a dose-response relationship in two ways: first, by estimating the dose-response function using general additive models, proposed in Zhao et al.⁵². This is one method for generalizing propensity scores to continuous treatments. Second, we estimate the effect of “extreme distances” on FV spending.

F.1 Continuous effect of store distance

We provide a brief overview of the method employed, and refer details to Zhao et al.⁵² for more details. Suppose we have n households and p covariates for each household. Let \mathbf{X} be the $n \times p$ matrix of covariates, let \mathbf{Y} be the response vector, and let \mathbf{D} be the vector of distances to the nearest grocery store (i.e. the households’ treatment level). The goal is to estimate either the function $Y(d, X_i) = \mathbb{E}[Y(d)|X_i]$ (i.e., the outcome as a function of store distance given covariates) or the average of this function over all X , given as $Y(d) := \mathbb{E}_X[Y(d)]$. The fundamental problem is that for a given household we only observe $Y(D_i)$ and *not* the full function $Y(d)$.

The *propensity function* is defined as $e(d, X_i) = \mathbb{P}(D_i = d|X_i)$ which is the probability density function of store distance for a household given their covariates. We assume that this PDF can be fully parametrized by θ , and the only dependency of $e(d, X_i)$ on X_i is through $\theta(X_i)$. For example, if the propensity functions for each individual follows a Gaussian distribution with constant variance, $e(d, X_i) \sim N(\beta^T X_i, \sigma^2)$, so $\theta(X_i) = \beta^T X_i$.

To proceed, we make the standard assumption of strong ignorability of treatment assignment, which states that $\mathbb{P}(D|X, Y) = \mathbb{P}(D|X)$. Under this assumption, it can be shown that conditioning on $e(d, X_i)$ has the same nice properties as conditioning on the propensity score. First, the balancing property: $D_i \perp X_i | e(d, X_i)$. Second, strong ignorability of treatment assignment given the propensity function: $\mathbb{E}[Y(d)|D_i, e(d, X_i)] = \mathbb{E}[Y(d)|e(d, X)]$. In other words, the potential outcomes are independent of the treatment assignment given the propensity function.

Based on this second property, the function of interest is given by

$$\mathbb{E}[Y(d)] = \int \mathbb{E}[Y(d)|D_i, \theta] p(\theta) d\theta$$

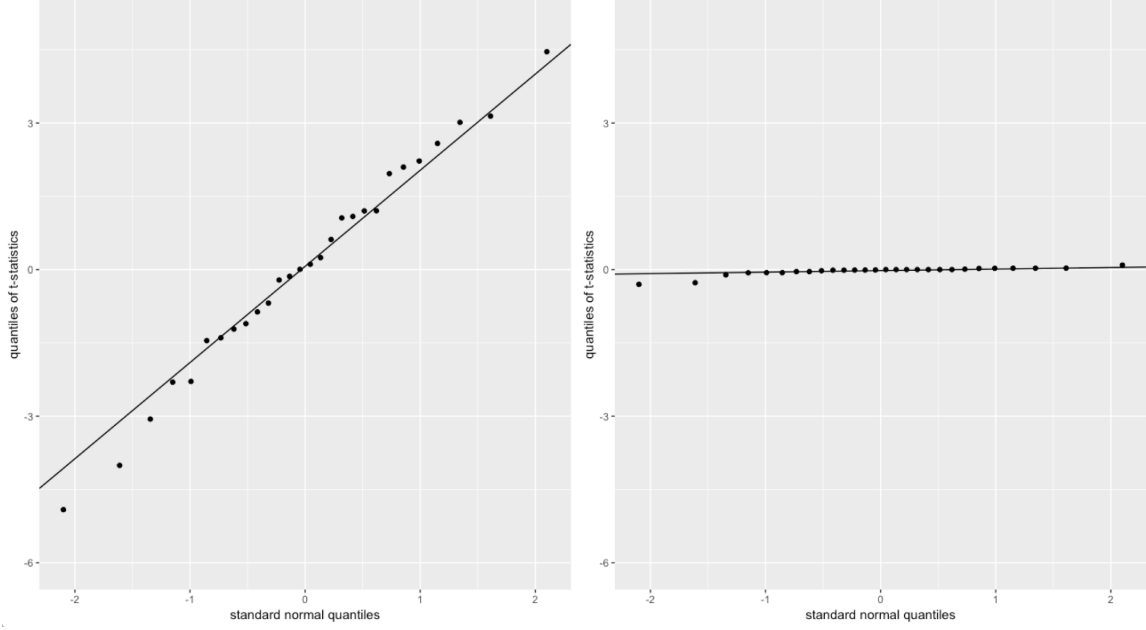


Figure 8: Quantile-quantile plots of the t-values for the coefficient on \mathbf{D} compared to a normal distribution. On the left, we show the results when the regressions do not adjust for $\hat{\theta}$ and on the right, we show the results when we adjust for $\hat{\theta}$.

where $\theta = \theta(\mathbf{X})$ uniquely defines the propensity function. We assume that $\mathbb{E}[Y(d)|D_i, \theta] = f(d, \theta)$ for some smooth function $f(\cdot)$ and compute the integral above using this function.

In practice, $e(d, X_i)$ is estimated by assuming a given parametric distribution and performing maximum likelihood estimation to find the parameters. Let $\hat{\theta}$ denote the estimated values of θ . To estimate $\mathbb{E}[Y(d)|D, \hat{\theta}]$ we assume that $\mathbb{E}[Y(d)|D, \hat{\theta}] = f(\hat{\theta}) + g(\hat{\theta}) * D$ where both f and g are smooth functions. These are typically estimated using penalized splines. We implement this using R's `mgcv` package. Finally, we can estimate $\mathbb{E}[Y(d)]$ using the empirical distribution of $\hat{\theta}$:

$$\mathbb{E}[Y(d)] = \frac{1}{n} \sum f(d, \hat{\theta}_i)$$

This method was found by Zhao et al.⁵² to outperform the standard method for propensity functions (which does not use general additive models), as well as the methods of generalized propensity score and inverse probability weighting.

First, we check our propensity function estimation. With a binary treatment, we checked our propensity score estimation by examining whether the underlying covariates in each subclass were balanced. In this case, we perform a similar check also based on the balancing property of the propensity function. The balancing property states that the treatment value is independent of the covariates given the propensity function. To check this, Imai and Van Dyk⁴⁸ suggests regressing each covariate on both \mathbf{D} and $\hat{\theta}$, and comparing the t-values of the coefficients on \mathbf{D} to a standard normal distribution. Figure 8 shows a Q-Q plot of the t-values of the coefficients for \mathbf{D} , both without adjusting for $\hat{\theta}$ (left) and after adjusting for $\hat{\theta}$ (right). It is clear that the adjustment results in t-values which are much closer to a normal distribution (although even without adjustment the t-values appear close to normal). This is the expected result if our propensity function estimation is indeed correct.

Second, we need to be wary of extrapolation using our estimation of $\mathbb{E}[Y(d)|D, \hat{\theta}]$. This is rather straightforward to check—we can simply plot \mathbf{D} versus $\hat{\theta}$ and look at the coverage. We should be cautious about extrapolation to areas where there is low coverage. This is shown in Figure 9. It appears that for distances under 1 mile, the range of $\hat{\theta}$ is relatively consistent (ranging from .25 to 1.0), however after 1 mile, the range of $\hat{\theta}$ is limited, indicating possible biases in extrapolation.

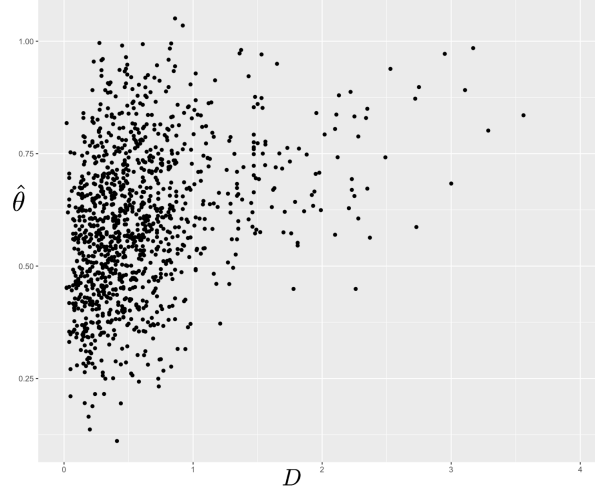


Figure 9: Plot of actual distances versus $\hat{\theta}$. It appears that for store distances greater than 1 mile, the estimated DRF may contain bias since the coverage of $\hat{\theta}$ for store distances over 1 mile is limited.

F.2 Extreme distances

To provide additional evidence for the causal effect of distance, we can compare people who live *very* close to the nearest grocery store to those who live *very* far. In the paper, we define the control group as those with store distances greater than τ and the treatment group as those with store distances less than τ , where τ is the chosen threshold. Now, we consider τ_1 and τ_2 such that $\tau_1 < \tau < \tau_2$, and define the control group as those with store distances greater than τ_2 and the treatment group as those with store distances under τ_1 . If we believe that the outcome (FV spending) is decreasing with distance, then we would expect to see a larger treatment effect when estimated using “extreme” distances.

If we do see such an effect, then in order for hidden biases to explain our initial results as well as these results, there would need to exist a confounder that is not only correlated with the original binary treatment, but that is also systematically increasing (or decreasing) with distance and also has an increasing (or decreasing) effect on FV spending and the probability of visiting the store.

For this analysis we define τ_1 to be the first quartile store distance (.3 miles for those with a vehicle and .2 miles for those without a vehicle), and τ_2 to be the third quartile store distance (.8 miles for those with a vehicle and .6 miles for those without a vehicle). In the paper we see that store distance has the largest estimated effect on households with either low value of nutrition, no vehicle, or both. Therefore, for this analysis we focus on these subpopulations. Using extreme distances, we obtain the results shown in Table 6. Since we are being more selective with what constitutes the control group and treated group, the sample sizes have greatly decreased for this analysis. Thus, we would not necessarily expect p-values to be smaller. Instead, we focus on the magnitude of the effect estimates. Comparing the estimated treatment effects using extreme distances to the those estimated using our original definition of high/low access, we see that extreme distances correspond to larger effects (Table 6). These results are in line with our hypothesis that distance has an *increasing effect* on fruit and vegetable spending.

G Post-hoc analyses

G.1 Distance thresholds

In this section we conduct a post-hoc analysis of the chosen distance threshold used to define “low access.” This analysis considers what would have happened if we had chosen different distance thresholds for defining “low access.” This exercise will not only tell us what the results of this study *would have been* if different distance thresholds had been chosen, but will also inform future studies and intervention design by elucidating the importance of choosing appropriate distance thresholds.

We focus specifically on the test that was conducted for estimating the impact of access on households with a low value of nutrition. In the study, “low access” was defined as residing more than 0.5 or 0.4 miles from a large grocery store, depending on whether the household owns a vehicle. In this post-hoc study, the analysis from the main study

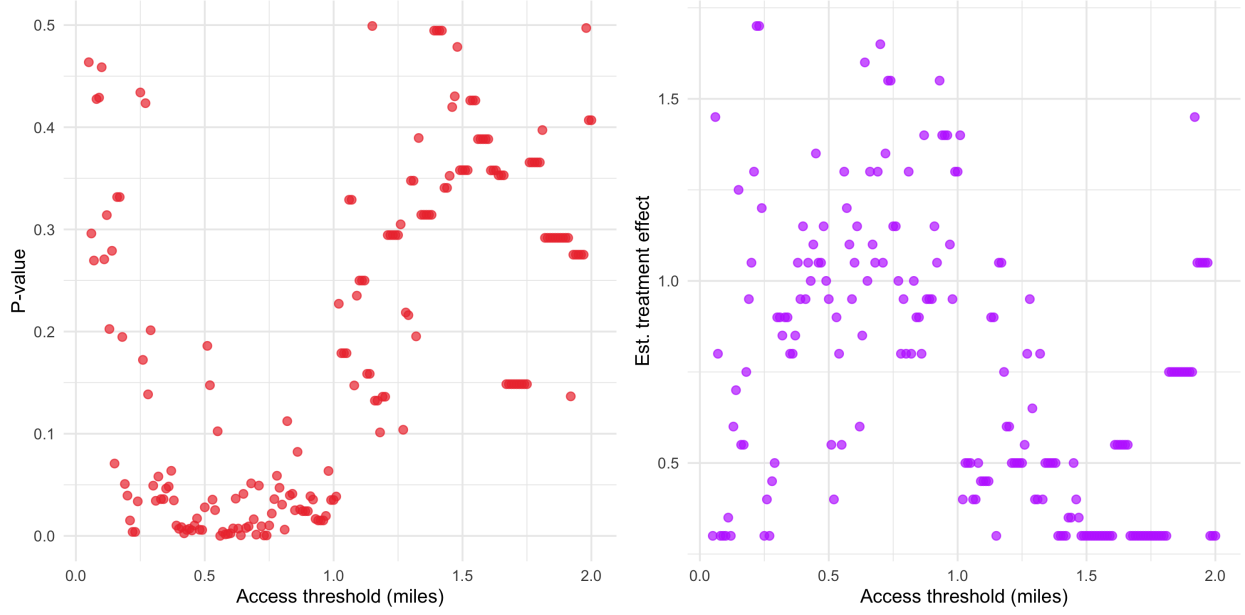


Figure 10: Post-hoc analysis of various distance thresholds that could have been used to define “low access” and “high access”.

is replicated for varying choices of the distance threshold. For each threshold, households are classified as either “low access” or “high access” based on the given threshold. Next, propensity score matching is performed. Finally, a stratified Wilcoxon test is performed to test for a significant effect of access on FV spending, and the Hodges-Lehman point estimate is also calculated. These results, for threshold ranging between 0.05-2.0 miles, is shown in Figure 10.

The smallest p-values, and largest effect estimates, occur with thresholds between 0.4 and 0.75 miles. In Figure 10, the same matching procedure is used for every threshold. These results largely align with the estimated dose-response function shown in Figure 1, which shows that the impact of distance appears to taper off around 1.0 miles. If a threshold over 1.0 mile is chosen, it is unlikely that the impact of access can be detected.

G.2 Public Transportation

One potential issue in the study is that the presence of public transportation is not considered. Since the locations of the households contained in the survey is not known, nor is the presence of public transportation recorded in the dataset, this covariate could not be systematically controlled for in the analysis. However, the survey does ask each household about their usual means of transport to their primary food store. This question contains the possible responses: drive own car, drive someone else’s car, someone else drives me, walk, bus, taxi, ride bicycle, and other.

This can be used, post-hoc, to better understand whether the presence of public transportation could be biasing the results of the study. Two possible ways this could occur are:

1. Perhaps households with low access and no vehicle who live near public transit should be treated as if they own a vehicle (i.e., classifying these households as “no vehicle” maybe a misrepresentation).
2. Depending on the efficiency of public transportation, it is possible that some households with seemingly low access could actually arrive at the grocery store quite quickly by using public transportation.

In order for public transportation to bias the results of the study, households must actually use public transportation to conduct their food shopping. In the dataset, only 5.3% of urban SNAP households report using a bus, taxi or “other” to usually conduct their food shopping. Therefore, it is unlikely that public transportation is not significantly biasing the results of the study.

However, to be sure, a post-hoc analysis is conducted that repeats the study after excluding any households that report using a bus, taxi, or “other” means of transportation to grocery shop. Table 7 reports some of the results obtained by excluding households who use public transportation. The impact of access on FV spending appears to be slightly larger than in the original study, although the results are quite similar. This indicates that public transportation may be

important to consider in the definition of access, especially in areas where many households use public transportation to grocery shop. In this study, very few households use public transportation to grocery shop, so its impact is limited.

G.3 Impact of superstores

Because some studies have found a negative impact of supercenters on diet and health, this post-hoc analysis removes superstores from the definition of a “large grocery store”^{37,38}.

In the original study, access is defined as the household’s minimum distance to a large grocery store, where a “large grocery” store was defined as a supermarket, medium or large grocery store, or supercenter. This aligns with the typical definition of access used at the federal level as well as in the literature. A superstore is the closest “large supermarket” for 32% of the urban, SNAP households in the dataset. Among these households, excluding superstores from the definition of “store distance” results in a minimum store distances that is typically 0.3 miles longer. Therefore, removing superstores from the definition of “large grocery store” substantially increases the minimum store distance for almost 1/3 of the households in the dataset, and could therefore significantly alter the results.

If the presence of superstores causes less (or no change in) FV spending, then removing superstores from the definition of access should increase the observed effect of access. The results—estimated treatment effects and p-values—are shown in Table 8 for certain subpopulations. These results, surprisingly, are not substantially different than the results of the main study. With respect to the impact of access, excluding superstores results in a slightly larger effect estimate among households with low value of nutrition. The estimated effect of value of nutrition on total FV spending as well as the likelihood of a store visit in a given week are also slightly larger with the exclusion of superstores.

G.4 Measures of value of nutrition

In this post-hoc additional analysis, we look at the specific covariates that were used to estimate value of nutrition, in order to better understanding which ones are the most impactful. The covariates are listed in Table 3. First, consider the correlations between each of the covariates, shown in Figure 11. Figure 11 shows that all covariates used in the construction of value of nutrition are positively correlated, which is unsurprising. Furthermore, Figure 11 shows that the covariates `nutritionfacts` and `nutritionsearch` seem to have the strongest correlation with the other covariates, which may suggest that these covariates are the best measures of value of nutrition.

However, the strong correlation also may suggest that these two covariates (`nutritionfacts` and `nutritionsearch`) are simply exaggerated or misreported more often than the others. For example, if many households who report the other three behaviors (those associated with `nutritioneduc`, `myplatefollow`, and `mypyramidfollow`) also tend to report the behaviors associated with `nutritionfacts` and `nutritionsearch`, then we would observe a strong correlation between `nutritionfacts` and `nutritionsearch` and the other three covariates.

Therefore, from the correlation table it is not clear whether `nutritionfacts` and `nutritionsearch` are the best indicators of value of nutrition or are simply the most likely to be misreported. We hypothesize that these two covariates—`nutritionfacts` and `nutritionsearch`—will either have the strongest or weakest effect on FV spending out of all five covariates.

Table 9 shows the weighted average FV spending among households who report each of the behaviors of the five covariates. These weighted averages are calculated by using the matched dataset, and associated weights, obtained in the original study for testing the impact of value of nutrition on all urban SNAP households. The covariates `nutritioneduc` and `myplatefollow` are associated with the largest average FV spending, and the covariates `nutritionsearch` and `nutritionfacts` are associated with the lowest. Furthermore the behaviors associated with the covariates `nutritionsearch` and `nutritionfacts` are the most commonly reported value-related behaviors. Fifty-eight percent of households with a high value of nutrition report reading nutrition facts “always” or “most of the time”, and 45% of households with a high value of nutrition report searching for nutrition-related information online. We hypothesize that these two behaviors are the most commonly over-reported.

We also hypothesize that the behaviors associated with `nutritioneduc` and `myplatefollow` are the most unlikely to be over-reported. Whether or not a household has taken a nutrition education class could, in theory, be verified, and therefore households may not misreport this behavior as often. Furthermore, it is likely that many households who do not use MyPlate have not even heard of MyPlate, and would therefore be unlikely to misreport this behavior. MyPlate replaced MyPyramid in 2011, and therefore many households may be unfamiliar with MyPlate since it is most commonly taught to children in school. Therefore, the covariates `nutritioneduc` and `myplatefollow` appear to be the best indicators of value of nutrition. These covariates should be included in future nutrition- and food-policy related studies and used as proxies for value of nutrition.

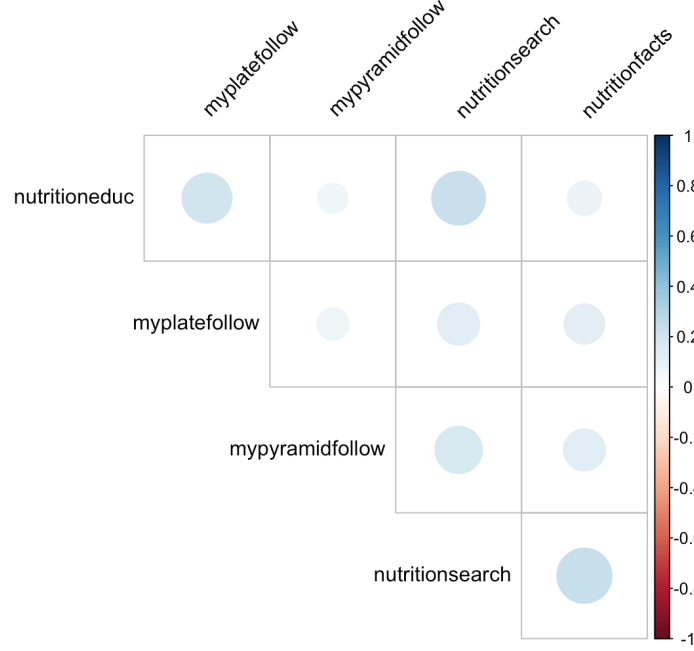


Figure 11: Correlation between variables used to construct value of nutrition

G.5 Substitution effects

In this section we explore which food groups, if any, households are likely to spend *less* money on if they were to increase their spending on fruits and vegetables. Note that it is possible for households to increase FV spending without decreasing spending in another food category, by allocating more of their own (non-SNAP) money towards food purchasing.

In order to address this question, we consider two approaches. First, we look for patterns in food expenditures across households as the absolute amount of money spent on FVs increases. Second, we consider patterns in the portion of food expenditures dedicated to different food groups as the portion spent on FVs increases. The food groups that are considered are: Fruits and Vegetables, Protein, Grains, Beverages, Snacks/sweets, and Dairy.

We find that absolute spending in all food groups is positively correlated. Therefore, households who spend more on FVs do not generally spend less than other households in any particular food category. This is likely due to the fact that households with higher incomes are likely to buy more food in general, and also may tend to eat healthier. Thus, by looking at this correlation it is difficult to predict which items households would move away from if they were to increase their spending on FVs, without an increase in total food budget.

However, we can normalize spending by total food expenditures, and consider the *percentage* of total food expenditures households allocate towards different food groups. Allocation of spending on Beverages has the strongest negative correlation to allocation of spending on FVs. The correlation with allocation of spending on Proteins and Snacks/sweets is also strong and negative. Therefore, households that allocate a larger *percentage* of their food budget to FVs are more likely to spend less (percentage-wise) on Beverages, Proteins and Snacks/sweets. Allocations to the Dairy and Grains food groups have the weakest correlation to FV allocation. This suggests that households may move away from unhealthy foods (Beverages and Snacks/sweets) when they spend more on FVs, or may also move away from other expensive items (such as Proteins). Spending on staple products such as Dairy and Grains is not largely impacted by spending on FVs.

Table 1: Descriptive statistics for outcomes and covariates

Variable	High access [†]	Low access [†]	p -value [‡] (if < .1)	High access	Low access
Number of households	575	583			
	Mean			Std dev	
Household Income (monthly)*	2,157	2,033		2,113	2,173
Closest large store (miles)	0.27	0.916	$< 2e^{-16}$	0.12	0.54
# Fast food within 1 mile	9.61	5.34	$< 2e^{-16}$	6.88	4.77
# Non-fast food restaurants within 1 mile	50.3	19.6	$< 2e^{-16}$	69.1	20.2
Standardized FV spending	4.07	3.83		5.78	6.16
Household SNAP benefit*	247	252		190	198
# Meals eaten out per week	1.09	1.00	0.066	1.33	1.31
# Of large store visits	2.65	2.23	0.0006	2.19	1.97
Household size*	3.56	3.40		2.00	1.90
Number of children*	1.58	1.55		1.61	1.58
	Proportion				
Visited a large grocery store	.866	.860	0.006		
Value of nutrition	.477	.513			
Whether SNAP benefit receipt overlapped with audit week*	.231	.222			
Season during survey*					
Winter	.038	.022			
Spring	.177	.190			
Fall	.437	.393			
Summer	.348	.395			
Region*					
Northeast	.181	.153			
Midwest	.198	.189			
South	.344	.408			
West	.277	.250			
Race (PR [§])*			$2e^{-4}$		
White	.521	.623			
Black	.243	.240			
Asian	.162	.089			
Other	.073	.048			
Hispanic (PR)*	.381	.232	$5e^{-8}$		
Housing ownership*			$6e^{-4}$		
Rent	.770	.674			
Own	.198	.266			
Employment (PR)*					
Working	.306	.305			
Looking for work	.151	.155			
Not working	.514	.507			
Education (PR)*					
≤ 10th grade	.202	.161			
11th or 12th grade, no diploma	.115	.108			
H.S. diploma or G.E.D	.299	.330			
Some college or associate's degree	.301	.328			
Bachelor's degree	.070	.062			
Master's degree	.012	.012			
Male*	.193	.209			
Age (PR)*					
18 yrs	.003	.002			
19 yrs	.010	.005			
20-35 yrs	.388	.353			
36-59 yrs	.471	.489			
60-65 yrs	.056	.079			
66-70 yrs	.033	.027			
≥ 71 yrs	.038	.045			
Marital status*			0.009		
Married	.247	.245			
Widowed	.057	.082			
Divorced	.176	.232			
Separated	.097	.108			
Never married	.423	.333			
Any vehicle*	.668	.707			

* Indicates a covariates that was used in the propensity-score matching.

[†] "Low access" indicates a store distance greater than 0.5 or 0.4 miles for households with and without a vehicle, respectively. Households that are not low access have "high access."

[‡] p -value is for the difference between the Short and Long values.

[§] PR = Primary Respondent—the household member who filled out the survey and does the majority of the households' grocery shopping.

Table 2: Estimated effects of access and value of nutrition

	Population	Access				Value of nutrition			
		d.1	d.2	d.3	d.4	v.1	v.2	v.3	v.4
		τ^{\S}	δ^{\P}	$\tau_{pv}^{\dagger\dagger}$	$N^{\ddagger\dagger}$	τ^{\S}	δ^{\P}	$\tau_{pv}^{\dagger\dagger}$	$N^{\ddagger\dagger}$
1.	Overall	0.24	0.08**	< 0	1,149	0.97***	0.03	0.20*	1,153
2.	No vehicle	1.25**	0.15**	< 0	354	0.55	0.05	0.20	347
3.	Vehicle	0.24	0.05*	< 0	785	0.85**	0.03	0.32*	789
4.	Low value	0.76*	0.09**	0.02	583	—	—	—	—
5.	High value	< 0	0.02	< 0	438 ^{§§}	—	—	—	—
6.	Low access [†]	—	—	—	—	1.32***	0.08*	0.39**	574
7.	High access [†]	—	—	—	—	0.55	0.0	0.09	565
8.	Vehicle, LV [‡]	1.76**	0.12**	< 0	386	—	—	—	—
9.	No vehicle, LV [‡]	2.3	0.20**	< 0	174	—	—	—	—
10.	Vehicle, HV [‡]	< 0	0.02	< 0	401	—	—	—	—
11.	No vehicle, HV [‡]	< 0	0.13	< 0	108 ^{§§}	—	—	—	—
12.	Vehicle, High access	—	—	—	—	0.29	< 0	0.16	366
13.	No vehicle, High access	—	—	—	—	0.11	.05	< 0	186
14.	Vehicle, Low access	—	—	—	—	1.36**	0.09*	0.42	407
15.	No vehicle, Low access	—	—	—	—	1.93*	0.07	0.89	145

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, exact p-values given in Table 4

[†] “Low access” is defined as a store distance greater than 0.5 miles with a vehicle, or 0.4 miles without a vehicle.

[‡] LV=Low value, HV=High value

[§] τ is the estimated Tobit effect of distance on FV spending across all households in the given populations.

[¶] δ is the estimated risk difference for a grocery store visit.

^{††} τ_{pv} is the estimated Tobit effect of distance on FV spending *per store visit* across all households in the given subgroups who made at least one trip to the grocery store.

^{‡‡} N shows number of units included in the matching to give a sense for the size of each subpopulation

^{§§} Indicates nearest neighbor PSM with caliper matching using Mahalanobis distance.

Table 3: Description of covariates used as proxy for value of healthy food.

Variable name	Type	Descripton
nutritionsearch	Binary	Whether a households searched for nutritional information online in the past two months
nutritioneduc	Binary	Whether a households participated in a nutrition education event in the past two months
nutritionfacts	Ordinal (1-5)	How often household uses Nutrition Facts panel
myplatefollow	Binary	Whether households tried to follow MyPlate recommendations
mypyramidfollow	Binary	Whether households tried to follow MyPyramid recommendations

Table 4: P-values

Population		Access			Value of nutrition		
		τ	δ	τ_{pv}	τ	δ	τ_{pv}
1.	Overall	0.237	0.001	0.496	< 0.001	0.114	0.046
2.	No vehicle	0.006	0.002	0.796	0.148	0.182	0.246
3.	Vehicle	0.275	0.048	0.833	0.003	0.126	0.017
4.	Low value	0.037	0.009	0.402	—	—	—
5.	High value	0.141	0.300	0.766	—	—	—
6.	Low access	—	—	—	0.005	0.007	0.007
7.	High access	—	—	—	0.327	0.118	0.219
8.	Vehicle, LV	0.004	0.005	0.594	—	—	—
9.	No vehicle, LV	0.062	0.009	0.566	—	—	—
10.	Vehicle, HV	0.675	0.310	0.950	—	—	—
11.	No vehicle, HV	0.886	0.063	0.714	—	—	—
12.	Vehicle, High access	—	—	—	0.268	0.549	0.158
13.	No vehicle, High access	—	—	—	0.487	0.231	0.799
14.	Vehicle, Low access	—	—	—	0.004	0.024	0.062
15.	No vehicle, Low access	—	—	—	0.034	0.196	0.055

Table 5: critical Γ for tests when p -value < .05

Population		Access			Value of nutrition		
		τ	δ	τ_{pv}	τ	δ	τ_{pv}
1.	Overall	—	1.310	—	1.213	—	1.005
2.	No vehicle	1.218	1.439	—	—	—	—
3.	Vehicle	—	1.004	—	1.159	—	1.075
4.	Low value	1.023	1.177	—	—	—	—
5.	High value	—	—	—	—	—	—
6.	Low access	—	—	—	1.286	1.171	1.152
7.	High access	—	—	—	—	—	—
8.	Vehicle, LV	1.258	1.339	—	—	—	—
9.	No vehicle, LV	—	1.459	—	—	—	—
10.	Vehicle, HV	—	—	—	—	—	—
11.	No vehicle, HV	—	—	—	—	—	—
12.	Vehicle, High access	—	—	—	—	—	—
13.	No vehicle, High access	—	—	—	—	—	—
14.	Vehicle, Low access	—	—	—	1.224	1.111	—
15.	No vehicle, Low access	—	—	—	1.057	—	—

Table 6: τ for the effect of distance on FV spending using extreme doses (ED). The last row shows the estimated effect sizes using the original treatment definitions (RD).

	No vehicle	Low value
τ (ED)	2.24	1.72
p-value	.016	.0005
N	164	279
τ (RD)	1.25	.76

Table 7: Effect estimates and p-values (in parentheses, if less than 0.1) for the effect of access on FV spending, excluding households who use public transportation

Population	Access			Value of nutrition		
	τ	δ	τ_{pv}	τ	δ	τ_{pv}
Overall	0.52 (.062)	0.07 (.004)	<0	1.26 (< .0001)	.04 (.060)	0.28 (.007)
Low value	1.29 (.001)	0.07 (.041)	0.12			
Low access				1.23 (.014)	0.08 (.026)	0.35 (.014)

Table 8: Effect estimates and p-values (in parentheses, if less than 0.1) for the effect of access on FV spending, without superstores

Population	Access			Value of nutrition		
	τ	δ	τ_{pv}	τ	δ	τ_{pv}
Overall	0.20	0.06 (.006)	0.0	0.97 (.0004)	.03	.20 (.046)
Low value	0.92 (.037)	0.0 (.009)	0.50			
Low access				1.68 (.001)	0.10 (.007)	0.38 (.020)

Table 9: Average FV spending for households who report the behaviors associated with each covariate used to measure value of nutrition.

Variable name	Weighted avg FV spending	% Among high val. hh's
nutritionsearch	\$3.78	45%
nutritioneduc	\$5.17	13%
nutritionfacts	\$3.73	58%
myplatefollow	\$5.20	33%
mypyramidfollow	\$4.37	16%
Low value of nutrition	\$2.45	—