The Effect of SNAP on the Composition of Purchased Foods: Evidence and Implications

Justine Hastings
Brown University and NBER

Ryan Kessler
Brown University

Jesse M. Shapiro*
Brown University and NBER

March 2020

Abstract

We use detailed data from a large retail panel to study the effect of participation in the Supplemental Nutrition Assistance Program (SNAP) on the composition and nutrient content of foods purchased for at-home consumption. We find that the effect of SNAP participation is small relative to the cross-sectional variation in most of the outcomes we consider. Estimates from a model relating the composition of a household's food purchases to the household's current level of food spending imply that closing the gap in food spending between high- and low-SES households would not close the gap in summary measures of food healthfulness.

Keywords: SNAP, nutrition, diet
JEL: D12, H31, I12, I38

*This work has been supported (in part) by awards from the Laura and John Arnold Foundation, the National Science Foundation under Grant No. 1658037, the Robert Wood Johnson Foundation's Policies for Action program, the Russell Sage Foundation, and the Smith Richardson Foundation. Any opinions expressed are those of the authors alone and should not be construed as representing the opinions of these Foundations. We also appreciate support from the Population Studies and Training Center at Brown University. We thank Dan Björkegren and Stefan Hut for help in obtaining some of the product nutrition information used in this project, Jessica Todd and Shelly Ver Ploeg for providing additional information regarding the estimates in their article, Sarah Bleich and Alyssa Moran for helping us to understand summary measures of food healthfulness, and participants at math.stackexchange.com for help with matrix calculus. This project has benefited from the comments of Chloe East, discussant Ariella Kahn-Lang Spitzer, and seminar audiences at the NBER Public Economics Meeting, Boston University's Population Health Science Research Workshop, University of Maryland, University of Nebraska Lincoln, Yale University, the London School of Economics, University College London, the Philadelphia Fed, UC Berkeley, the National Tax Association, and the University of Illinois at Urbana-Champaign. We thank our dedicated research assistants for their contributions. All estimates and analyses in this paper based on Information Resources, Inc. data are by the authors and not by Information Resources, Inc.

E-mail: justine_hastings@brown.edu, ryan_kessler@brown.edu, jesse_shapiro_1@brown.edu.

1 Introduction

The Supplemental Nutrition Assistance Program (SNAP) is the second-largest means-tested program in the United States (Falk et al. 2018), enrolling 18.6 percent of households in the average month of fiscal 2016.[1] The program provides households with an electronic benefit transfer (EBT) card that can be used to purchase food for at-home consumption at participating retailers.

Among the stated aims of SNAP are to reduce hunger and to improve nutrition by allowing households to spend more on food.[2] Consistent with the latter aim, a number of studies find a positive association between grocery spending and markers of diet quality (see, for example, Mabli et al. 2010; Anderson and Butcher 2016).[3] Motivated in part by this evidence, recent policy reports advocate increasing SNAP benefits or enrollment as a way to improve diet-related health (Hartline-Grafton 2013; Woolf et al. 2013; Harvard T. H. Chan School of Public Health 2017).[4] Policies that reduce socioeconomic disparities in food expenditure may also reduce socioeconomic disparities in diet quality (Rehm et al. 2011; Monsivais et al. 2012).

The analysis in this paper has two main objectives. The first is to estimate the effect of SNAP participation on the composition and nutrient content of foods purchased for at-home consumption. The second is to estimate the contribution of differences in food-at-home spending to socioeconomic differences in measures of food healthfulness.

Our analysis uses detailed transaction records from February 2006 through December 2012 for nearly half a million regular customers of a large US grocery retailer. Hastings and Shapiro (2018) use these data to estimate the effect of SNAP participation on food spending. The data contain information on method of payment, which we use to infer participation in SNAP. The data contain identifiers for products purchased, which we join to information from several sources on food types and nutrient content. The resulting panel allows us to track the composition and nu-

[1] There were 21,777,938 participating households in the average month of fiscal 2016 (FNS 2016) and 116,926,305 households in the US on average from 2011-2015 (US Census Bureau 2016).

[2] The Food and Nutrition Act of 2008, which created SNAP as the successor to the Food Stamp Program, states, "To alleviate ... hunger and malnutrition, a supplemental nutrition assistance program is herein authorized which will permit low-income households to obtain a more nutritious diet through normal channels of trade by increasing food purchasing power for all eligible households who apply for participation."

[3] There is also evidence that healthier diets cost more (Jetter and Cassady 2006; Aggarwal et al. 2012; Rao et al. 2013; Rehm et al. 2015).

[4] For example, Hartline-Grafton (2013) cites the positive association between grocery spending and diet quality documented in Mabli et al. (2010) as evidence that "more adequate [SNAP] benefits improve dietary quality" (p. 6).

trient content of households' grocery purchases at the retailer over nearly seven years, including many thousands of transitions on to and off of SNAP.

Our outcome measures include the share of kilocalories devoted to different types of foods (e.g., fruits and non-starchy vegetables) and the ratio of different nutrients (e.g., fat) to total kilocalories. We also consider two summary measures: a nutrient density score (NDS) measuring compliance with the Food and Drug Administration's (FDA) Daily Value (DV) bounds (Hansen et al. 1979; Fulgoni et al. 2009; Drewnowski and Fulgoni 2014), and the 2010 version of the Healthy Eating Index (HEI-2010) measuring compliance with the USDA's 2010 Dietary Guidelines for Americans (Guenther et al. 2014).

We adapt two research designs from Hastings and Shapiro (2018) to estimate the causal effect of SNAP participation on these outcomes. The first is a panel event-study design exploiting the fine timing of entry into SNAP and using a proxy for income to control for the endogeneity of program entry (Freyaldenhoven et al. 2019). The second is a quasi-experimental design exploiting plausibly exogenous variation in the timing of program exit driven by the fact that the lengths of SNAP spells are often divisible by six months (Klerman and Danielson 2011; Mills et al. 2014; Scherpf and Cerf 2019; Gray 2019).

We find that the effect of SNAP is small relative to the cross-sectional variation in most of the outcome measures we consider. For example, using our first research design, we estimate that SNAP reduces the share of kilocalories from fruits and non-starchy vegetables by 0.0009, with a standard error of 0.0004. The cross-sectional interquartile range (IQR) of the average share of kilocalories from fruits and non-starchy vegetables across all households in the retail panel is 0.031, two orders of magnitude greater than the estimated effect. Likewise, we estimate that SNAP increases the ratio of kilocalories from total fat to total kilocalories by 0.014 of its DV upper bound, with a standard error of 0.0033. This estimate can be compared to an IQR of 0.19.

Turning to our summary measures, using our first research design we estimate that SNAP reduces healthfulness as measured by the NDS by 0.009 (with a standard error of 0.004) and increases healthfulness as measured by the HEI-2010 by 0.173 (with a standard error of 0.14). These estimates are of lower order than the respective IQRs of 0.29 and 10.2. We show that these estimates are small when compared to a variety of other cross-sectional and time-series benchmarks drawn both from our own calculations and from the literature. Estimates from our second research design imply more negative effects of SNAP participation, with effects on the NDS and HEI-2010 of -0.040 and -0.232, respectively, but with less statistical precision: standard errors on these estimates are 0.013 and 0.481, respectively.

We use our two research designs to estimate a model in which the healthfulness of a household's food purchases depends on the household's contemporaneous food spending. We use the estimated model to simulate the effect on food healthfulness of closing the spending gap between lower and higher socioeconomic status (SES) households. We find that closing the socioeconomic gap in mean spending would widen the gap in mean NDS by 4.7 percent (with a standard error of 4.5 percent) and narrow the gap in mean HEI-2010 by 3.6 percent (with a standard error of 3.0 percent).

Our analysis benefits from the length and detail of the retail panel, but it is also limited by some aspects of the data that are worth noting. First, because our data do not come from a representative sample of households, we cannot (without additional assumptions) generalize our findings to the broader population of SNAP recipients. As one way to assess representativeness, we compare the distribution of our outcome variables between our data and nationally representative survey data. The two distributions are similar for many outcomes, but we also highlight some differences.

Second, because our data come from a single retail chain, we only observe a portion of households' purchases for at-home consumption. Hastings and Shapiro (2018) provide evidence that panelists devote a large share of their grocery budget to the retailer, and use longitudinal data from the Nielsen Homescan Consumer Panel (NHCP) to show that SNAP participation is not associated with significant changes in choice of retailer. In the online Appendix, we revisit the NHCP and find little evidence of a longitudinal association between SNAP participation and measures of food healthfulness, consistent with our estimates based on the retail panel. While these findings are reassuring, they do not exclude the possibility that the changes we observe in the retail data do not reflect the full effect of SNAP participation on the composition of foods purchased for at-home consumption.

Third, because our data measure purchases rather than consumption, we cannot make definitive statements about what food is eaten. We focus on intensive measures of food healthfulness such as ratios of nutrients to kilocalories purchased so that our inferences are valid under the assumption that wastage (and shopping at other retailers) affects all inputs to these measures in equal proportion.

Fourth, because our data come from a grocery retailer, we cannot study the effect of SNAP participation on food consumed away from home (FAFH),[5] which research has found to be less healthy than food purchased for at-home consumption (FAH).[6] Recent research does not consistently find evidence of an effect of SNAP on FAFH expenditure.[7] If SNAP participation were to decrease consumption of FAFH, this could lead to an improvement in overall healthfulness of consumed foods.

Fifth, our data are best suited to studying the effects of SNAP on relatively short time horizons. We estimate effects at both monthly and quarterly levels of aggregation, and our event-study plots report estimated effects up to four quarters (one year) after the quarter of program adoption. Our data and research designs do not permit informative inference on effects of SNAP over very long periods (e.g., decades). Although SNAP is primarily a safety-net program (FNS 2012, p. 1), it has a nutrition education component (FNS 2012, p. 24), and long-term effects may be important in light of evidence that household diet choices are persistent in the short run (see, e.g., Hut 2019 and the references therein).

This paper contributes to a large literature on the effects of SNAP and the predecessor Food Stamp Program on diet quality.[8] Our main contribution to this literature is a set of relatively precise estimates of the causal effect of SNAP on several measures of the composition of foods

[5]We estimate in survey data that FAFH accounts for 24 percent of food spending and 19 percent of calorie acquisitions for SNAP recipients. Clay et al. (2016) find that FAFH accounts for 22 percent of food spending among SNAP households in the Consumer Expenditure Survey (CE) and 18 percent of food spending among SNAP households in the National Health and Nutrition Examination Survey (NHANES).

[6]For example, Lin and Guthrie (2012) find that FAFH tends to contain more saturated fat, sodium, and cholesterol and less dietary fiber per calorie than FAH.

[7]Hoynes and Schanzenbach (2009, Table 3 panel C) estimate a statistically insignificant effect on the propensity to eat a meal away from home of the initial rollout of the Food Stamp Program. Beatty and Tuttle (2015, Table 7) estimate a statistically insignificant effect on FAFH expenditure of the increases in SNAP benefits associated with the American Recovery and Reinvestment Act (ARRA). Burney (2018, Table 3) estimates a statistically insignificant effect on FAFH expenditure and a statistically significant negative effect on the FAFH share of changes in SNAP participation during the recession in the early 2000s. Liu et al. (2013a, 2013b) estimate inconsistently signed (though mostly negative) effects of SNAP participation on FAFH expenditure using a multivariate Tobit system. Liu et al. describe the estimated effects as “negligible” (2013a, p. 162) and “limited” (2013b, p. 210); the estimates are not consistently statistically significant. Yen et al. (2012) estimate statistically significant negative effects of SNAP participation on FAFH expenditure for the elderly using a similar methodology.

[8]Related literatures consider effects on food insecurity (Gregory et al. 2016) and obesity (Gundersen 2016); see also Kreider et al. (2012).

purchased for at-home consumption. We obtain these estimates by applying research designs that explicitly account for the endogeneity of program participation to data from a large and detailed retail panel.

One strand of the prior literature, reviewed for example in Fox et al. (2004) and Andreyeva et al. (2015), compares the diet and food purchases of SNAP participants to those of subgroups of nonparticipants. Many studies in this strand use survey data, but some recent ones use data more similar to those we analyze here. Garasky et al. (2016) and Franckle et al. (2017) study the association between SNAP participation and the healthfulness of purchased foods using scanner data from a grocery retailer. Grummon and Taillie (2017) study the association between SNAP participation and the nutritional content of household food purchases using data from the Nielsen Homescan Consumer Panel.[9]

The potential for selection on unobservables means that comparisons between SNAP recipients and non-recipients may not be interpretable as estimates of the causal effect of SNAP (see, for example, the discussion in Bitler 2016, pp. 151-153). A second strand of literature, reviewed for example in Meyerhofer and Yang (2011), Andreyeva et al. (2015), and Bitler (2016), uses various research designs to account for the endogeneity of SNAP participation and benefits. Yen (2010) uses survey data and an empirical selection model, and finds that SNAP has minimal effects on children's nutrient intake over and above the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). Gregory et al. (2013) and Todd and Ver Ploeg (2014) use survey data and treat state-level SNAP policy variables as excluded instruments for SNAP participation. Gregory et al. (2013) find that SNAP participation increases the consumption of whole fruits but decreases consumption of dark-green vegetables, leading to an overall decrease in diet quality. Todd and Ver Ploeg (2014) find that SNAP participation reduces caloric intake from sugar-sweetened beverages. Bronchetti et al. (2017, 2019) use variation in local food prices to isolate plausibly exogenous variation in the real value of SNAP benefits.[10]

[9]Garasky et al. (2016) find that SNAP households and non-SNAP households purchase similar foods at the grocery store. Franckle et al. (2017) find that grocery items purchased with SNAP benefits tend to be less healthful than grocery items not purchased with SNAP benefits. Grummon and Taillie (2017) find that, along several dimensions, the grocery purchases of households participating in SNAP are less healthful than the grocery purchases of income-eligible non-participating households.

[10]Using food acquisition data, Bronchetti et al. (2017) find that increases in SNAP purchasing power raise HEI-2010 scores among children by a small amount but have no detectable effect among adults. Using survey data, Bronchetti et al. (2019) find that increases in SNAP purchasing power reduce the likelihood of food insecurity.

We study a wide range of outcome measures that include many of those studied in past work, and our estimates are in many cases more precise than in past work. For example, Yen (2010) estimates an effect of SNAP participation on the log of children's dietary fiber intake (as a proportion of the daily recommended intake) of -0.045 with a standard error of 0.178 (right-most column of Table 1). Expressing our estimates as a fraction of the baseline mean, we estimate an effect of SNAP participation on dietary fiber purchases (as a proportion of the DV bound) of -0.043 to -0.004 with standard errors as low as 0.004. Likewise, Gregory et al. (2013) estimate an effect of SNAP participation on the 2005 HEI of -1.4 points with a standard error of 4.5 (left-most column of Table 5). We estimate an effect of SNAP participation on the 2010 HEI of -0.23 to 0.17 points with standard errors as low as 0.14.[11] We believe that more precise estimates are valuable in light of what Bitler (2016, p. 153) calls a "lack of good causal evidence about the effects of SNAP on nutrition."

This paper also contributes to the large literature studying the association between food spending and diet quality.[12] Within this literature, ours is one of only a few papers we are aware of that explicitly addresses the endogeneity of food spending. Using survey data and a random-effects model, Carlson et al. (2014) find that increases in food spending result in only small increases in overall diet quality. Closest to our approach, Griffith et al. (2012) use scanner data from the UK to estimate a demand system, with food spending as an endogenous variable and non-food spending as an excluded instrument. Griffith et al. (2012) find that differences in food spending explain little of the SES gradient in diet quality.

More broadly, this paper contributes to a large literature studying the determinants of diet quality and the effect of policies intended to improve it. Recent studies find relatively small responses to changes in the shopping environment (e.g., Allcott et al. 2019) and the arrival of health information (e.g., Oster 2018; Hut and Oster 2019). Other studies find that financial

[11] Todd and Ver Ploeg (2014) estimate using an instrumental variables approach that SNAP reduces kilocalories from sugar-sweetened beverages, among those consuming them, by 130 log points, or 0.73 of the baseline level, with a standard error of 24 log points (0.21 of baseline) (second column in lower panel of Table 3). We estimate an effect of SNAP participation on the share of kilocalories from soft drinks, sodas, fruit drinks, and ades of -0.005 to 0.036 of the baseline mean, with standard errors as low as 0.013 of the baseline mean.

[12] Blisard et al. (2004) and Frazao et al. (2007) study the association between income and fruit and vegetable spending and conclude that increases in food budgets are unlikely to increase diet quality among low-income households. Mabli et al. (2010) and Anderson and Butcher (2016) find that, among low-SES households, higher food spending is associated with higher diet quality.

incentives can change diet choices (e.g., Bartlett et al. 2014). Verghese et al. (2019) review evidence on the effect of interventions that aim to change the diet quality of SNAP recipients.

The remainder of this paper is structured as follows. Section 2 describes the data and introduces important definitions. Section 3 outlines our empirical framework. Sections 4 and 5 present our results. Section 6 concludes.

2 Data and Definitions

We conduct our primary analysis on the transaction-level data from a large US grocery retailer introduced in Hastings and Shapiro (2018). We augment these data with additional product information, including data on food categories and nutrient content. When describing elements of the data inherited from Hastings and Shapiro (2018), we sometimes quote Hastings and Shapiro (2018) without attribution.

2.1 Purchases and SNAP Use

The retail panel data consist of all purchases in five states made using loyalty cards by customers who shop at one of the retailer's stores at least every other month. We refer to these customers as *households*. Hastings and Shapiro (2018) report that at least 90 percent of purchases at the retailer involve the use of a loyalty card.

We observe 6.02 billion purchases made on 608 million purchase occasions by $486,570$ households from February 2006 through December 2012. We exclude from our analysis the $1,214$ households who spend more than \$5,000 in a single month.

For each item purchased, we observe the quantity, the pre-tax amount paid, and a flag for the use of WIC. For each purchase occasion, we observe the date and a classification of the main payment method used for the purchase, defined as the payment method accounting for the greatest share of expenditure. The main payment method categories include cash, check, credit, debit, and a government benefit category that consists of SNAP, WIC, cash benefits (e.g., TANF) delivered by EBT card, and a number of other, smaller government programs.

We classify a purchase occasion as a *SNAP purchase occasion* if the main payment method is a government benefit and WIC is not used. Figure 9 reports results excluding all households that ever use WIC in a transaction.

We define a *SNAP month* as any household-month with positive total spending across SNAP purchase occasions.[13] Of the household-months in our panel, 7.7 percent are SNAP months. Hastings and Shapiro (2018) report that this fraction is below the SNAP penetration estimated in administrative data, and show evidence consistent with the hypothesis that the retailer's customers tend to have higher incomes than the general population. We define an indicator called *SNAP use* equal to one in any SNAP month and zero otherwise. This indicator serves as our main measure of SNAP participation.

We define a *SNAP adoption* as a period of six or more consecutive non-SNAP months followed by a period of six or more consecutive SNAP months. We refer to the first SNAP month in an adoption as an *adoption month*. We define a *SNAP adopter* as a household with at least one SNAP adoption. Our panel contains a total of 24,456 SNAP adopters. Figure 9 reports results using alternative definitions of SNAP adoption.

For the purposes of quarterly analysis, we define a quarter as a *SNAP quarter* if it contains a SNAP month and as an *adoption quarter* if it contains an adoption month. Figure 9 reports results using an alternative definition of quarterly SNAP use. We include only complete calendar quarters in our analysis.

Hastings and Shapiro (2018) investigate whether SNAP adoption can be taken as a proxy for new enrollment in SNAP. Using administrative records of all debits and credits to SNAP EBT cards of Rhode Island residents from September 2012 through October 2015, Hastings and Shapiro (2018) estimate that the fraction of SNAP adoptions that are actually new SNAP enrollments ranges from 87 to 96 percent depending on the sample used.

Hastings and Shapiro (2018) also investigate what fraction of panelists' total grocery budget is spent at the retailer. They report evidence suggesting that this fraction is large. For example, following SNAP adoption the average retailer panelist spends SNAP benefits at the retailer equivalent to more than 80 percent of average benefits received in a sample of publicly available

[13]For purchase occasions in March 2009 and later, we further observe the exact breakdown of spending according to a more detailed classification that itemizes specific government programs. Using these data, Hastings and Shapiro (2018) report that, excluding WIC transactions, SNAP accounts for 99.3 percent of expenditures classified as a government benefit, that SNAP is used in only 0.23 percent of the purchase occasions that are not classified as SNAP purchase occasions, and that a definition of SNAP month based on the detailed payment data agrees with our principal definition in all but 0.27 percent of household-months.

administrative data. The online Appendix reports results restricting to households with relatively few supermarkets in their county, and hence with presumably fewer opportunities to substitute across retailers.

2.2 Product Classification

The retailer data include characteristics of each product purchased, including the Universal Product Code (UPC), a text description of the product, the product's size, and the product's location within a taxonomy. We refer to locations within the taxonomy as *product categories*. Across all products sold at the retailer there are 6,623 unique product categories.

We use the retailer's product taxonomy along with the SNAP-eligibility classification of products developed and validated in Hastings and Shapiro (2018) to identify food products purchased for at-home consumption.[14] In particular, we restrict attention to purchased SNAP-eligible products and alcoholic beverage products and hereafter refer to them as *food products* for simplicity. Across all food products there are 2,650 unique product categories.

In the retail panel, 83.7 percent of food spending goes to products with a UPC; the remaining 16.3 percent goes to "random-weight" products such as fresh produce or deli meats. We refer to these products as *UPC food products* and *random-weight food products*, respectively.

We classify food products according to the product categories underlying the USDA's Thrifty Food Plan (TFP) (USDA 2007).[15] TFP product categories include whole grains, dark-green vegetables, whole-milk products, and sugars, sweets, and candies. We fail to classify food products that account for 5.8 percent of food spending. The online Appendix presents estimates of the effect of SNAP on the share of purchased kilocalories going to these unclassified products.

We classify each TFP category as healthful or unhealthful based on whether the category is recommended for increased consumption by the 2010 Dietary Guidelines for Americans (DGAs) (HHS and USDA 2010). Appendix A details our procedure for assigning food products to TFP categories and for classifying the healthfulness of TFP categories.

[14] Grocery and prepared food items intended for home consumption are generally SNAP-eligible (FNS 2017). Alcohol, tobacco, pet food, and prepared food intended for on-premise consumption are SNAP-ineligible (FNS 2017).

[15] The TFP specifies the types and quantities of foods households can purchase to obtain a nutritious diet at minimal cost. The TFP is used as the basis for legislated maximum SNAP benefit levels.

2.3 Product Nutritional Information

We obtain nutritional information for food products from several sources. For UPC food products, our primary data source is a nutritional database maintained by Information Resources, Inc (IRI). The IRI nutritional database contains nutritional information for over 260,000 UPCs obtained directly from product labels. For each UPC, the data contain product size, kilocalories, macronutrients (e.g., total fat, total carbohydrates, protein) and micronutrients (e.g., vitamin A, iron, calcium) per serving, and the number of servings per container. We use a single extract of the database as of January 2017.

We supplement the IRI data with similar UPC-level nutritional information from the USDA Branded Food Products Database (USDA 2018), the public websites of Walmart (a retailer) and ShopWell (a personalized nutrition platform), and a file covering store-brand products provided by the retailer.[16,17] We process each data set according to FDA rules governing Nutrition Facts labels (FDA 2013), and we exclude from each data set UPCs with anomalous values.[18]

For the remaining UPC food products, we impute nutritional information based on matched products in the same product category with the same product size unit. Specifically, for each unmatched product for which there are at least 20 matched products, we impute the amount of each nutrient by multiplying the product's size by the median value of the nutrient per product size unit among matched products.

We use product size from the retailer and IRI data to calculate the edible weight of each UPC food product, assuming that there is no inedible portion and that all liquid products share the

[16]For UPC food products for which we have multiple UPC-level data sources, we prioritize data inputs as follows: (1) IRI, (2) ShopWell, (3) USDA Branded Food Products, (4) retailer store brand, (5) Walmart.

[17]The data file covering store-brand products includes nutritional information per serving but not servings per container. For store-brand products for which there are at least 20 matched products in the other UPC-level data files that share the same product category and product size unit, we impute servings per container by multiplying the product's size by the median value of servings per container per product size unit among the matched products.

[18]The FDA requires Nutrition Facts labels to contain the following fields: kilocalories, total fat, saturated fat, trans fat, cholesterol, sodium, total carbohydrates, dietary fiber, sugar, protein, vitamin A, vitamin C, calcium, and iron. However, a simplified Nutrition Facts label containing five "core" nutrients — kilocalories, total fat, sodium, total carbohydrates, and protein — may be used when at least eight of the required fields are present in "insignificant" amounts (FDA 2013). Given these FDA guidelines, we exclude from our UPC-level data sets products for which any of the five core nutrients are missing, and among the remaining products, we set to zero any missing values of the non-core required fields. To limit the role of anomalous values, we further exclude UPCs for which any of the required fields exceed the 99.99th percentile among products in the given data set and UPCs for which the kilocalories implied by the macronutrients exceed reported kilocalories by at least 500.

density of water.[19]

For random-weight food products, our primary data source is release 28 of the USDA National Nutrient Database for Standard Reference (SR28) (USDA 2016a). The SR28 provides information on the nutritional content and weight of nearly 9,000 food items. For each food item, it contains the amount of kilocalories, macronutrients, and micronutrients per 100 edible grams of the item, as well as the typical weight of the edible and inedible portions of the item.

We link to the SR28 the set of retailer products in product categories for which the IRI and retailer store brand UPC-level data cover less than half of category food spending. Appendix A details the linking procedure. We use the resulting links to assign nutritional content and weight information to random-weight food products. We also use the resulting links, together with our estimates of the edible weight of each UPC food product, to supplement the UPC-level and imputed nutritional information for UPC food products. In particular, for UPC food products, we prioritize the UPC-level data sources, then the imputations, and then the links with the SR28.

The UPC-level data sources, imputations, and links with the SR28 provide nutritional information for UPC food products that account for 79.5, 9.0, and 11.0 percent of UPC food spending, respectively. The links with the SR28 provide nutritional information for random-weight food products that account for 96.3 percent of random-weight food spending.

To assess the sensitivity of our estimates to the data assignment scheme, we conduct an experiment in which, for a randomly-chosen subset of UPC food products that together account for 10 percent of UPC food spending, we replace the nutritional information observed in our UPC-level data sources with the imputed or SR28 counterpart, following the priority outlined above.[20] Our main dependent variables at the household-month level have a correlation of at least 0.97 with their counterparts constructed in this experiment. Figure 9 shows estimates of the effect of SNAP using the counterpart of the NDS constructed in this experiment.

Measuring conformance to the DGAs requires information on USDA Food Patterns, which are units (e.g., cup equivalents of vegetables, ounce equivalents of grains) used to help convey desirable daily amounts of foods. Appendix A details our sources and procedures for obtaining the USDA Food Patterns. These sources provide USDA Food Patterns information for UPC food

[19]We exclude from these calculations products with product sizes equivalent to more than 5000 ounces.

[20]The online Appendix reports product-level rank correlations between observed nutrient quantities and their counterparts constructed in this experiment.

products that account for 98.1 percent of UPC food spending, and they provide USDA Food Patterns information for random-weight food products that account for 98.7 percent of random-weight food spending.

Taken together, these data sources provide nutritional information and USDA Food Patterns information for products that account for 97.6 percent of food spending in the retail panel. We exclude from our analysis the products that account for the remaining 2.4 percent of food spending.

2.4 Monthly Food Spending, Food Attributes, and Food Healthfulness

For each household in our panel, we calculate total food spending and kilocalories, macronutrients, micronutrients, and USDA Food Patterns purchased in each calendar month. We also compute several measures of the healthfulness of purchased foods. For the purposes of quarterly analysis, we compute the quarterly average of the monthly variables defined below.

2.4.1 Thrifty Food Plan Kilocalorie Shares

Following Oster (2018), we measure how kilocalorie purchases are distributed across the product categories underlying the TFP. For each household and calendar month with nonzero kilocalories purchased, we calculate the share of total kilocalories purchased going to each of the TFP product categories including an unclassified category. We also calculate the share of total kilocalories purchased going to a composite *fruits and non-starchy vegetables* category comprised of the dark-green vegetables, orange vegetables, other vegetables, and whole fruits categories.

2.4.2 Nutrient Density Indexes and Score

Following Hansen et al. (1979), Fulgoni et al. (2009), Drewnowski and Fulgoni (2014), Handbury et al. (2016), and others, we measure the extent to which a household's food purchases deviate from the nutrient density recommended by the FDA's DV bounds (FDA 2013). We focus on nutrients that are generally required to appear on the Nutrition Facts label and for which the FDA recommends either increased or limited consumption.

Nutrients that the FDA recommends for increased consumption $\mathcal{N}_H = \{$dietary fiber, calcium, iron, vitamin A, vitamin C$\}$ are assigned a DV lower bound indicating the minimum amount that should be consumed per 2,000 kilocalories. Nutrients that the FDA recommends for limited consumption $\mathcal{N}_U = \{$total fat, saturated fat, sodium, cholesterol$\}$ are assigned a DV upper bound indicating the maximum amount that should be consumed per 2,000 kilocalories. We use the DV bounds for adults and children four or more years of age (FDA 2013, Appendix F).

As in Hansen et al. (1979), we calculate, for each household i and calendar month t with nonzero kilocalories purchased and for each nutrient $n \in \left(\mathcal{N}_H \cup \mathcal{N}_U\right)$, a *nutrient density index* δ_{it}^n reflecting the amount of the nutrient purchased per kilocalorie relative to the nutrient density implied by the corresponding DV bound (i.e., the DV bound divided by 2,000). For each $n \in \mathcal{N}_H$, higher values of δ_{it}^n reflect higher healthfulness, with values less than one indicating that household food purchases contain less than the recommended minimum amount of the nutrient per kilocalorie. For each $n \in \mathcal{N}_U$, higher values of δ_{it}^n reflect lower healthfulness, with values greater than one indicating that household food purchases contain more than the recommended maximum amount of the nutrient per kilocalorie.

As in Fulgoni et al. (2009) and Drewnowski and Fulgoni (2014) we summarize the nutrient density indexes using a composite *nutrient density score* (NDS) of the form:

$$\delta_{it} = \frac{\frac{1}{|\mathcal{N}_H|}\sum_{n\in\mathcal{N}_H} \delta_{it}^n}{\frac{1}{|\mathcal{N}_U|}\sum_{n\in\mathcal{N}_U} \delta_{it}^n} \tag{1}$$

where by construction δ_{it} is increasing in δ_{it}^n for $n \in \mathcal{N}_H$ and decreasing in δ_{it}^n for $n \in \mathcal{N}_U$. Fulgoni et al. (2009) and Drewnowski and Fulgoni (2014) compute an NDS of the form in equation (1) under several different specifications of $\mathcal{N}_H$ and $\mathcal{N}_U$.[21] We exclude from our analysis values of the NDS above the 99.9th percentile. Figure 9 shows estimates of the effect of SNAP on the NDS in a sample that includes these values and in a sample that excludes any household that ever has a month in which zero kilocalories are purchased. Figure 9 also shows estimates of the effect of SNAP on an alternative form of the NDS considered in the literature (Fulgoni et al. 2009; Drewnowski and Fulgoni 2014).

[21] For example, Fulgoni et al. (2009) consider a specification in which $\mathcal{N}_H = \{$dietary fiber, calcium, iron, vitamin A, vitamin C, protein$\}$ and $\mathcal{N}_U = \{$saturated fat, sodium, added sugar$\}$.

2.4.3 Healthy Eating Index 2010

The Healthy Eating Index 2010 (HEI-2010) assesses conformance to the 2010 DGAs (Guenther et al. 2014). The HEI-2010 and its predecessor, the HEI-2005, are widely used in studies of the determinants and correlates of diet quality (e.g., Mabli et al. 2010; Gregory et al. 2013; Wang et al. 2014; Condon et al. 2015; Drewnowski et al. 2016).

The HEI-2010 ranges from 0 to 100 and is the sum of 12 component scores, each of which measures conformance to a different aspect of the 2010 DGAs. Nine of the 12 component scores (e.g., whole fruit, total vegetables, whole grains) assess adequacy of diet. The remaining three component scores (empty calories, sodium, and refined grains) assess moderation of diet. All component scores are computed such that higher scores represent higher diet quality. See Guenther et al. (2014) for details regarding the definition of each of the 12 component scores.

For each household and calendar month with nonzero kilocalories purchased, we calculate the HEI-2010. Figure 9 shows estimates of the effect of SNAP on the HEI-2010 in a sample that treats the HEI-2010 as zero in the 0.07 percent of household-months in which zero kilocalories are purchased and in a sample that excludes any household that ever has a month in which zero kilocalories are purchased.

2.4.4 Interpretation of Summary Measures

The NDS and HEI-2010 are related but distinct summaries of diet healthfulness. Whereas the NDS depends only on nutrient density, the HEI-2010 also depends on food categories. The online Appendix presents plots depicting the relationship between cross-sectional averages of and quarter-over-quarter changes in the NDS and HEI-2010. We estimate the correlation between the cross-sectional averages and quarter-over-quarter changes to be 0.67 and 0.37, respectively.

The HEI-2005 has been shown to predict markers of diet-related health. Gao et al. (2008) find that a one standard deviation increase in the HEI-2005 is associated with a 0.4 to 1.9 percent reduction in the average body mass index (BMI) and a 0 to 28.3 percent reduction in the ratio of the probability of being obese to the probability of being normal weight, across different racial and ethnic groups.[22] In our data one standard deviation in the HEI-2010 is equivalent to 0.75 of

[22]Gao et al. (2008) report an estimate of the change in average BMI associated with a one standard deviation change in the HEI-2005 for each racial and ethnic group (first column of Table 4). Expressing these estimates relative

an IQR.

2.5 FoodAPS Data

Portions of our analysis rely on the public-use release of the USDA's National Household Food Acquisition and Purchase Survey (FoodAPS) (USDA 2016b). The FoodAPS data contain detailed information regarding the food acquisitions of a nationally-representative sample of 4,826 households over a seven-day period between April 2012 and January 2013. Food acquisitions include purchases of foods and foods obtained for free.

Survey households record food acquisitions in a "food book" according to whether the acquisition was food at home (FAH) or food away from home (FAFH). Each food book entry corresponds to an "event" such as a trip to the grocery store or a meal at a restaurant. For each FAH event, households are asked to record total spending and scan all acquired items using a handheld scanner. Similarly, for each FAFH event, households are asked to record total spending and write down all items acquired.

For each household, we observe the education of each household member and whether, according to self-reports and administrative SNAP records, the household is currently participating in SNAP.[23] We define a household to be *college-educated* if the household's main food shopper or meal planner reports having a bachelor's degree or higher. We define a household to be *non-college-educated* if this person reports having less than a bachelor's degree.

For each food acquisition, we observe the quantity obtained, expenditure made, and the food group, USDA Food Pattern, and nutritional information associated with each item acquired.[24]

to the corresponding baseline mean BMI (third row of Table 3) implies an estimate of the percent change in average BMI for each racial and ethnic group. Gao et al. (2008) also report an estimate μ of the change in the the ratio of the probability of being obese relative to to the probability of being normal weight associated with a one unit increase in the HEI-2005 for each racial and ethnic group (second column of Table 5). Evaluated at the corresponding standard deviation σ in the HEI-2005 (12th row of Table 2), these estimates imply an estimate $\mu^{\sigma} - 1$ of the percent change in the relative likelihood of being obese associated with a one standard deviation change in the HEI-2005 for each racial and ethnic group.

[23]Estimates of the association between diet-related outcomes and SNAP participation are sensitive to the measurement of SNAP participation (Courtemanche et al. 2019). Following Todd and Scharadin (2016) and Tiehen et al. (2017), we measure current SNAP participation using information from the administrative SNAP records, supplementing with self-reports only for the 122 households that did not consent to being matched with the administrative SNAP records.

[24]Quantity and nutritional information are missing for some items. We replace missing quantities with imputed quantities made available by the USDA Economic Research Service (Mancino et al. 2018) and exclude from our analysis the remaining 4.9 percent of FAH items and 1.1 percent of FAFH items with missing quantity or nutritional

We assign FoodAPS food items to TFP product categories using FoodAPS food group and TFP product category descriptors, following a procedure similar to the one described in appendix A for random-weight retailer products.

For each household, we calculate total food spending, kilocalories, macronutrients, and micronutrients, separately for FAH and FAFH acquisitions. For each household with nonzero acquired FAH kilocalories, we calculate our measures of healthfulness using the nutritional information for FAH acquisitions.

We exclude from our analysis the 107 households who do not have at least one FAH or FAFH food acquisition of an item with valid quantity and nutrition information, the 2 remaining households who report acquiring more than 500,000 kilocalories during the sample week, and the 73 remaining households whose food collection week falls outside the time range covered by the retail panel. This leaves us with a final sample of 4,644 households.[25]

2.6 Comparing Food Healthfulness Between Retailer and FoodAPS Data

We use the FoodAPS data to assess the representativeness of the retail panel. To facilitate comparison, we randomly assign each household in the retail panel a pseudo-survey week within the FoodAPS data collection period such that the distribution of pseudo-survey weeks in the retail panel matches the distribution of actual survey weeks in the FoodAPS data. Then, for each household in the retail panel, we reconstruct our measures of food healthfulness using only transactions within their given pseudo-survey week. We compare the distributions of these pseudo-survey week-based outcomes to their counterparts in the FoodAPS data.

Figure 1 presents cumulative distribution functions of select measures of healthfulness in both data sets. The online Appendix presents analogous plots for all other outcomes. In general, the retailer and FoodAPS data sets paint a fairly similar picture of healthfulness. The 25th, 50th, and 75th percentiles of the distribution of the share of kilocalories from fruits and non-starchy vegetables in the retailer data are 0.01, 0.03, and 0.07, respectively. The corresponding percentiles in the FoodAPS data are 0.01, 0.03, and 0.08. The 25th, 50th, and 75th percentiles of

information.

[25] Among households participating in SNAP in our sample, FAFH accounts for an average of 24 percent of food spending and 19 percent of kilocalorie acquisitions.

the distribution of the share of kilocalories from total fat relative to the DV upper bound in the retailer data are 0.92, 1.16, and 1.39, respectively. The corresponding percentiles in the FoodAPS data are 0.89, 1.17, and 1.45.

There are, however, outcomes for which these comparisons reveal meaningful differences between the two data sets. For example, the online Appendix figures show that, relative to FoodAPS households, retailer households devote a notably larger fraction of purchased kilocalories to non-whole grains and a notably smaller fraction of purchased kilocalories to frozen or refrigerated entrees.

Turning to summary measures of healthfulness, the 25th, 50th, and 75th percentiles of the distribution of the NDS are 0.48, 0.70, and 1.04 in the retailer data and 0.54, 0.82, and 1.30 in the FoodAPS data. The 25th, 50th, and 75th percentiles of the distribution of the HEI-2010 are 43.5, 53.2, and 62.7 in the retailer data and 41.6, 52.1, and 63.0 in the FoodAPS data.

The online Appendix presents results of Kolmogorov-Smirnov tests of the equality of the distributions across data sets for all outcomes. In most cases, we can confidently reject the null hypothesis that the two distributions are the same.

2.7 Administrative Data on Earnings and SNAP Participation

Following Hastings and Shapiro (2018), we use Rhode Island state administrative records housed in a secure facility. These records are not linked to our retail panel.

From these records we construct a panel of households containing, in each quarter from the second quarter of 2006 through the fourth quarter of 2012, an indicator for participation in SNAP and the average monthly sum of total unemployment insurance benefits received by and total earnings reported for all individuals who are in the household as of the quarter's end. We refer to this total as *in-state earnings* for short, and we note that it excludes income sources such as social security benefits and out-of-state earnings.

We restrict attention to households who participate in SNAP at least once during our sample period. We define a SNAP spell to be a contiguous period of SNAP participation. We define a SNAP adoption quarter to be the first quarter of a SNAP spell.

The resulting panel consists of 143,929 households observed in 3,886,083 household-quarters.

Appendix A contains additional details on the construction of this panel.

3 Model and Assumptions

We describe the at-home food consumption of household i in period t by a vector $\mathbf{d}_{it}$ of attributes. We summarize the healthfulness of at-home food consumption by a scalar $h_{it} = H(\mathbf{d}_{it})$ where $H(\cdot)$ is a known function that is homogeneous of degree zero. For example, the attributes $\mathbf{d}_{it}$ might be the total number of kilocalories in each of the TFP categories, and the summary h_{it} might be the share of kilocalories from fruits and non-starchy vegetables. Like the TFP kilocalorie share, all of the summaries of healthfulness that we consider are homogeneous of degree zero in the corresponding attributes, consistent with our assumption on $H(\cdot)$.

Letting Δ denote the first-difference operator, we model the evolution of healthfulness over time as

$$\Delta h_{it} = \beta \Delta s_{it} + \mathbf{q}_{it}^{'} \rho + \gamma \eta_{it} + \varepsilon_{it} \tag{2}$$

where s_{it} is an indicator for participation in SNAP, $\mathbf{q}_{it}$ is a vector of controls such as indicators for time period, η_{it} is an income shock, ε_{it} is a preference shock satisfying $\mathrm{E}(\varepsilon_{it}|\mathbf{q}_{it}) = 0$, and ρ and γ are parameters. The parameter β captures the causal effect of SNAP and is our target.

The econometrician has data $\{\lambda_{it}\mathbf{d}_{it}, s_{it}, \mathbf{q}_{it}, z_{it}\}_{i=1,\ldots,N}^{t=1,\ldots,T}$ where z_{it} is an indicator for SNAP adoption and $\lambda_{it} > 0$ is an unknown scalar reflecting the ratio of food purchased at the retailer to food consumed. The assumption that the econometrician observes a scalar proportion λ_{it} of each attribute in $\mathbf{d}_{it}$ rules out, for example, that a household sources its healthy food from one retailer and its unhealthy food from another, or that a household substitutes toward buying a relatively greater share of its unhealthy food at a given retailer when it enters SNAP. Formally, because $h_{it} = H(\mathbf{d}_{it}) = H(\lambda_{it}\mathbf{d}_{it})$, observing $\lambda_{it}\mathbf{d}_{it}$ amounts to observing h_{it}. Thus, we may think of h_{it} as a direct measure of the healthfulness of food purchased at a given retailer, or, under the stated assumptions, as an indirect measure of the healthfulness of all foods eaten at home.

The shocks η_{it} and ε_{it} are unobserved. A fundamental concern is that income shocks affect both healthfulness and SNAP participation, i.e., that $\gamma \neq 0$ and $\mathrm{E}(\Delta s_{it}\eta_{it}) \neq 0$. We adopt two research designs for identification and estimation of the parameter of interest β.

3.1 Research Design Based on Fine Timing of Program Adoption

In this research design we use an observed proxy to learn the evolution of the income shock η_{it} around changes in SNAP participation, as in Freyaldenhoven et al. (2019). Let x_{it} be an observable measure of income that obeys

$$\Delta x_{it} = \mathbf{q}_{it}^{'} \psi + \varphi \eta_{it} + \zeta_{it} \tag{3}$$

where ζ_{it} is an unobserved measurement error satisfying $\mathrm{E}\left(\zeta_{it} | \mathbf{q}_{it}\right) = 0$, ψ and φ are parameters, and $\varphi \neq 0$. Equation (3) allows that Δx_{it} is an imperfect (noisy) proxy for the underlying income shock η_{it} that influences healthfulness in equation (2).

Let z_{it} be an indicator for whether household i adopts SNAP in period t. We assume the exclusion restriction that

$$\mathrm{E}\left(\left(z_{it}, z_{it+1}\right)' \left(\varepsilon_{it}, \zeta_{it}\right) \right) \quad = \quad 0, \tag{4}$$

i.e., that the SNAP adoption indicator and its first lead are orthogonal to the preference shock ε_{it} and the measurement error ζ_{it}.

Under suitable relevance conditions, equation (4) justifies a two-stage least squares (2SLS) regression of Δh_{it} on Δs_{it}, Δx_{it}, and $\mathbf{q}_{it}$, using $\left(z_{it}, z_{it+1}\right)$ as excluded instruments (Freyaldenhoven et al. 2019). Importantly, equation (4) allows the timing of SNAP adoption to be related to the timing of income shocks. We expect the appropriate relevance conditions to be satisfied because entry into SNAP in the near future tends to be associated with lower income in the present (Hastings and Shapiro 2018; Freyaldenhoven et al. 2019).

Intuitively, equation (4), coupled with our other assumptions, implies that the average dynamics of the proxy Δx_{it} in the periods before and during SNAP adoption mirror those of the confound η_{it} (Freyaldenhoven et al. 2019). In Section 4 we discuss the sensitivity of our conclusions to relaxation of this restriction.

In our application, x_{it} is the household's in-state earnings. Because this is observed in the administrative data but not in the retail data, we estimate the model via two-sample two-stage least squares (TS2SLS) (Inoue and Solon 2010). Our first-stage data consist of $\left\{x_{it}, s_{it}, \mathbf{q}_{it}, z_{it}\right\}_{i=N+1,\ldots,.N+M}^{t=1,\ldots,T}$

where M is the number of households in the administrative data panel.

Our main approach to inference for the TS2SLS estimator is an asymptotic approximation described in Appendix B. Figure 9 reports estimates with standard errors calculated via a non-parametric bootstrap.

3.2 Research Design Based on Exogenous Timing of Program Exit

In this research design we exploit the fact that a large fraction of SNAP spells' lengths are divisible by six months to isolate variation in Δs_{it} that is not related to the unobserved shocks $(\eta_{it}, \varepsilon_{it})$, as in Hastings and Shapiro (2018). For each household i and calendar month t, define a clock c_{it} that begins in the sixth month following SNAP adoption and resets every six months until two years following SNAP adoption. Formally

$$c_{it} = \mod\left(t - \max\left\{t' \leq t : z_{it'} = 1\right\}, 6\right) + 1 \tag{5}$$

when $(t - \max\{t' \leq t : z_{it'} = 1\}) \in \{6, ..., 23\}$, and $c_{it} = 0$ otherwise.

Let $m_{it} = \mathbf{1}_{c_{it=1}}$ be an indicator for whether household i is in the first month of the clock in period t. We assume the exclusion restriction that

$$\mathrm{E}\left(m_{it}\left(\eta_{it}, \varepsilon_{it}\right)\right) = 0, \tag{6}$$

i.e., that the timing of income and preference shocks is unrelated to the timing of the clock.

Under a suitable relevance condition, equation (6) justifies a 2SLS regression of Δh_{it} on Δs_{it} and $\mathbf{q}_{it}$, using m_{it} as an excluded instrument. We expect the appropriate relevance condition to be satisfied because aspects of the program's structure mean that many households exit SNAP after 6, 12, 18, etc. months on the program (Klerman and Danielson 2011; Mills et al. 2014; Scherpf and Cerf 2019; Gray 2019; Hastings and Shapiro 2018). Intuitively, equation (6), coupled with our other assumptions, implies that there should be no six-month cycle in food healthfulness after SNAP adoption, absent a causal effect of SNAP on healthfulness.

4 Estimated Effect of SNAP on the Composition of Purchased Foods

4.1 Research Design Based on Fine Timing of Program Adoption

Figure 2 illustrates the first stage of this research design. The figure plots estimates that summarize the evolution of in-state earnings before and after entry into SNAP in the administrative data panel.

Figure 2 shows that, as expected, in-state earnings fall during the quarter in which a household enters SNAP. If healthfulness h_{it} is a normal good ($\gamma > 0$), then a naive estimate of equation (2) that ignores the income confound η_{it} will tend to understate the true effect β of SNAP on healthfulness.

Figure 2 also shows that in-state earnings fall in the quarters preceding a household's entry into SNAP. Under the assumptions of this research design, the dynamics of the income confound around SNAP adoption mirror those of in-state earnings. We can therefore learn how the income confound affects healthfulness h_{it} by looking at how healthfulness evolves before SNAP adoption. This, in turn, allows us to adjust our estimate of the causal effect β of SNAP on healthfulness to account for the role of the income confound.

Figure 3 illustrates this logic. Panel A of Figure 3 plots estimates from a dynamic analogue of equation (2) estimated on monthly data for three select outcome variables h_{it}: the share of kilocalories devoted to fruits and non-starchy vegetables, the share of kilocalories from total fat relative to the DV upper bound, and the nutrient density score (NDS). Each plot summarizes the evolution of an outcome variable before and after entry into SNAP in the retail panel. The online Appendix presents analogues of Figure 3 for all outcome variables.

Both the share of kilocalories from fruits and non-starchy vegetables and the NDS exhibit a trend prior to entry into SNAP in the direction of declining healthfulness. Under the model in equation (2), this reflects the causal effect of declining income on healthfulness. The share of kilocalories from total fat relative to the DV upper bound exhibits a less consistent trend.

All three outcomes exhibit a statistically significant and visually clear change upon entry into SNAP, in the direction of declining healthfulness. Under the model in equation (2), this reflects a combination of the causal effect of SNAP and the causal effect of the decline in income that accompanies entry into the program.

Panel B of Figure 3 repeats the specification from panel A of Figure 3 at quarterly resolution. The plots also overlay the estimated trend in in-state earnings from Figure 2, rescaled so that, for each outcome, the change in the outcome matches the change in in-state earnings between the two quarters prior to entry into SNAP. In this research design, the divergence between the outcome series and the (rescaled) in-state earnings series upon entry into SNAP reveals the causal effect of SNAP on healthfulness.

Panel C of Figure 3 plots estimates from a dynamic analogue of equation (2), using in-state earnings as a proxy for the income confound and using the first lead of SNAP adoption as an excluded instrument for in-state earnings following a dynamic analogue of the exclusion restriction in equation (4). In this research design, these plots reveal the causal effect of SNAP on each outcome.

In the cases of the share of kilocalories from fruits and non-starchy vegetables and the NDS, adjusting for the income confound reduces the estimated decline in healthfulness relative to the unadjusted estimates in panels A and B of Figure 3. In the case of the share of kilocalories from total fat relative to the DV upper bound, adjusting for the income confound increases the estimated decline in healthfulness relative to the unadjusted estimates in panels A and B of Figure 3. In all three cases, adjusting for the income confound increases the standard error on the estimated effect of SNAP.

Panel D of Figure 3 repeats the plots of panel C of Figure 3 with the y-axis range scaled to match the cross-sectional IQR of the average of the outcome variable across all households in the retail panel. In each case, the effect of SNAP appears small compared to the cross-sectional variation in the outcome.

As Figure 3 illustrates, the essence of this research design is to infer the dynamics of the income confound from those of in-state earnings. This approach fails if, for example, in-state earnings do not correctly capture the dynamics of income, or if the important confound is not income but some other factor (e.g., family size) that evolves differently around SNAP adoption. In the online Appendix, we show the sensitivity of our estimates of the causal effect of SNAP to different assumptions about the dynamics of the confound. The estimated effect of SNAP remains quantitatively small relative to the IQR even if we assume dynamics far more extreme than those exhibited by in-state earnings, though our confidence intervals are wider in such cases.

Figure 4 presents estimates of the causal effect β of SNAP in equation (2) on the full set of TFP kilocalorie shares, the full set of nutrient density indexes, the NDS, and the HEI-2010. For each outcome variable we report the estimated effect β, its confidence interval, and the cross-sectional IQR of the average of the outcome variable across all households in the retail panel, signed so that a positive IQR indicates that higher values of this outcome are associated with greater healthfulness. Thus, if the estimated effect of SNAP is on the same side of zero as the IQR, SNAP is estimated to improve healthfulness along the given dimension. If the estimated effect is large in absolute value relative to the IQR, SNAP is estimated to have a large effect relative to the cross-sectional variation in the given outcome. Section 4.3 compares our estimates to several other benchmarks. The online Appendix shows the fit of the static model in equation (2) to the dynamics of healthfulness depicted in Figure 3.

Figure 4 shows that SNAP is estimated to improve healthfulness along some dimensions and worsen it along others. In most cases, the estimated effect of SNAP is small relative to the cross-sectional variation in the outcome variable. For example, the confidence interval for the effect of SNAP on the share of kilocalories going to fruits and non-starchy vegetables ranges from -0.0017 to -0.0001, or from -0.056 to -0.003 of an IQR.[26] The confidence interval for the effect of SNAP on the share of kilocalories from total fat relative to the DV upper bound ranges from 0.0074 to 0.0203, or from 0.039 to 0.106 of an IQR.

Turning to our summary measures, the confidence interval for the effect of SNAP on the NDS ranges from -0.017 to -0.001, or from -0.057 to -0.005 of an IQR. The confidence interval for the effect of SNAP on the HEI-2010 ranges from -0.101 to 0.448, or from -0.010 to 0.044 of an IQR. (See also Table 1.)

Figure 9 reports estimates of the average marginal effect of SNAP from models in which the dependent variable is either the natural logarithm of the NDS or the natural logarithm of the HEI-2010. The online Appendix presents estimates of the effect of SNAP on an additional set of macronutrient kilocalorie shares. The online Appendix also presents an estimate of the size

[26]Figure 4 does not present estimates for the composite fruits and non-starchy vegetables category. The figure instead presents estimates for the dark-green vegetables, orange vegetables, other vegetables, and whole fruits TFP product categories, which together make up the composite category.

distortion cutoff, a measure of instrument strength proposed by Andrews (2018), for the models involving our summary measures.

To explore heterogeneity in the effect of SNAP, the online Appendix reports estimates of the effect of SNAP on our summary measures for different subgroups of households. To explore whether SNAP affects aspects of the distribution of healthfulness other than the mean, the online Appendix reports estimates of the effect of SNAP on the probability that a household achieves different levels of our summary measures.

Interpretation of the Coefficient on Earnings

Table 1 reports estimates of the coefficient $100\gamma/\varphi$ on in-state earnings (in hundreds of dollars) from this research design for our two summary measures. For the NDS, we estimate a coefficient of 0.0036, with a confidence interval of -0.0012 to 0.0085, or -0.0041 to 0.0288 of an IQR. For the HEI-2010, we estimate a coefficient of 0.2755, with a confidence interval of 0.1020 to 0.4490, or 0.0101 to 0.0442 of an IQR. Given knowledge of the effect φ of the income confound on in-state earnings, these estimates can be translated into estimates of the effect γ of the income confound on healthfulness.

Table 1 also reports estimates of the ratio $100\left(\gamma/\varphi\right)/\beta$ of the coefficient on in-state earnings to the coefficient on SNAP use. Given knowledge of φ, these estimates can be translated into estimates of the ratio $100\gamma/\beta$ of the effect of income on healthfulness to the effect of SNAP use on healthfulness. Under the restriction that both SNAP use and income affect healthfulness only through their effect on food spending, the ratio $100\gamma/\beta$ should equal the ratio of the effect of a \$100 increase in income on food spending to the effect of SNAP participation on food spending. Taking $\varphi = 1$ as a benchmark, the confidence intervals on $100\left(\gamma/\varphi\right)/\beta = 100\gamma/\beta$ admit a wide range of values for this ratio, including values consistent with the estimates in Hastings and Shapiro (2018).[27]

The restriction that SNAP use and income affect healthfulness only through their effect on food spending can be tested directly, by testing the overidentifying restriction that the instruments (z_{it}, z_{it+1}) are jointly exogenous in a model in which food spending is the only endogenous re-

[27] Hastings and Shapiro (2018) argue that \$100 in additional income increases food spending by about \$10, whereas SNAP participation increases food spending by about \$110, implying that $100\gamma/\beta \approx 0.09$ if both SNAP use and income affect healthfulness only through their effect on food spending.

gressor. Table 1 reports the results of this test. The restriction is not rejected in the case of the NDS but is rejected in the case of the HEI-2010.

4.2 Research Design Based on Exogenous Timing of Program Exit

Panel A of Figure 5 illustrates the first stage of this research design. The plot shows coefficients from a regression of the change in SNAP participation Δs_{it} on indicators for months of the six-month SNAP clock c_{it}. As expected, households are especially likely to transition from a SNAP month to a non-SNAP month in the first month of the clock, i.e., after completing a six-month block of their SNAP spell.

Panel B of Figure 5 illustrates the second stage of this research design. The plots show estimated coefficients from a regression of the change in a measure of healthfulness Δh_{it} on indicators for months of the SNAP clock c_{it}. We scale these estimates by the absolute value of the estimated change in the probability of SNAP participation in the first month of the clock from panel A of Figure 5. In this research design, the difference between the coefficient on the indicator for the first month of the clock and the coefficient on the indicators for the other months reflects the causal effect of a change in SNAP participation on healthfulness h_{it}. The healthfulness measures h_{it} are the same as those in Figure 3, and the online Appendix presents an analogue of panel B of Figure 5 for all outcome variables. The plots show that there is little evidence of a systematic effect of SNAP on the outcomes depicted.

Figure 6 presents estimates of the causal effect of SNAP β in equation (2) on the full set of outcomes, following the format of Figure 4. The estimates are generally less precise than those from the program adoption research design. The confidence interval for the effect of SNAP on the NDS ranges from -0.065 to -0.014, or from -0.220 to -0.049 of an IQR. The confidence interval for the effect of SNAP on the HEI-2010 ranges from -1.176 to 0.712, or from -0.116 to 0.070 of an IQR.

Figure 9 reports estimates of the effect of SNAP among the sample of SNAP adopters who exhibit a period of at least six consecutive non-SNAP months following their initial SNAP adoption. Figure 9 also reports estimates of the effect of SNAP from an alternative specification in which all variables are averaged to the level of the household-quarter.

4.3 Comparison of Magnitudes

Figure 7 compares the estimated effect of SNAP on our two summary measures of healthfulness to several benchmarks drawn both from our own calculations and from the literature. The first set of comparisons are to measures of the cross-sectional variation in the outcome. The first comparison is to the cross-sectional IQR. This repeats the comparison from Figures 4 and 6. The second comparison is to the cross-sectional standard deviation. The third and fourth comparisons are to the gradients with respect to two markers of socioeconomic status, education and income, with the latter measured relative to a poverty line. The fifth comparison is to the gradient with respect to diet cost, measured as total expenditure relative to total kilocalories for all food acquired.

The second set of comparisons are to measures of the time-series changes in the outcome. The sixth and seventh comparisons are to measures of the within-household IQR, computed as the average, across households, of the household-level IQR of the outcome over calendar months and calendar quarters, respectively. The eighth and final comparison is to the trend in the outcome over the first decade of the 2000s, which is available from the literature only for the HEI-2010.

In all cases, the effects we estimate are small relative to the given benchmark. In the case of our program adoption design, our confidence intervals exclude positive effects of SNAP larger than 15 percent of any benchmark. In the case of our program exit design, our confidence intervals exclude positive effects of SNAP larger than 24 percent of any benchmark.

4.4 Supplemental Analysis of the Nielsen Homescan Consumer Panel

In the online Appendix, we study the relationship between food healthfulness and SNAP participation in data from the Nielsen Homescan Consumer Panel (NHCP). NHCP panelist households are asked to record all consumer packaged good purchases, regardless of the store where they were purchased. In a quarterly supplement to the NHCP, a subset of panelist households are also asked whether they are currently using SNAP.

For each panelist household in each calendar quarter, we use data on household food purchases to construct many of the measures of food healthfulness introduced in Section 2.4. We define a SNAP quarter to be any household-quarter in which the household reports currently using SNAP.

We estimate a two-way fixed effects regression of each measure of healthfulness on an indicator for whether the current quarter is a SNAP quarter, including both a household fixed effect and a calendar quarter fixed effect. We take the estimated coefficient on the SNAP quarter indicator as our summary of the longitudinal association between SNAP participation and healthfulness.

In most cases, we find that the estimated association is small in magnitude relative to cross-sectional variation in measured food healthfulness. For example, in the case of the NDS, the confidence interval on the estimated association ranges from -0.0309 to 0.0079, or from -0.0982 to 0.0251 of an IQR.

5 Implications for the Socioeconomic Gradient in Food Healthfulness

Consider a model of the form

$$\Delta h_{it} = \tilde{\beta}\Delta f_{it} + \mathbf{q}_{it}^{'}\tilde{\rho} + \tilde{\gamma}\tilde{\eta}_{it} + \tilde{\varepsilon}_{it} \tag{7}$$

where f_{it} is food-at-home (FAH) spending and the remaining objects are defined by analogy to equation (2). We assume that the data consist of $\left\{\lambda_{it}\mathbf{d}_{it}, s_{it}, \mathbf{q}_{it}, z_{it}, \tilde{\lambda}_{it}f_{it}\right\}_{i=1,\ldots,N}^{t=1,\ldots,T}$ where $\tilde{\lambda}_{it} \in [0,1]$ is the share of total FAH spending devoted to the retailer.

We can estimate equation (7) by adopting exclusion restrictions analogous to those in Section 3. For the first research design, the excluded instruments are an indicator for SNAP adoption and its first lead. Hastings and Shapiro (2018) show that food spending at the retailer increases significantly upon SNAP adoption, and Figure 2 shows that income declines in the period prior to adoption. For the second research design, the excluded instrument is an indicator for the first month of the SNAP clock. Hastings and Shapiro (2018) show that food spending at the retailer decreases significantly in the first month of the SNAP clock. We can therefore expect our instruments to be relevant in both research designs.

Table 2 reports estimates of $\tilde{\beta}$ for each research design for both the NDS and the HEI-2010 under various assumptions about $\tilde{\lambda}_{it}$. The online Appendix compares the estimates of $\tilde{\beta}$ from our program adoption research design under a benchmark assumption about $\tilde{\lambda}_{it}$ to an estimate of the cross-sectional association between the given measure of healthfulness and FAH spending in the FoodAPS data. We are able to reject the equality of $\tilde{\beta}$ and its cross-sectional analogue for both the NDS and HEI-2010.

We use these estimates of $\tilde{\beta}$ to simulate the effect on the socioeconomic gap in food healthfulness of eliminating the socioeconomic gap in FAH spending. We conduct these simulations on FoodAPS data, treating educational attainment as the socioeconomic variable of interest. We conduct two simulations.

First, we simulate equating the mean level of FAH spending between college-educated and non-college-educated households. To implement this counterfactual, we multiply the mean difference in FAH spending between college-educated and non-college-educated households in FoodAPS by the estimated value of $\tilde{\beta}$, and divide the resulting product by the mean difference in healthfulness between college-educated and non-college-educated households in FoodAPS. The resulting value estimates the share of the gap in mean healthfulness that would be closed if the gap in mean FAH spending were closed.

Second, we simulate equating the entire distribution of FAH spending between college-educated and non-college-educated households. To implement this counterfactual, we assign to each non-college-educated household a college-educated counterpart whose percentile in the distribution of FAH spending among college-educated households is closest to that of the non-college-educated household among non-college-educated households, breaking ties at random. We then multiply the estimated value of $\tilde{\beta}$ by the difference in FAH spending between the two households to predict how much the non-college-educated household's healthfulness would change if the household's FAH spending were equal to that of the household's college-educated counterpart.

Figure 8 reports the results of the two simulations for the NDS. The online Appendix reports the results of the two simulations for the HEI-2010.

The first simulation shows that closing the gap in mean FAH spending between college-educated and non-college-educated households would widen the mean difference in the NDS by 4.7 percent of its baseline value, with a standard error of 4.5 percent. Closing the gap in FAH spending would narrow the mean difference in the HEI-2010 by 3.6 percent of its baseline value with a standard error of 3.0 percent. These estimates are statistically indistinguishable from zero, and we are able to reject that closing the gap in mean FAH spending would narrow the healthfulness gap by more than 10 percent of its baseline value. The second simulation likewise shows that equalizing the distribution of FAH spending between the two groups would do little to reduce the large difference in the distribution of healthfulness.

The online Appendix presents an analogue of the distribution plot in Figure 8 based on a model in which which we allow $\tilde{\beta}$ in equation (7) to vary with baseline FAH spending. This exercise thus allows for heterogeneity in the effect of food spending on food healthfulness.

6 Conclusion

We use data from a large retail panel, and two research designs, to study the effect of SNAP participation on the composition of foods purchased for at-home consumption. The effect of SNAP is inconsistent in sign and, for most of the outcomes we consider, is small in magnitude relative to cross-sectional variation. Counterfactual simulations imply that closing the gap in food spending between college-educated and non-college-educated households would not close the gap in two summary measures of food healthfulness.

We study the effect of SNAP on the composition of foods purchased for at-home consumption. If SNAP has small effects on the composition of foods eaten, then SNAP's effects on diet healthfulness may turn on how SNAP influences the amount of food eaten. By making more food available to the household, SNAP can reduce food insecurity (Gregory et al. 2016). Greater food intake can also contribute to the calorie imbalance responsible for high prevalence of obesity and related diet-related health risk factors in the United States (HHS and USDA 2010). Although a body of existing research (reviewed, for example, in Gundersen 2016) studies the effect of SNAP participation on obesity, data and other limitations mean that this topic remains, in our view, an important one for future work.

Some existing research evaluates proposals to modify the design of SNAP to increase the healthfulness of purchased foods. Examples of such proposals include adding incentives to purchase healthier foods (Bartlett et al. 2014) and making some less healthy foods ineligible (Basu et al. 2014; but see also Schanzenbach 2013). We believe that further evidence on the consequences of such program changes would be valuable.

References

Aggarwal, Anju, Pablo Monsivais, and Adam Drewnowski. 2012. Nutrient intakes linked to better health outcomes are associated with higher diet costs in the US. *PLoS ONE* 7(5): e37533.

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell. 2019. Food deserts and the causes of nutritional inequality. *Quarterly Journal of Economics* 134(4): 1793-1844.

Anderson, Patricia M. and Kristin F. Butcher. 2016. The relationships among SNAP benefits, grocery spending, diet quality, and the adequacy of low-income families' resources. Center on Budget and Policy Priorities: Policy Futures.

Andreyeva, Tatiana, Amanda S. Tripp, and Marlene B. Schwartz. 2015. Dietary quality of Americans by Supplemental Nutrition Assistance Program participation status: A systemic review. *American Journal of Preventative Medicine* 49(4): 594-604.

Andrews, Isaiah. 2018. Valid two-step identification-robust confidence sets for GMM. *Review of Economics and Statistics* 100(2): 337-348.

Bartlett, Susan, Jacob Klerman, Lauren Olsho, Christopher Logan, Michelle Blocklin, Marianne Beauregard, Ayesha Enver, Parke Wilde, Cheryl Owens, and Margaret Melhem. 2014. Evaluation of the Healthy Incentives Pilot (HIP): Final report. United States Department of Agriculture Food and Nutrition Service.

Basu, Sanjay, Hilary Kessler Seligman, Christopher Gardner, and Jay Bhattacharya. 2014. Ending SNAP subsidies for sugar-sweetened beverages could reduce obesity and type 2 diabetes. *Health Affairs* 33(6): 1032-1039.

Beatty, Timothy K.M. and Charlotte J. Tuttle. 2015. Expenditure response to increases in in-kind transfers: Evidence from the Supplemental Nutrition Assistance Program. *American Journal of Agricultural Economics* 97(2): 390-404.

Bitler, Marianne P. 2016. The health and nutrition effects of SNAP: Selection into the program and a review of the literature on its effects. In J. Bartfeld, C. Gundersen, T. Smeeding, and J. P. Ziliak (eds.), *SNAP Matters: How Food Stamps Affect Health and Well Being*: 134-160. Stanford: Stanford University Press.

Blisard, Noel, Hayden Stewart, and Dean Jolliffe. 2004. Low-income households' expenditures on fruits and vegetables. United States Department of Agriculture Economic Research Service Agricultural Economic Report No. 833.

Bowman, Shanthy A., John C. Clemens, James E. Friday, Rachel C. Thoerig, and Alanna J. Moshfegh. 2014. Food Patterns Equivalents Database 2011-12: Methodology and User Guide. Food Surveys Research Group, Agricultural Research Service, United States Department of Agriculture. Accessed at <http://www.ars.usda.gov/nea/bhnrc/fsrg> on March 27, 2017.

Bronchetti, Erin T., Garret Christensen, Benjamin Hansen. 2017. Local food prices, SNAP, the National School Lunch and School Breakfast Programs, and nutritional outcomes. Working paper.

Bronchetti, Erin T., Garret Christensen, and Hilary W. Hoynes. 2019. Local food prices, SNAP purchasing power, and child health. *Journal of Health Economics* 68: 1-17.

Burney, Shaheer. 2018. In-kind benefits and household behavior: The impact of SNAP on food-away-from-home consumption. *Food Policy* 75: 134-146.

Carlson, Andrea, Diansheng Dong, and Mark Lino. 2014. Association between total diet cost and diet quality is limited. *Journal of Agricultural And Resource Economics* 39(1): 47-68.

Clay, Marie, Michele Ver Ploeg, Alisha Coleman-Jensen, Howard Elitzak, Christian Gregory, David Levin, Constance Newman, and Matthew P. Rabbitt. 2016. Comparing National Household Food Acquisition and Purchase Survey (FoodAPS) data with other national food surveys' data. United States Department of Agriculture Economic Research Service Economic Information Bulletin No. 157.

Condon, Elizabeth, Susan Drilea, Keri Jowers, Carolyn Lichtenstein, James Mabli, Emily Madden, and Katherine Niland. 2015. Diet quality of Americans by SNAP participation status: Data from the National Health and Nutrition Examination Survey, 2007-2010. Walter R. McDonald & Associates, Inc. and Mathematica Policy Research for the Food and Nutrition Service.

Courtemanche, Charles, Augustine Denteh, and Rusty Tchernis. 2019. Estimating the associations between SNAP and food insecurity, obesity, and food purchases with imperfect administrative measures of participation. *Southern Economic Journal* 86(1): 202-228.

Drewnowski, Adam and Victor L. Fulgoni, III. 2014. Nutrient density: Principles and evaluation tools. *The American Journal of Clinical Nutrition* 99(5): 1223S-1228S.

Drewnowski, Adam, Anju Aggarwal, Andrea Cook, Orion Stewart, and Anne Vernez Moudon. 2016. Geographic disparities in Healthy Eating Index scores (HEI-2005 and 2010) by residential property values: Findings from Seattle Obesity Study (SOS). *Preventive Medicine* 83: 46-55.

Falk, Gene, Karen E. Lynch, and Jessica Tollestrup. 2018. Federal spending on benefits and services for people with low income: In brief. Congressional Research Service Report No. 45097.

Food and Drug Administration (FDA). 2013. A food labeling guide: Guidance for industry. Food and Drug Administration Center for Food Safety and Applied Nutrition.

Food and Nutrition Act of 2008, 7 U.S.C. § 2011 (2017).

Food and Nutrition Service (FNS). 2012. Building a healthy America: A profile of the Supplemental Nutrition Assistance Program. Accessed at <http://www.fns.usda.gov/building-healthy-america-profile-supplemental-nutrition-assistance-program> on November 5, 2019.

———. 2016. Supplemental Nutrition Assistance Program. Accessed at <http://www.fns.usda.gov/pd/supplemental-nutrition-assistance-program-snap> on May 25, 2016.

———. 2017. Supplemental Nutrition Assistance Program (SNAP): Eligible food items. Accessed at <http://www.fns.usda.gov/snap/eligible-food-items> on January 14, 2017.

Fox, Mary Kay, William Hamilton, and Biing-Hwan Lin. 2004. Effects of food assistance and nutrition programs on nutrition and health: Volume 4, Executive summary of the literature. United States Department of Agriculture Economic Research Service Food Assistance and Nutrition Research Report No. 19-4.

Franckle, Rebecca L., Alyssa Moran, Tao Hou, Dan Blue, Julie Greene, Anne N. Thorndike, Michele Polacsek, and Eric B. Rimm. 2017. Transactions at a Northeastern supermarket chain: Differences by Supplemental Nutrition Assisstance Program use. *American Journal of Preventative Medicine* 53(4): e131-e138.

Frazao, Elizabeth, Margaret Andrews, David Smallwood, and Mark Prell. 2007. Food spending patterns of low-income households: Will increasing purchasing power result in healthier food choices? United States Department of Agriculture Economic Research Service

Economic Information Bulletin No. 29-4.

Freyaldenhoven, Simon, Christian Hansen, and Jesse M. Shapiro. 2019. Pre-event trends in the panel event-study design. *American Economic Review* 109(9): 3307-3338.

Fulgoni, III, Victor L., Debra R. Keast, and Adam Drewnowski. 2009. Development and validation of the nutrient-rich foods index: A tool to measure nutritional quality of foods. *The Journal of Nutrition* 139(8): 1549-1554.

Gao, Sue K., Shirley AA. Beresford, Laura L. Frank, Pamela J. Schreiner, Gregory L. Burke, and Annette L. Fitzpatrick. 2008. Modifications to the Healthy Eating Index and its ability to predict obesity: The Multi-Ethnic Study of Atherosclerosis. *American Journal of Clinical Nutrition* 8: 64-69.

Garasky, Steven, Kassim Mbwana, Andres Romualdo, Alex Tenaglio, and Manan Roy. 2016. Foods typically purchased by Supplemental Nutrition Assistance Program (SNAP) households. United States Department of Agriculture, Food and Nutrition Service.

Gray, Colin. 2019. Leaving benefits on the table: Evidence from SNAP. *Journal of Public Economics* 179: 1-15.

Gregory, Christian, Michelle Ver Ploeg, Margaret Andrews, Alisha Coleman-Jensen. 2013. Supplemental Nutrition Assistance Program (SNAP) participation leads to modest changes in diet quality. United States Department of Agriculture Economic Research Service Economic Research Report No. 147.

Gregory, Christian, Matthew P. Rabbitt, and David C. Ribar. 2016. The Supplemental Nutrition Assistance Program and food insecurity. In J. Bartfeld, C. Gundersen, T. Smeeding, and J. P. Ziliak (eds.), *SNAP Matters: How Food Stamps Affect Health and Well Being*: 74-106. Stanford: Stanford University Press.

Griffith, Rachel, Martin O'Connell, and Kate Smith. 2012. The socioeconomic gradient in diet. Presentation to the National Bureau of Economic Research (NBER) Summer Institute.

Grummon, Anna H. and Lindsey Smith Taillie. 2017. Nutritional profile of Supplemental Nutrition Assistance Program household food and beverage purchases. *The American Journal of Clinical Nutrition* 105(6): 1433-1442.

Guenther, Patricia M., Sharon I. Kirkpatrick, Jill Reedy, Susan M. Krebs-Smith, Dennis W. Buckman, Kevin W. Dodd, Kellie O. Casavale, and Raymond J. Carroll. 2014. The Health

Eating Index-2010 is a valid and reliable measure of diet quality according to the 2010 Dietary Guidelines for Americans. *The Journal of Nutrition* 144(3): 399-407.

Gundersen, Craig. 2016. SNAP and obesity. In J. Bartfeld, C. Gundersen, T. Smeeding, and J. P. Ziliak (eds.), *SNAP Matters: How Food Stamps Affect Health and Well Being*: 161-185. Stanford: Stanford University Press.

Handbury, Jessie, Ilya Rahkovsky, and Molly Schnell. 2016. Is the focus on food deserts fruitless? Retail access and food purchases across the socioeconomic spectrum. NBER Working Paper 21126.

Hansen, Lars Peter. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50(4): 1029-1054.

Hansen, R. Garuth, Bonita W. Wyse, and Ann W. Sorenson. 1979. Nutritional quality index of foods. Westport, CT: AVI Publishing Co.

Hartline-Grafton, Heather. 2013. SNAP and public health: The role of the Supplemental Nutrition Assistance Program in improving the health and well-being of Americans. Food Research and Action Center.

Harvard T.H. Chan School of Public Health. 2017. Food Assistance Programs. Accessed at <https://www.hsph.harvard.edu/obesity-prevention-source/obesity-prevention/food-environment/food-assistance-programs/#source-list> on August 3, 2017.

Hastings, Justine S. and Jesse M. Shapiro. 2018. How are SNAP benefits spent? Evidence from a retail panel. *American Economic Review* 108(12): 3493-3540.

Hastings, Justine S., Mark Howison, Ted Lawless, John Ucles, and Preston White. 2019. Unlocking data to improve public policy. *Communications of the ACM* 62(10): 48-53.

Hoynes, Hilary W. and Diane Whitmore Schanzenbach. 2009. Consumption responses to in-kind transfers: Evidence from the introduction of the Food Stamp Program. *American Economic Journal: Applied Economics* 1(4): 109-139.

Hut, Stefan. 2019. Determinants of dietary choice in the US: Evidence from consumer migration. Working paper.

Hut, Stefan and Emily Oster. 2019. Changes in household diet: Determinants and predictability. Working paper.

Inoue, Atsushi and Gary Solon. 2010. Two-sample instrumental variables estimators. *Review of*

Economics and Statistics 92(3): 557-561.

Jetter, Karen M., and Diana L. Cassady. 2006. The availability and cost of healthier food alternatives. *American Journal of Preventive Medicine* 30(1): 38-44.

Klerman, Jacob A. and Caroline Danielson. 2011. The transformation of the Supplemental Nutrition Assistance Program. *Journal of Policy Analysis and Management* 30 (4): 863-888.

Kreider, Brent, John V. Pepper, Craig Gundersen, and Dean Jolliffe. 2012. Identifying the effects of SNAP (food stamps) on child health outcomes when participation is endogenous and misreported. *Journal of the American Statistical Association* 107(499): 958-975.

Lin, Biing-Hwan and Joanne Guthrie. 2012. Nutritional quality of food prepared at home and away from home, 1977-2008. United States Department of Agriculture Economic Research Service Economic Information Bulletin No. 105.

Liu, Miaoru, Panagiotis Kasteridis, and Steven T. Yen. 2013a. Breakfast, lunch, and dinner expenditures away from home in the United States. *Food Policy* 38: 156-164.

———. 2013b. Who are consuming food away from home and where? Results from the Consumer Expenditure Survey. *European Review of Agricultural Economics* 40(1): 191-213.

Mabli, James, Laura Castner, James Ohls, Mary Kay Fox, Mary Kay Crepinsek, and Elizabeth Condon. 2010. Food expenditures and diet quality among low-income households and individuals. Mathematica Policy Research.

Mancino, Lisa, Jessica E. Todd, and Benjamin Scharadin. 2018. USDA's National Household Food Acquisition and Purchase Survey: Methodology for imputing missing quantities to calculate Healthy Eating Index-2010 scores and sort foods into ERS food groups. United States Department of Agriculture Economic Research Service Technical Bulletin No. 1947.

Meyerhoefer, Chad D. and Muzhe Yang. 2011. The relationship between food assistance and health: A review of the literature and empirical strategies for identifying program effects. *Applied Economic Perspectives and Policy* 33(3): 304-344.

Mills, Gregory, Tracy Vericker, Heather Koball, Kye Lippold, Laura Wheaton, and Sam Elkin. 2014. Understanding the rates, causes, and costs of churning in the Supplemental Nutrition Assistance Program (SNAP): Final report. U.S. Department of Agriculture, Food and Nutrition Service.

Monsivais, Pablo, Anju Aggarwal, and Adam Drewnowski. 2012. Are socioeconomic disparities in diet quality explained by diet cost? *Journal of Epidemiology and Community Health* 66(6): 530-535.

Montiel Olea, José Luis and Mikkel Plagborg-Møller. 2019. Simultaneous confidence bands: Theory, implementation, and an application to SVARs. *Journal of Applied Econometrics* 34(1): 1-17.

Newey, Whitney K. and Daniel McFadden. 1994. Large sample estimation and hypothesis testing. In R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics* 4: 2111-2245. Amsterdam: Elsevier.

Oster, Emily. 2018. Diabetes and diet: Purchasing behavior change in response to health information. *American Economic Journal: Applied Economics* 10(4): 308-348.

Rao, Mayuree, Ashkan Afshin, Gitanjali Singh, and Dariush Mozaffarian. 2013. Do healthier foods and diet patterns cost more than less healthy options? A systematic review and meta-analysis. *BMJ Open* 3(12): 1-16.

Rehm, Colin D., Pablo Monsivais, and Adam Drewnowski. 2011. The quality and monetary value of diets consumed by adults in the United States. *The American Journal of Clinical Nutrition* 94(5): 1333-1339.

———. 2015. Relation between diet cost and Healthy Eating Index 2010 scores among adults in the United States 2007-2010. *Preventive Medicine* 73: 70-75.

Rhode Island Department of Labor and Training. 2016. 2016 UI and TDI Quick Reference. Accessed at <https://web.archive.org/web/20160104022814/http://www.dlt.ri.gov/lmi/news/quickref.htm> on December 22, 2016.

Schanzenbach, Diane Whitmore. 2013. Proposals to ban purchase of sugary drink with food stamps won't work. *The Christian Science Monitor*.

Scherpf, Erik and Benjamin Cerf. 2019. Local labor demand and program participation dynamics: Evidence from New York SNAP administrative records. *Journal of Policy Analysis and Management* 38(2): 394-425.

Tiehen, Laura, Constance Newman, and John A. Kirlin. 2017. The food-spending patterns of households participating in the Supplemental Nutrition Assistance Program: Findings from USDA's FoodAPS. United States Department of Agriculture Economic Research Service Economic Information Bulletin No. 176.

Todd, Jessica E. and Benjamin Scharadin. 2016. Where households get food in a typical week: Findings from USDA's FoodAPS. United States Department of Agriculture Economic Research Service Economic Information Bulletin No. 156.

Todd, Jessica E., Lisa Mancino, Ephraim Leibtag, and Christina Tripodo. 2010. Methodology behind the Quarterly Food-at-Home Price Database. United States Department of Agriculture Economic Research Service Technical Bulletin No. 1926.

Todd, Jessica E. and Michele Ver Ploeg. 2014. Caloric beverage intake among adult Supplemental Nutrition Assistance Program participants. *American Journal of Public Health* 104(9): e80-e85.

US Department of Agriculture (USDA). 2007. Thrifty Food Plan, 2006. Accessed at <https://fns-prod.azureedge.net/sites/default/files/usda_food_plans_cost_of_food/TFP2006Report.pdf> on February 25, 2020.

———. 2014. USDA Food and Nutrient Database for Dietary Studies 2011-2012. US Department of Agriculture Agricultural Research Service.

———. 2016a. USDA National Nutrient Database for Standard Reference, Release 28 (Slightly revised). US Department of Agriculture Agricultural Research Service Nutrient Data Laboratory.

———. 2016b. National Household Food Acquisition and Purchase Survey (FoodAPS): User's guide to survey design, data, collection, and overview of datasets. Accessed at <https://www.ers.usda.gov/data-products/foodaps-national-household-food-acquisition-and-purchase-survey/> on April 12, 2017.

———. 2016c. Quarterly Food-at-Home Price Database. Accessed at <https://www.ers.usda.gov/data-products/quarterly-food-at-home-price-database> on December 4, 2017.

———. 2018. USDA Branded Food Products Database. US Department of Agriculture Agricultural Research Service Nutrient Data Laboratory.

US Department of Health and Human Services (HHS) and US Department of Agriculture (USDA). 2010. Dietary Guidelines for Americans, 2010. 7th edition. Washington, D.C.: US Government Printing Office.

US Census Bureau. 2016. QuickFacts: United States. Accessed at <http://www.census.gov/quickfacts/fact/table/US/PST045216> on August 2, 2017.

Verghese, Anila, Margaret Raber, and Shreela Sharma. 2019. Interventions targeting diet quality of Supplemental Nutrition Assistance Program (SNAP) participants: A scoping review. *Preventive Medicine* 119: 77-86.

Volpe, Richard and Abigail Okrent. 2012. Assessing the healthfulness of consumers' grocery purchases. United States Department of Agriculture Economic Research Service Economic Information Bulletin No. 102.

Volpe, Richard, Abigail Okrent, and Ephraim Leibtag. 2013. The effect of supercenter-format stores on the healthfulness of consumers' grocery purchases. *American Journal of Agricultural Economics* 95(3): 568-589.

Wang, Dong D., Cindy W. Leung, Yanping Li, Eric L. Ding, Stephanie E. Chiuve, Frank B. Hu, and Walter C. Willett. 2014. Trends in dietary quality among adults in the United States, 1999 through 2010. *JAMA Internal Medicine* 174(10): 1587-1595.

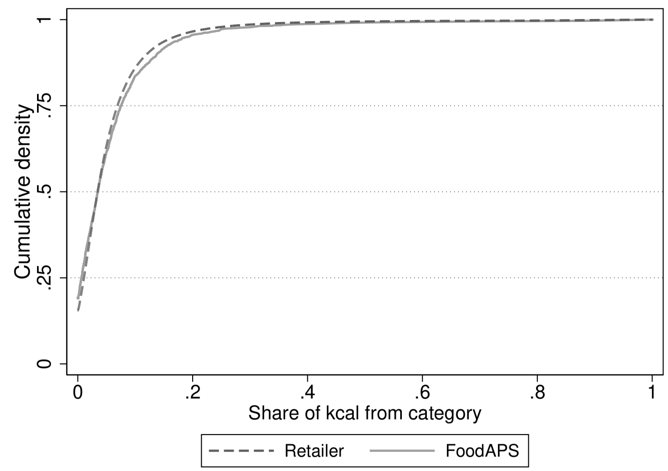
Woolf, Steven H., Paula Braveman, and Benjamin F. Evans. 2013. The health implications of reduced food stamp eligibility. Virginia Commonwealth University Center on Human Needs.

Yen, Steven T. 2010. The effects of SNAP and WIC programs on nutrient intakes of children. *Food Policy* 35(6): 576-583.

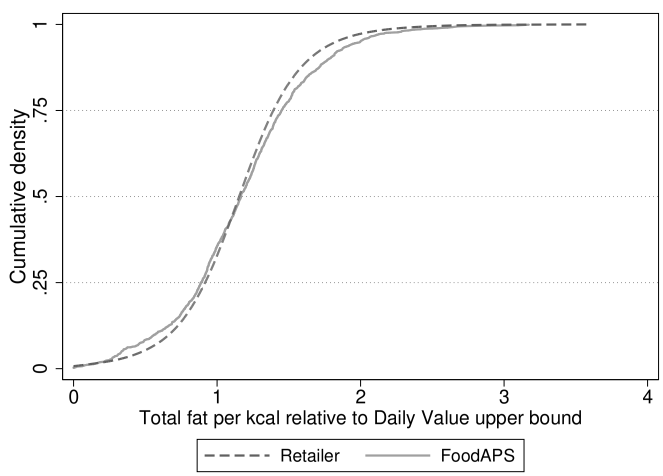
Yen, Steven T., Panagiotis P. Kasteridis, and John B. Riley. 2012. Food expenditures away from home by elderly households. 2012 Annual Meeting, August 12-14, 2012, Seattle, Washington 124981, Agricultural and Applied Economics Association.

Figure 1: Comparisons of Select Outcomes in the Retail Panel and FoodAPS Data

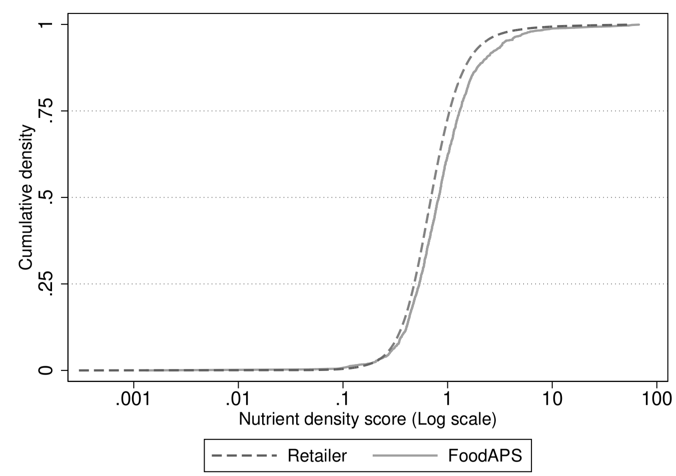
Panel A: Share of kcal from fruits and non-starchy vegetables



Panel B: Share of kcal from total fat relative to the Daily Value upper bound

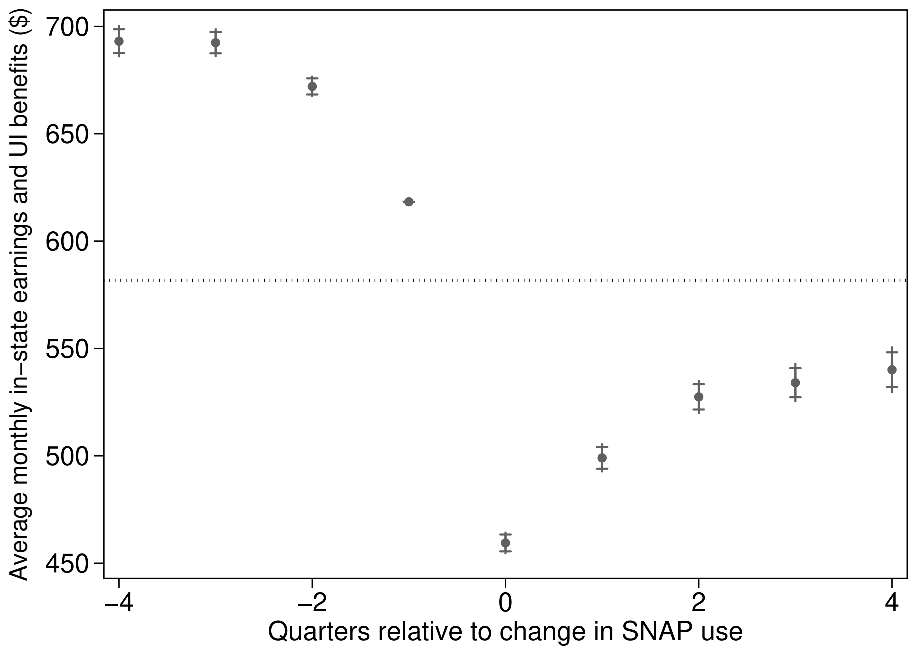


Panel C: Nutrient density score



Notes: Each panel plots the cumulative distribution function of a measure of food healthfulness across households in two different samples. For the line labeled "FoodAPS", the sample is the set of households surveyed in the FoodAPS data, described in Section 2.5. Each FoodAPS household is weighted according to the FoodAPS household weights such that the overall sample is nationally representative. For each household, the measure of food healthfulness is calculated from all observed food-at-home acquisitions during the survey week. For the line labeled "Retailer", the sample is all households in the retail panel during a randomly-assigned pseudo-survey week. Pseudo-survey weeks are randomly assigned to retailer households such that the distribution of pseudo-survey weeks in the retail panel equals the distribution of actual survey weeks in the FoodAPS data. For each household, the measure of food healthfulness is calculated from all food purchases at the retailer during their given pseudo-survey week. In panel A, the measure of healthfulness is the share of kilocalories from fruits and non-starchy vegetables, described in Section 2.4.1. In panel B, the measure of healthfulness is the share of kilocalories from total fat relative to the Daily Value upper bound, described in Section 2.4.2. In panel C, the measure of healthfulness is the nutrient density score, described in Section 2.4.2. The nutrient density score is shown on a log scale. The horizontal dotted lines intersect the 25th, 50th, and 75th percentiles of the distributions.

Figure 2: Dynamics of In-state Earnings Before and After Entry Into SNAP

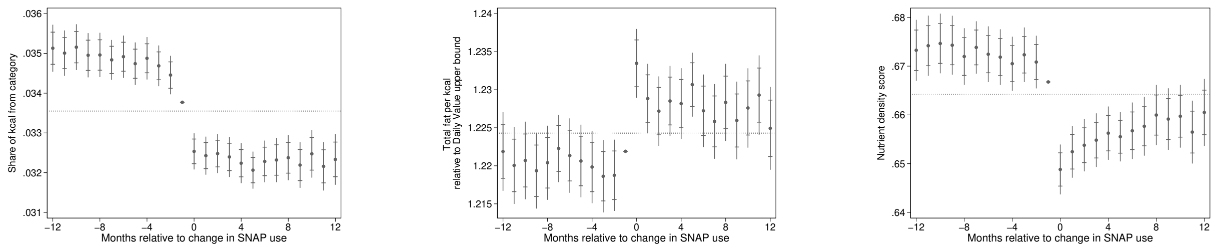


Notes: Data are from Rhode Island administrative records from the second quarter of 2006 through the fourth quarter of 2012. See Section 2.7 and Appendix A for details on sample definition and variable construction. The figure plots coefficient estimates from a two-stage least squares regression of average monthly in-state earnings plus unemployment insurance benefits on a vector of leads and lags of the contemporaneous change in SNAP use, with leads and lags of a contemporaneous indicator for whether the current quarter is a SNAP adoption quarter as excluded instruments. The sample is the set of SNAP adopters. The unit of observation is the household-quarter. Each regression includes controls for the sum of the change in SNAP use before the start of the plot window and after the end of the plot window, with the number of SNAP adoption quarters before the start of the plot window and after the end of the plot window as excluded instruments. The coefficient on the first lead of the contemporaneous change in SNAP use is normalized to zero. The change in SNAP use and the SNAP adoption indicator are treated as zero outside of the sample period. The coefficient estimates are shifted by a constant such that the mean of the coefficient estimates is equal to the mean of the outcome in the estimation sample. This mean is marked by a dotted line within the plot. The inner error bars represent 95 percent pointwise confidence intervals based on asymptotic standard errors clustered by household. The outer error bars represent 95 percent uniform sup-t confidence intervals computed as outlined in Montiel Olea and Plagborg-Møller (2019) based on an asymptotic variance-covariance matrix clustered by household.

Figure 3: Dynamics of Select Outcomes Before and After Entry Into SNAP

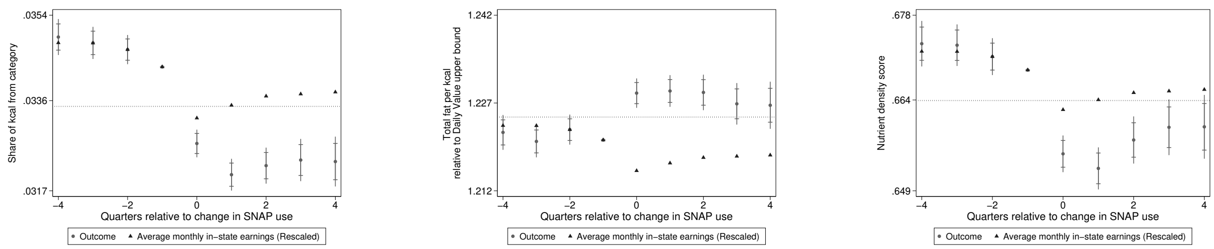
Panel A: Monthly frequency

(I) Share of kcal from fruits and non-starchy vegetables (II) Share of kcal from total fat relative to the Daily Value upper bound (III) Nutrient density score



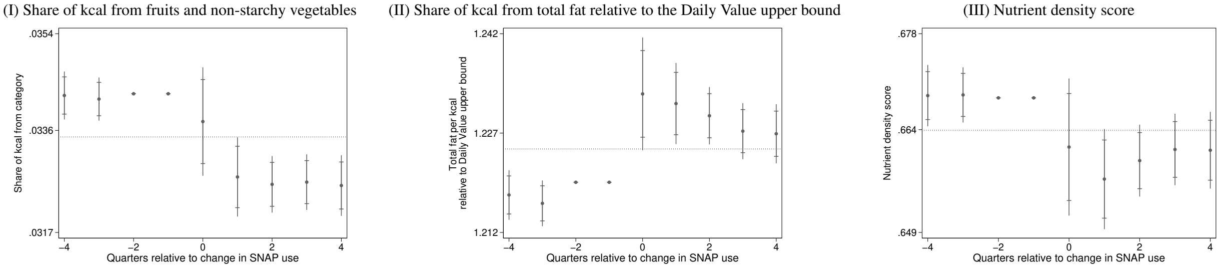
Panel B: Quarterly frequency with in-state earning dynamics

(I) Share of kcal from fruits and non-starchy vegetables (II) Share of kcal from total fat relative to the Daily Value upper bound (III) Nutrient density score



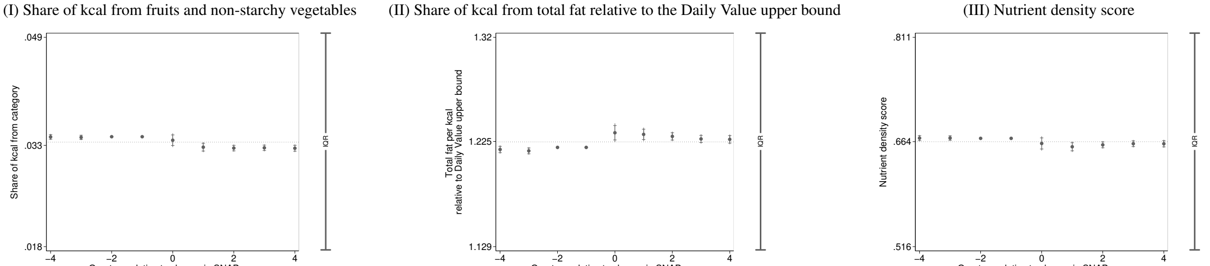
Panel C: Two stage least squares estimator

(I) Share of kcal from fruits and non-starchy vegetables (II) Share of kcal from total fat relative to the Daily Value upper bound (III) Nutrient density score



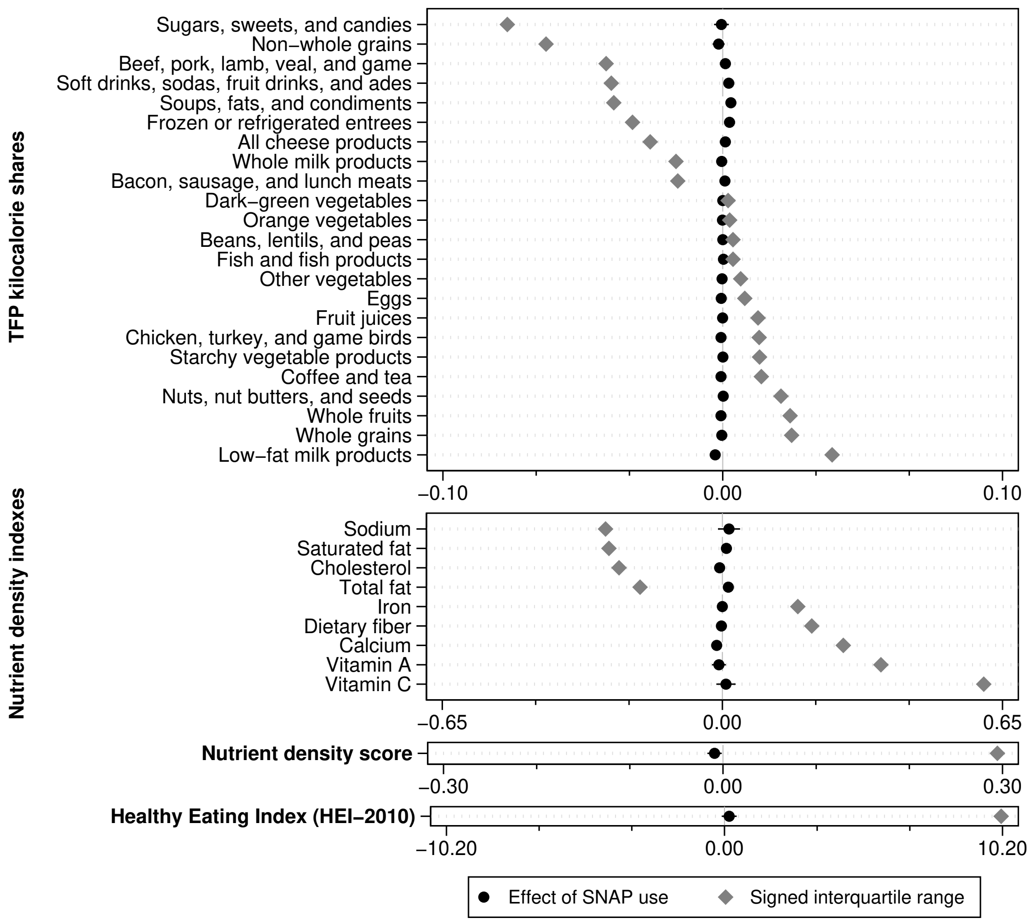
Panel D: Two stage least squares estimator relative to the IQR

(I) Share of kcal from fruits and non-starchy vegetables (II) Share of kcal from total fat relative to the Daily Value upper bound (III) Nutrient density score



Notes: Each figure plots coefficient estimates from a two-stage least squares regression of a measure of healthfulness on a vector of leads and lags of the contemporaneous change in SNAP use. The sample is the set of SNAP adopters. The unit of observation is the household-time period. Each regression includes controls for the sum of the change in SNAP use before the start of the plot window and after the end of the plot window, with the number of SNAP adoption periods before the start of the plot window and after the end of the plot window as excluded instruments. The change in SNAP use and the SNAP adoption indicator are treated as zero outside of the sample period. Each regression includes household and time period fixed effects. The coefficient estimates are shifted by a constant such that the mean of the coefficient estimates is equal to the mean of the outcome in the estimation sample. This mean is marked by a dotted line within each plot. The inner error bars represent 95 percent pointwise confidence intervals based on asymptotic standard errors clustered by household. The outer error bars represent 95 percent uniform sup-t confidence intervals computed as outlined in Montiel Olea and Plagborg-Møller (2019) based on an asymptotic variance-covariance matrix clustered by household. In panel A, the time period is a calendar month. In panels B-D, the time period is a calendar quarter. In panel A and panel B, the endogenous variables are a vector of leads and lags of the contemporaneous change in SNAP use, with leads and lags of a contemporaneous indicator for whether the current time period (i.e., month or quarter) is a SNAP adoption period as excluded instruments. The coefficient on the first lead of the contemporaneous change in SNAP use is normalized to zero. In panel B, in addition to the dynamics of the outcomes, the plots show the dynamics of in-state earnings (from Figure 2) rescaled such that the change in in-state earnings matches the change in the outcome between two and one periods prior to the change in SNAP use. In panel C and panel D, the estimates are based on the research design described in Section 3.1. The model is estimated in two samples using the TS2SLS estimator defined in Inoue and Solon (2010). Standard errors are calculated as outlined in Appendix B. The endogenous variables are a vector of leads and lags of the contemporaneous change in SNAP use and average monthly in-state earnings, with leads and lags of a contemporaneous indicator for whether the current quarter is a SNAP adoption quarter as excluded instruments. The first stage for in-state earnings is estimated on the sample of SNAP adopters in the Rhode Island administrative data described in Section 2.7. The first stage for the leads and lags of the contemporaneous change in SNAP use and the second stage are estimated in the retail panel. The coefficients on the first and second leads of the contemporaneous change in SNAP use are normalized to zero. In panel D, we repeat the plots in panel C, setting the y-axis range to be the interquartile range of the average of the outcome across calendar months for each retailer household. In the first column, the measure of healthfulness is the share of kilocalories from fruits and non-starchy vegetables, described in Section 2.4.1. In the second column, the measure of healthfulness is the share of kilocalories from total fat relative to the Daily Value upper bound, described in Section 2.4.2. In the third column, the measure of healthfulness is the nutrient density score, described in Section 2.4.2.

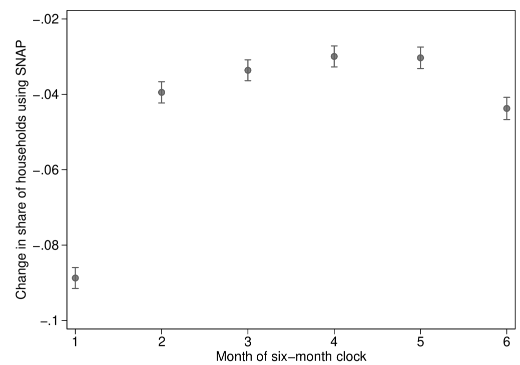
Figure 4: Effect of SNAP Use on Food Healthfulness, Program Adoption Research Design



Notes: Each box presents the signed interquartile range (IQR) of and the estimated effect of SNAP use on the given outcome(s). For the signed IQR series, the sample is all retailer households and the unit of observation is the household. For the estimated effect of SNAP use series, the sample is the set of SNAP adopters and the unit of observation is the household-quarter. For each outcome, the signed IQR is the IQR of the average of the outcome across calendar months for each household, signed to reflect a one IQR increase in food healthfulness. For each outcome, the causal effect of SNAP use on the change in the outcome is estimated in two samples using the TS2SLS estimator defined in Inoue and Solon (2010) in a model that includes calendar quarter fixed effects. Standard errors are calculated as outlined in Appendix B. The endogenous variables are the change in an indicator for whether the current quarter is a SNAP quarter and change in average monthly in-state earnings. The excluded instruments are an indicator for whether the current quarter is a SNAP adoption quarter and its first lead. The first stage for the change in in-state earnings is estimated on the sample of SNAP adopters in the Rhode Island administrative data described in Section 2.7. The first stage for the change in an indicator for whether the current quarter is a SNAP quarter and the second stage are estimated in the retail panel. In the first box, the outcomes are the shares of kilocalories going to each of the product categories that underlie the Thrifty Food Plan (TFP), and the IQR is signed according to the TFP healthfulness classification described in Appendix A. In the second box, outcomes are nutrient density indexes, and the IQR is signed according to whether the corresponding Daily Value bound represents a lower or upper bound. In the third and fourth boxes, the outcomes are the nutrient density score (NDS) and Healthy Eating Index (HEI-2010), respectively, and the IQR is signed to reflect the fact that both the NDS and the HEI-2010 are increasing in food healthfulness by construction. Error bars represent 95 percent confidence intervals based on asymptotic standard errors clustered by household.

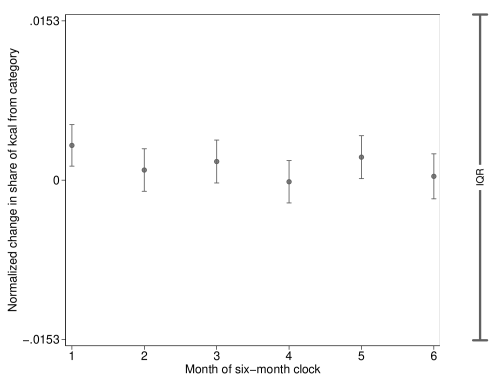
Figure 5: Dynamics of Select Outcomes Over the Six-month SNAP Clock

Panel A: SNAP use

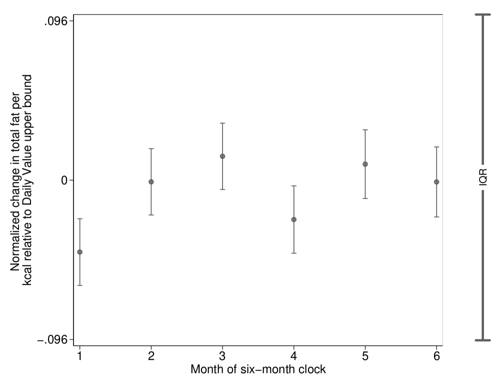


Panel B: Select outcomes

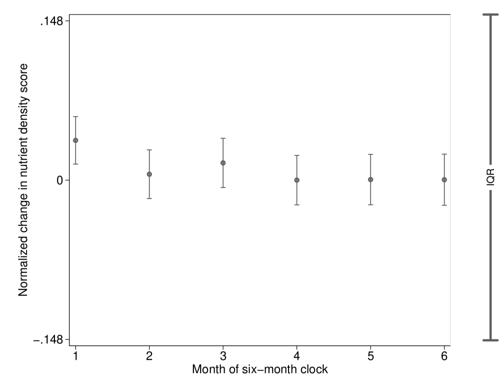
(I) Share of kcal from fruits and non-starchy vegetables



(II) Share of kcal from total fat relative to the Daily Value upper bound

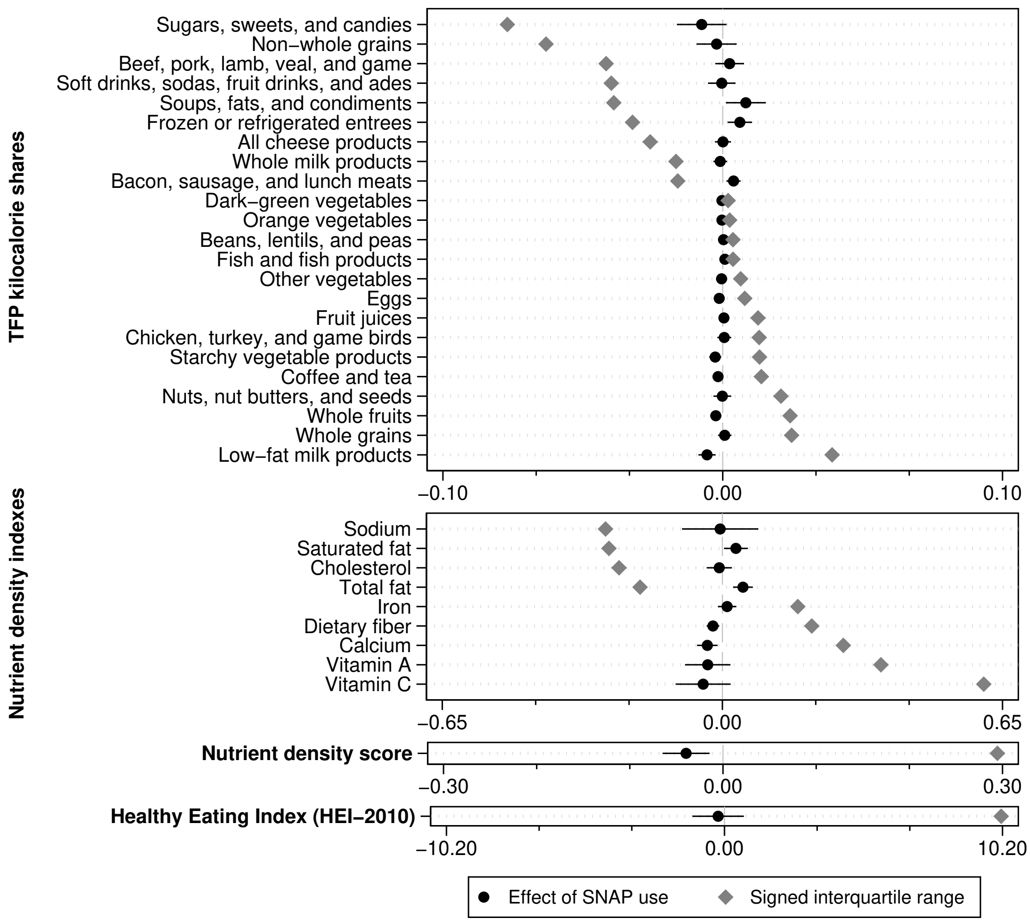


(III) Nutrient density score



Notes: Panel A plots coefficients from a regression of the change in an indicator for whether the current month is a SNAP month on a vector of indicators for the position of the current month in a monthly clock that begins in the most recent adoption month and resets every six months or at the next SNAP adoption, whichever comes first. So, for example, the first month of the clock corresponds to months 7, 13, 19, etc. following SNAP adoption. The sample is the set of SNAP adopters. The unit of observation for each regression is the household-month. Each regression includes calendar month fixed effects. The omitted category consists of the first six months (inclusive of the adoption month) after the household's most recent SNAP adoption, all months after the first 24 months (inclusive of the adoption month) following the household's most recent adoption, and all months for which there is no preceding adoption. Error bars represent 95 percent confidence intervals based on asymptotic standard errors clustered by household. For each outcome in panel B, we first divide the change in the outcome by the absolute value of the coefficient on clock month 1 in the regression from panel A. We then use this normalized change as the dependent variable in a regression specified analogously to that of panel A. In the first column of panel B, the outcome is the share of kilocalories from fruits and non-starchy vegetables, described in Section 2.4.1. In the second column of panel B, the outcome is the share of kilocalories from total fat relative to the Daily Value upper bound, described in Section 2.4.2. In the third column of panel B, the outcome is the nutrient density score, described in Section 2.4.2. In each column of panel B, the range of the y-axis is the interquartile range of the average of the outcome across calendar months for each retailer household.

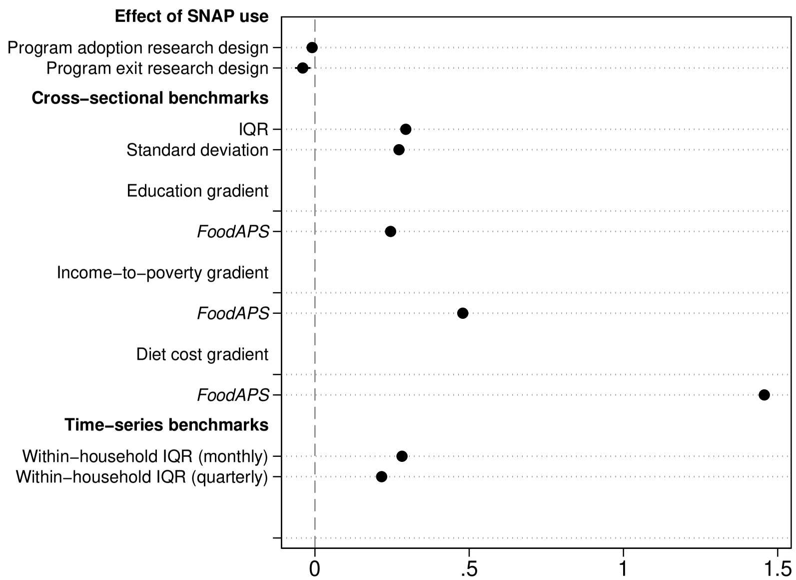
Figure 6: Effect of SNAP Use on Food Healthfulness, Program Exit Research Design



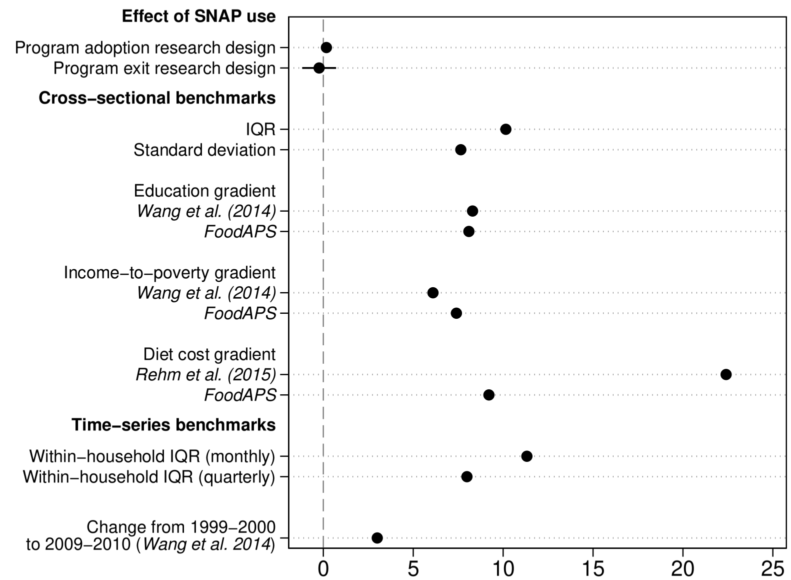
Notes: Each box presents the signed interquartile range (IQR) of and the estimated effect of SNAP use on the given outcome(s). For the signed IQR series, the sample is all retailer households and the unit of observation is the household. For the estimated effect of SNAP use series, the sample is the set of SNAP adopters and the unit of observation is the household-month. For each outcome, the signed IQR is the IQR of the average outcome across calendar months for each household, signed to reflect a one IQR increase in food healthfulness. For each outcome, the causal effect of SNAP use is estimated via a two-stage least squares regression of the change in the outcome on the change in an indicator for whether the current month is a SNAP month, with an indicator equal to one in the first month of a six-month clock that begins in the most recent adoption month as the excluded instrument and calendar month fixed effects as exogenous controls. The clock indicator is set to zero in the first six months (inclusive of the adoption month) following the most recent adoption, in any month after the first 24 months (inclusive of the adoption month) following the recent adoption, and in any month for which there is no preceding adoption. In the first box, the outcomes are the shares of kilocalories going to each of the product categories that underlie the Thrifty Food Plan (TFP), and the IQR is signed according to the TFP healthfulness classification described in Appendix A. In the second box, the outcomes are nutrient density indexes, and the IQR is signed according to whether the corresponding Daily Value bound represents a lower or upper bound. In the third and fourth boxes, the outcomes are the nutrient density score (NDS) and Healthy Eating Index (HEI-2010), respectively, and the IQRs are signed to reflect the fact that both the NDS and the HEI-2010 are increasing in food healthfulness by construction. Error bars represent 95 percent confidence intervals based on asymptotic standard errors clustered by household.

Figure 7: Effect of SNAP Use on Food Healthfulness, Comparison of Magnitudes

Panel A: Nutrient density score

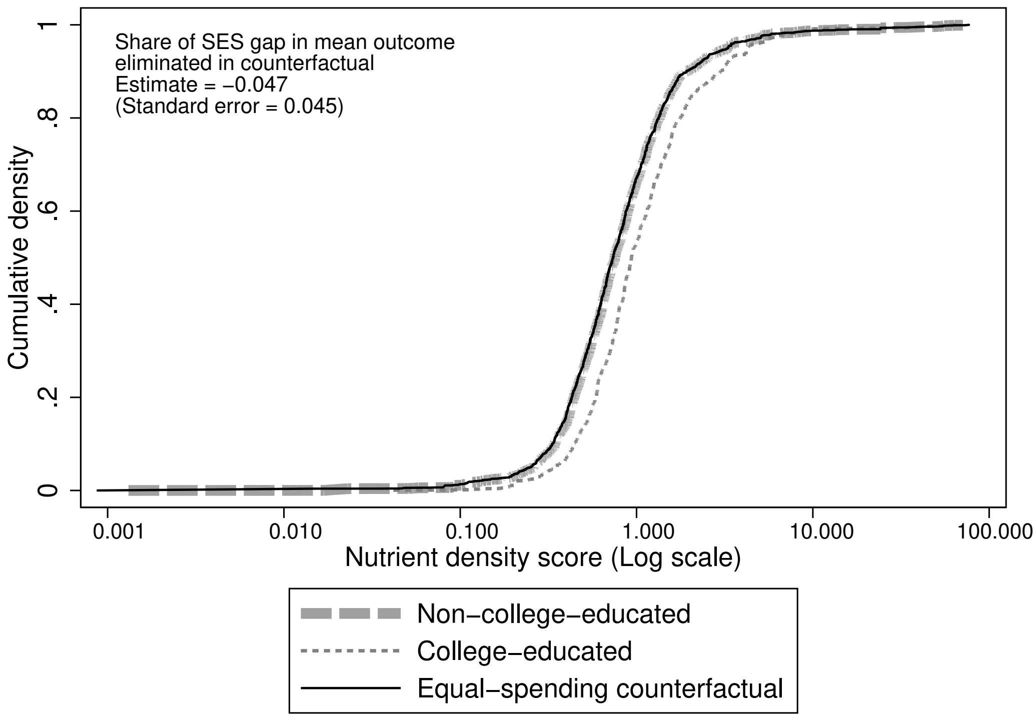


Panel B: Healthy Eating Index (HEI-2010)



Notes: Each figure plots estimates of the effect of SNAP use on a measure of healthfulness alongside benchmarks drawn from both our own calculations and from the literature. In panel A, the measure of healthfulness is the nutrient density score (NDS), described in Section 2.4.2. In panel B, the measure of healthfulness is the Healthy Eating Index (HEI-2010), described in Section 2.4.3. The "Program adoption research design" and "Program exit research design" rows report the estimates from Figures 4 and 6, respectively. Error bars represent 95 percent confidence intervals based on asymptotic standard errors clustered by household. The "IQR" and "Standard deviation" rows present the interquartile range (IQR) and standard deviation, respectively, of the average of the given outcome across calendar months for each household in the retail panel. The "Education gradient" rows present estimates of the difference in the mean of the given outcome between units with at least a bachelor's degree and units with less than a high school degree. For the "Wang et al. (2014)" row, the unit is an individual and the estimate is from the "2009-2010" column of Table 4 of the paper's supplementary online content. For the "FoodAPS" row, the unit is a household, the household's education level is defined to be the education level of the household's main food shopper or meal planner, and the estimate is computed using household survey weights. The "Income-to-poverty gradient" rows report the difference in the mean of the outcome between units whose income is greater than or equal to 350 percent of the federal poverty line and units whose income is less than 130 percent of the federal poverty line. For the "Wang et al. (2014)" row, the unit is an individual, the income measure is family income, and the estimate is from the "2009-2010" column of Table 4 of the paper's supplementary online content. For the "FoodAPS" row, the unit is a household, the income measure is average monthly household income, and the estimate is computed using household survey weights. The "Diet cost gradient" rows report the difference in the mean of the outcome between units in the top and bottom quintiles of energy-adjusted diet cost, defined as total food spending (on both food-at-home and food-away-from-home) per 2000 kilocalories. For the "Rehm et al. (2015)" row, the unit is an individual and the estimate is from Table 2 of the paper. For the "FoodAPS" row, the unit is a household and the estimate is computed using household survey weights. The "Within-household IQR (monthly)" and "Within-household IQR (quarterly)" rows report the average, across all households in the retail panel, of the household-level IQR of the given outcome over calendar months and calendar quarters, respectively. The "Change from 1999-2000 to 2009-2010 (Wang et al. 2014)" row reports the change in the average HEI-2010 between 1999-2000 and 2009-2010 reported in Table 4 of Wang et al. (2014)'s supplementary online content.

Figure 8: Role of Food Spending in Socioeconomic Disparities in the Nutrient Density Score



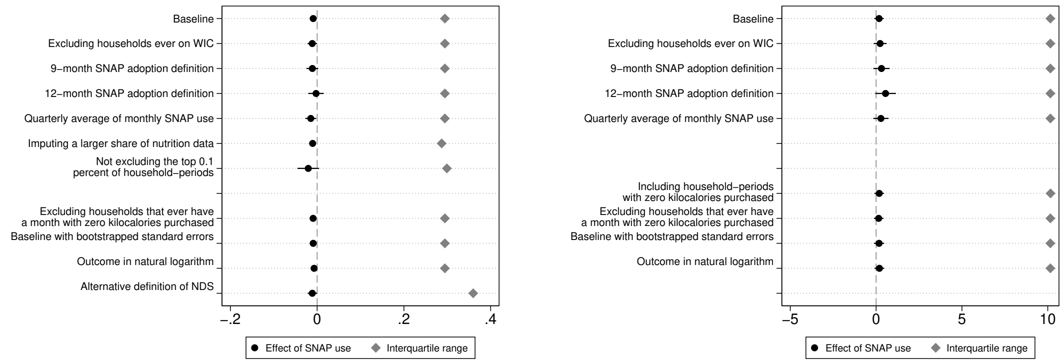
Notes: The figure plots cumulative distribution functions of the nutrient density score (NDS), described in Section 2.4.2, over a subset of households surveyed in the FoodAPS data, described in Section 2.5. The NDS is calculated from all food-at-home acquisitions during the survey week. The NDS is shown on a log scale. Each FoodAPS household is weighted according to the FoodAPS household weights such that the overall sample is nationally representative. For the line labeled "Non-college-educated," the sample is the set of FoodAPS households whose main food shopper or meal planner does not report having a bachelor's degree or higher. For the line labeled "College-educated," the sample is the set of FoodAPS households whose main food shopper or meal planner reports having a bachelor's degree or higher. For the line labeled "Equal-spending counterfactual," the sample is the set of FoodAPS households whose main food shopper or meal planner does not report having a bachelor's degree or higher. The "Equal-spending counterfactual" series is constructed as follows. First, among non-college-educated and college-educated households, we compute percentiles of each household by total food spending. Second, we assign to each non-college-educated household the food spending of the college-educated household at the closest percentile, breaking ties at random. We then use the estimates of the effect of food spending on the NDS from panel A of column (1) of Table 2 to compute counterfactual food healthfulness at the given counterfactual level of food spending. The "Share of SES gap in mean outcome eliminated in counterfactual" is the share of the difference in the average NDS across college-educated and non-college-educated households that would be eliminated if college-educated and non-college-educated households had the same average food spending. The share is estimated as the effect of food spending on the NDS (from panel A of column (1) of Table 2) times the difference in average food spending between college-educated and non-college-educated households divided by the difference in the average NDS between college-educated and non-college-educated households. The standard error associated with the estimated share is calculated via the delta method under the assumption that the estimate of the effect of food spending on the NDS is statistically independent from the estimated sample means.

Figure 9: Robustness of the Estimated Effect of SNAP Use

Panel A: Program adoption research design

(I) Nutrient density score

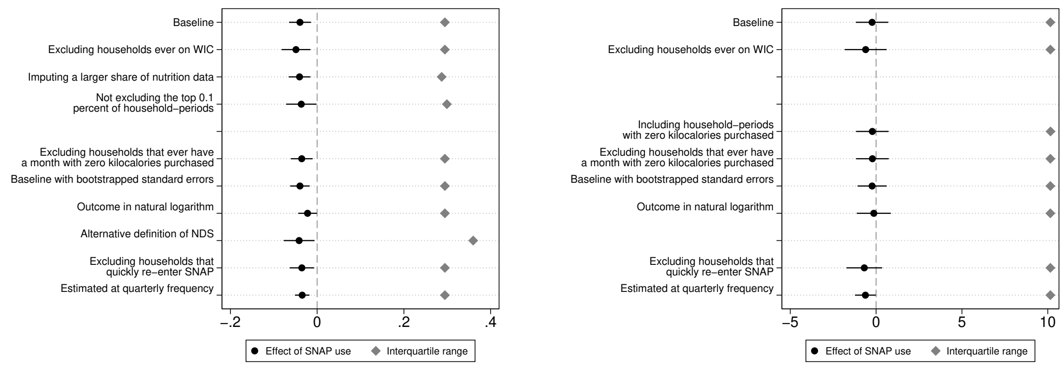
(II) Healthy Eating Index (HEI-2010)



Panel B: Program exit research design

(I) Nutrient density score

(II) Healthy Eating Index (HEI-2010)



Notes: Each panel presents the interquartile range (IQR) of and the effect of SNAP use on the given outcomes across a variety of specifications. For the interquartile range series, the sample is all retail households and the unit of observation is the household. For the effect of SNAP use series, the sample is described by the row label and the unit of observation is the household-time period. In panel A, the time period is a calendar quarter. In panel B, the time period is a calendar month. Error bars represent 95 percent confidence intervals based on asymptotic standard errors clustered by household. All models are estimated in first differences and include time period fixed effects. The estimates in panel A are based on the research design described in Section 3.1. The estimates in panel B are based on the research design described in Section 3.2. For each outcome, the IQR is the IQR of the average of the outcome across calendar months for each household. In the first column, the outcome is the nutrient density score. In the second column, the outcome is the Healthy Eating Index (HEI-2010). In both panels, the row labeled "Baseline" presents the results from Figure 4 and Figure 6 in panel A and panel B, respectively. In both panels, the row labeled "Excluding households ever on WIC" repeats the baseline specification using the sample of SNAP adopters who never use WIC in any transaction. In panel A, the row labeled "9-month SNAP adoption definition" repeats the baseline specification defining SNAP adoption as a period of nine or more consecutive non-SNAP months followed by a period of nine or more consecutive SNAP months. In panel A, the row labeled "12-month SNAP adoption definition" repeats the baseline specification defining SNAP adoption as a period of twelve or more consecutive non-SNAP months followed by a period of twelve or more consecutive SNAP months. In panel A, the row labeled "Quarterly average of monthly SNAP use" repeats the baseline specification defining SNAP use to be the quarterly average of monthly SNAP use. In the first column of both panels, the row labeled "Imputing a larger share of nutrition data" repeats the baseline specification using the alternative nutrition data assignment scheme described in Section 2.3. In the first column of both panels, the row labeled "Not excluding the top 0.1 percent of household-periods" repeats the baseline specification not excluding values of the nutrient density score above the 99.9th percentile. In the second column of both panels, the row labeled "Including household-periods with zero kilocalories purchased" repeats the baseline specification setting the HEI-2010 equal to zero for household-periods in which the household purchases zero kilocalories. In both panels, the row labeled "Excluding households that ever have a month with zero kilocalories purchased" repeats the baseline specification using the sample of SNAP adopters who never have a household-month in which zero kilocalories are purchased at the retailer during the sample period. In both panels, the row labeled "Baseline with bootstrapped standard errors" presents the results in the row labeled "Baseline" with standard errors calculated via a nonparametric bootstrap. In panel A, bootstrap standard errors are estimated as follows. In the Rhode Island administrative data, we sample 30 sets of households with replacement. For each set of households, we estimate the first stage regression of in-state earnings plus unemployment insurance benefits on the instruments. We then sample 30 sets of households in the retail panel with replacement, and randomly assign each set of households one of the 30 first-stage estimates obtained via the Rhode Island administrative data. We use these first-stage estimates to calculate the predicted in-state earnings plus unemployment insurance benefits in each of the 30 retailer replicates. We then estimate the second stage for each replicate and estimate the bootstrap standard error as the standard deviation of the second-stage estimates across the 30 bootstrap replicates. In panel B, bootstrap standard errors are estimated via a standard non-parametric bootstrap, with 30 replicates and sampling done by household with replacement. In both panels, the row labeled "Outcome in natural logarithm" repeats the baseline specification using the natural logarithm of the outcome as the dependent variable. For this specification, we report estimates of the average marginal effect. In the first column of both panels, the row labeled "Alternative definition of NDS" repeats the baseline specification defining the nutrient density score as $\delta_{it} = \left(\frac{1}{|\mathcal{N}_H|}\sum_{n\in\mathcal{N}_H}\delta_{it}^n\right) - \left(\frac{1}{|\mathcal{N}_U|}\sum_{n\in\mathcal{N}_U}\delta_{it}^n\right)$, with objects defined as in Section 2.4.2. In panel B, the row labeled "Excluding households that quickly re-enter SNAP" repeats the baseline specification using the sample of SNAP adopters who have at least one period of six consecutive non-SNAP months following the household's first SNAP adoption. In panel B, the row labeled "Estimated at quarterly frequency" repeats the baseline specification using quarterly averages of all variables.

Table 1: Coefficients on SNAP and In-state Earnings, Program Adoption Research Design

	(1) Nutrient density score	(2) Healthy Eating Index (HEI-2010)
Estimated coefficient on		
SNAP use $\left(\hat{\beta}\right)$	-0.0091	0.1734
	(0.0039)	(0.1400)
In-state earnings (in \$100) $\left(100\widehat{\left(\frac{\gamma}{\varphi}\right)}\right)$	0.0036	0.2755
	(0.0025)	(0.0885)
Number of household-quarters	611297	611363
Number of households	24456	24456
Interquartile range of outcome	0.2945	10.1522
Ratio of coefficient on in-state earnings	-0.4011	1.5893
to the coefficient on SNAP use $\left(100\widehat{\left(\frac{\gamma}{\varphi}\right)}\Big/\hat{\beta}\right)$	(0.4429)	(0.8048)
p-value for test of restriction that SNAP use and income affect the outcome only through food spending	0.4682	0.0057

Notes: The sample is the set of SNAP adopters. The unit of observation is the household-quarter. For each outcome, the coefficients on SNAP use and in-state earnings are estimated in two samples using the TS2SLS estimator defined in Inoue and Solon (2010). The endogenous variables are the change in an indicator for whether the current quarter is a SNAP quarter and the change in average monthly in-state earnings (in hundreds of dollars). The excluded instruments are an indicator for whether the current quarter is a SNAP adoption quarter and its first lead. The first stage for the change in in-state earnings is estimated on the sample of SNAP adopters in the Rhode Island administrative data described in Section 2.7. The first stage for the change in an indicator for whether the current quarter is a SNAP quarter and the second stage are estimated in the retail panel. All regressions include calendar quarter fixed effects. Standard errors on coefficients are calculated as outlined in Appendix B. Standard errors on the ratio of the coefficient on in-state earnings (in hundreds of dollars) to the coefficient on SNAP use are calculated via the delta method. The interquartile range (IQR) reported in each column represents the IQR of the cross-sectional distribution of the average outcome across all retail households. The final row presents the $p-$value from Hansen's overidentification test (Hansen 1982) based on estimates of a restricted model in which SNAP use and income are assumed to affect the outcome only through their effect on food spending at the retailer. We estimate the restricted model in the retail panel via a 2SLS regression of the change in the outcome on the change in food spending at the retailer, with an indicator for whether the current quarter is a SNAP adoption quarter and its first lead as excluded instruments and calendar quarter fixed effects as exogenous controls. In column (1), the dependent variable is the change in the nutrient density score, described in Section 2.4.2. Missing values arise from the trimming of extreme values as described in Section 2.4.2. In column (2), the dependent variable is the change in the Healthy Eating Index (HEI-2010), described in Section 2.4.3.

Table 2: Effect of Food-at-home Spending on Food Healthfulness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Outcome: Nutrient density score				Outcome: Healthy Eating Index (HEI-2010)			
	Panel A: Program adoption research design							
Effect of $100 increase in FAH spending	-0.0167	-0.0137	-0.0140	-0.0165	0.3198	0.2623	0.2674	0.3160
	(0.0072)	(0.0059)	(0.0060)	(0.0071)	(0.2584)	(0.2119)	(0.2160)	(0.2554)
Number of household-quarters	611297	611297	611297	611297	611363	611363	611363	611363
Number of households	24456	24456	24456	24456	24456	24456	24456	24456
	Panel B: Program exit research design							
Effect of $100 increase in FAH spending	-0.0368	-0.0302	-0.0305	-0.0330	-0.2154	-0.1767	-0.1784	-0.1929
	(0.0121)	(0.0099)	(0.0100)	(0.0109)	(0.4477)	(0.3672)	(0.3707)	(0.4009)
Number of household-months	2002182	2002182	2002182	2002182	2003811	2003811	2003811	2003811
Number of households	24456	24456	24456	24456	24456	24456	24456	24456
Interquartile range of outcome	0.295	0.295	0.295	0.295	10.152	10.152	10.152	10.152
Assumed retailer share of FAH spending when household is:								
Not on SNAP	1.000	0.820	0.818	0.800	1.000	0.820	0.818	0.800
On SNAP	1.000	0.820	0.820	0.820	1.000	0.820	0.820	0.820
Basis for assumed effect of SNAP on retailer share of FAH spending	No effect	No effect	Homescan lower bound	Homescan upper bound	No effect	No effect	Homescan lower bound	Homescan upper bound

Notes: The sample is the set of SNAP adopters. The unit of observation is the household-time period. In panel A, the time period is a calendar quarter. In panel B, the time period is a calendar month. Each column within panels A and B reports coefficient estimates from an instrumental variables regression, with standard errors in parentheses clustered by household. All models are estimated in first differences and include time period fixed effects. The estimates in panel A are based on the research design described in Section 3.1. The model is estimated in two samples using the TS2SLS estimator defined in Inoue and Solon (2010). Standard errors are calculated as outlined in Appendix B. In each column, the endogenous variables are food-at-home (FAH) spending (in hundreds of dollars), computed by dividing food spending at the retailer by the "On SNAP" share in SNAP months and the "Not on SNAP" share in non-SNAP months, and average monthly in-state earnings. The excluded instruments are the number of SNAP adoptions the household has experienced as of the given calendar quarter and its first lead. The first stage for in-state earnings is estimated on the sample of SNAP adopters in the Rhode Island administrative data described in Section 2.7. The first stage for FAH spending and the second stage are estimated in the retail panel. The estimates in panel B are based on the research design described in Section 3.2. In each column, the endogenous variable is FAH spending (in hundreds of dollars), computed by dividing food spending at the retailer by the "On SNAP" share in SNAP months and the "Not on SNAP" share in non-SNAP months. The excluded instrument is an indicator equal to one in the first month of a six-month clock that begins in the most recent adoption month. The indicator is set to zero in the first six months (inclusive of the adoption month) following the most recent adoption, in any month after the first 24 months (inclusive of the adoption month) following the most recent adoption, and in any month for which there is no preceding adoption. So, the first month of the clock corresponds to months 7, 13, 19, etc. following SNAP adoption. In columns (1) to (4), the dependent variable is the nutrient density score, described in Section 2.4.2. Missing values arise from the trimming of extreme values as described in Section 2.4.2. In columns (5) to (8), the dependent variable is the Healthy Eating Index (HEI-2010), described in Section 2.4.3. The interquartile range (IQR) reported in each column represents the IQR of the cross-sectional distribution of the average outcome across all retail households. In columns (1) and (5), we assume that all households devote all food spending to the retailer in all months. In columns (2) and (6), we assume that all households devote a constant share of food spending to the retailer, with the share given by the ratio of average SNAP benefits between retailer and SNAP Quality Control Data, as estimated in online Appendix Table 7 of Hastings and Shapiro (2018). In columns (3) and (7), we assume that the share of spending in SNAP months is the same as in columns (2) and (5), and that the difference in the share of spending between SNAP months and non-SNAP months is equal to the lower bound of the 95 percent confidence interval of the effect of SNAP participation on the share of spending devoted to the primary retailer reported in column (2) of Appendix Table 1 in Hastings and Shapiro (2018). In columns (4) and (8), we assume that the share of spending in SNAP months is the same as in columns (2) and (5), and that the difference in the share of spending between SNAP months and non-SNAP months is equal to the upper bound of the 95 percent confidence interval of the effect of SNAP participation on the share of spending devoted to the primary retailer reported in column (2) of Appendix Table 1 in Hastings and Shapiro (2018).

A Data Appendix

Linking Retailer Food Products to TFP categories

UPC Food Products

We assign UPC food products to TFP product categories in two steps. First, we join UPC food products to product categories in the Quarterly Food-at-Home Price Database (QFAHPD) (Todd et al. 2010) using a crosswalk between UPCs and QFAHPD product categories established by the USDA (Todd et al. 2010; USDA 2016c). Second, we join QFAHPD product categories to TFP product categories using the crosswalk between QFAHPD and TFP product categories established in Volpe and Okrent (2012).

The crosswalk between UPCs and QFAHPD product categories established by the USDA is based on version 2 of the QFAHPD, while the crosswalk between QFAHPD product categories and TFP product categories established in Volpe and Okrent (2012) is based on version 1 of the QFAHPD. Relative to version 1, version 2 has two additional product categories: "non-alcoholic diet carbonated beverages" and "unsweetened coffee and tea." We assign these version-2 product categories to the "soft drinks, sodas, fruit drinks, and ades (including rice beverages)" and "coffee and tea" TFP categories, respectively.

Due to imperfections inherent to the UPC-to-QFAHPD-to-TFP mapping, we make two changes to the TFP product categories. First, because the "all potato products" TFP category contains both potato products and other starchy vegetable products, we rename the category "all starchy vegetable products." Second, because the "soups" TFP category contains canned soups, sauces, and prepared foods we combine the "soups" and "fats and condiments" TFP categories into a single category which we denote as "soups, fats, and condiments."

Random-Weight Food Products

We assign random-weight food products directly to TFP product categories using retailer and TFP product category descriptors. The assignment was performed by hand by a coauthor and then refined following a review by a research assistant.

Classifying the Healthfulness of TFP Categories

We identify TFP product categories that are recommended by the 2010 Dietary Guidelines for Americans (DGAs) (HHS and USDA 2010) for increased consumption using the healthful classification of QFAHPD product categories established in Volpe et al. (2013).

In aggregating the QFAHPD-level healthful classification we encounter two issues. First, there are three TFP product categories that contain both healthful and unhealthful QFAHPD product categories: "all cheese (including cheese soup and sauce)," "beef, pork, veal, lamb, and game," and "fats and condiments." We follow Handbury et al. (2016) and mark all three categories as unhealthful. Second, since the healthful classification in Volpe et al. (2013) is based on version 1 of the QFAHPD, it does not suggest a classification for "coffee and tea." We mark this category as healthful.

Linking Retailer Food Products to the USDA SR28 and FNDDS

Here we outline our procedure for linking retailer products to food items in release 28 of the USDA National Nutrient Database for Standard Reference (SR28) (USDA 2016a) and the 2011-2012 version of the USDA Food and Nutrient Database for Dietary Studies (FNDDS) (USDA 2014).

We proceed in two rounds. In the first round, we link a subset of retailer products to the SR28. In the second round, we link the remaining products to the union of the SR28 and FNDDS.

Round 1: Linking Select Products to the USDA SR28

We link retailer food products in product categories for which the UPC-level nutrition data cover less than half of category food spending to the SR28 by hand in two steps.

First, we use retailer product category and SR28 food item descriptors to link retailer product categories to SR28 food items. Since some retailer product categories can reasonably be linked to multiple SR28 food items, we allow the mapping between product categories and food items to be one-to-many. For each link established between a retailer product category and an SR28 food item, we record the weight associated with a typical unit of the food item (e.g., the weight of a medium-sized banana).[28] We also flag product categories that contain nutritionally distinct products.

Second, for retailer product categories flagged in step one as containing nutritionally distinct products, we repeat step one at the product level, using retailer product and SR28 food item descriptors to link retailer products to SR28 food items, again allowing for the mapping to be one-to-many. We limit the scope of this second step to a subset of top-selling products, chosen to account for 95 percent of spending among the product categories flagged as containing nutritionally distinct products.

To help ensure a high degree of accuracy in our hand-coding, we complete steps one and two twice and reconcile any discrepancies. Given the product-category and product-level links, we assign to each retailer product the median nutritional content and weight across the SR28 food items to which it was linked, with priority given to the product-level links.

Round 2: Linking Remaining Products to the USDA SR28 and FNDDS

In round 1, we linked select retailer products to the SR28. In this round, we link the remaining products to the union of the SR28 and FNDDS.

We link retailer products to food items in the union of the SR28 and FNDDS by hand, mirroring our approach in step one of round one above. Unlike in round one, we do not capture information regarding weight, and we do not establish links at the product level.

To help ensure a high degree of accuracy in our hand-coding, all product categories are linked twice and discrepancies are reconciled. Given the product category level links, we assign to each retailer product the median nutritional content across the SR28 and/or FNDDS food items to which it was linked.

Obtaining Information on USDA Food Patterns

We obtain USDA Food Pattern information from the 2011-2012 version of the USDA Food Pattern Equivalents Ingredients Database (FPID) and the USDA Food Pattern Equivalents Database (FPED, Bowman et al. 2014). The FPID contains the amount of each of the thirty-seven USDA

[28]The weight of a typical unit is often identified by weight descriptors (`msre_desc`) containing "medium," "fruit," or the name of the item (e.g., "avocado," "melon," "potato").

Food Patterns per 100 edible grams of food items in the SR28. The FPED contains the same information but for food items in the FNDDS—a derivative of the SR28 containing more mixed-ingredient food items.

We join retailer products to food items in the FPID and FPED, using product category level links with the SR28 and FNDDS as crosswalks. We assign to each retailer product the median food pattern content across the FPID and/or FPED food items to which it was linked.

Rhode Island Administrative Data

We use Rhode Island state administrative records housed in a secure facility. Personally identifiable information has been removed from the data and replaced with anonymous identifiers that make it possible for researchers with approved access to join and analyze records associated with the same individual while preserving anonymity (Hastings et al. 2019).

The data include anonymized state SNAP records from April 2006 through December 2012, which indicate the months of benefit receipt and the collection of individuals associated with each household on SNAP in each month. We define a SNAP spell to be a contiguous period of benefit receipt. We assume that an individual belongs to the household of her most recent spell, does not change households between the end of any given spell and the start of the next spell, and belongs to the household of her first spell as of the start of the sample period. We determine each individual's age in each month, and we exclude from our sample any household whose adult (over 18) composition changes during the sample period. We also exclude from our sample any household whose membership we cannot uniquely identify in every month, which can occur either because we lack a unique identifier for an individual in the household or because a given individual is associated with multiple households in the same month.

The data also include anonymized administrative records of the state's unemployment insurance system joined via anonymized identifiers to the individuals in the SNAP records over the same period. We compute, for each household and quarter, the sum of total unemployment insurance benefits received by and total earnings reported for all individuals who are in the household as of the quarter's end. We exclude from our sample any household-quarter in which the household is not observed for all three months of the quarter. We also exclude from our sample any household-quarter in which the household's total quarterly earnings exceed the 99.9999th percentile or in which unemployment insurance benefits in any month of the quarter exceed three times the four-week equivalent of the 2016 maximum individual weekly benefit of $707 (Rhode Island Department of Labor and Training 2016).

B Inference for the TS2SLS Estimator

Let $\hat{\theta}_1$ denote the estimated coefficients on the excluded instruments in the first stage, with estimated variance $\hat{V}_1$. Let $\hat{\theta}_2$ denote the estimated structural parameters in the second stage, with unadjusted estimated variance $\hat{V}_2$. Let ∇_{21} denote the gradient of $\hat{\theta}_2$ with respect to $\hat{\theta}_1$ at the estimated value of the parameters. We compute the adjusted estimated variance $\hat{V}_2^*$ of $\hat{\theta}_2$ as

$$\hat{V}_2^* = \hat{V}_2 + \nabla_{21}\hat{V}_1\nabla_{21}'.$$

This follows Newey and McFadden (1994, equation 6.12) under the assumption of independent samples. Except where otherwise stated, we use asymptotic standard errors clustered at the household level for the inputs $\hat{V}_1$ and $\hat{V}_2$.