

# The Association Between SNAP Participation and the Composition of Purchased Foods in the Nielsen Homescan Consumer Panel

Justine Hastings  
*Brown University*  
*and NBER*

Ryan Kessler  
*Brown University*

Jesse M. Shapiro\*  
*Brown University*  
*and NBER*

May 2019

## Abstract

This note investigates the longitudinal association between SNAP participation and the composition and nutrient content of foods purchased for at-home consumption in data from the Nielsen Homescan Consumer Panel. We find little evidence of such an association for the measures that we study.

## 1 Data and definitions

Our discussion of the Nielsen Homescan Consumer Panel (NHCP) draws heavily on Bronnenberg et al. (2015) and Hastings and Shapiro (2018), and we sometimes quote these studies without specific attribution.

### 1.1 Purchases and SNAP participation

We obtained data from the NHCP from the Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at the University of Chicago Booth

---

\*E-mail: justine\_hastings@brown.edu, ryan\_kessler@brown.edu, jesse\_shapiro\_1@brown.edu. This note supplements Hastings et al. (2019). Researcher(s) own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

School of Business.<sup>1</sup> Panelist households are given an optical scanner and are asked to scan the barcode of every consumer packaged good they purchase, regardless of the store where it was purchased.<sup>2</sup>

Nielsen recruits its panelists by direct mail and through internet advertising, and provides incentives to recruit and retain panelists. Muth et al. (2007) and Kilts Center for Marketing (2016) describe the recruitment process in more detail.

We observe 663 million purchases made on 119 million purchase occasions by 158,830 households from January 2004 through September 2015.

For each product purchased on each shopping trip, we observe the date, the transaction price, the quantity of items purchased, and the total expenditure on the item.

We obtained from the Nielsen Company a quarterly supplement from the Homescan Panel Omnibus Survey. This supplement has been used to study the association between SNAP participation and the nutritional content of household grocery purchases (Grummon and Taillie 2017). The supplement is available for the fourth quarter of 2010 and for every other quarter from the fourth quarter of 2011 through the second quarter of 2015. It contains panelists' answers to the following question:

*Are you or anyone in your household currently using or have you ever used food stamps, which includes food stamp card or voucher or cash grant from the state for food (also known as Supplemental Nutritional Assistance Program (SNAP), Electronic Debit Card (EBT card))?*

*Please read all response options then select the one that best describes you.*

- 1) Currently using food stamps*
- 2) Have used food stamps, but not currently using them.*
- 3) Have never used food stamps.*

We define a *SNAP quarter* as any household-quarter in which the household's answer is "currently using food stamps." Of the household-quarters in our panel for which this question is asked, 7.3 percent are SNAP quarters.

---

<sup>1</sup>Information on the data is available at <http://research.chicagobooth.edu/nielsen/>.

<sup>2</sup>Beginning in 2007, a subset of panelists are asked to itemize purchases of products (e.g., produce) that do not have a barcode.

## 1.2 Product classification

The NCHP data include characteristics of each product purchased, including the Universal Product Code (UPC) if it has one, 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 observed in the data there are 1304 unique product categories. We exclude all products without a UPC from our calculations.

We classify UPC products as SNAP-eligible or SNAP-ineligible based on the product taxonomy and the guidelines for eligibility published on the USDA website (FNS 2017). Across all SNAP-eligible UPC products there are 608 unique product categories. We exclude all SNAP-ineligible UPC products from our calculations.

We classify SNAP-eligible UPC products according to the product categories underlying the USDA’s Thrifty Food Plan (TFP) (USDA 2007) using the procedure described in section 2.2 of Hastings et al. (2019).

## 1.3 Product nutritional information

We obtain nutritional information for SNAP-eligible UPC products from a nutritional database maintained by Information Resources, Inc (IRI), 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 at the large US grocery retailer used in the main analysis provided by the retailer. The processing of all five data sets is described in section 2.3 of Hastings et al. (2019).

For the remaining SNAP-eligible UPC 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.

For SNAP-eligible UPC products, we prioritize UPC-level data sources following the priority described in section 2.3 of Hastings et al. (2019), followed by the imputations. The UPC-level data sources and imputations provide nutritional information for SNAP-eligible UPC products

that account for 74.9 and 23.2 percent of SNAP-eligible UPC spending, respectively. We exclude from our analysis the products that account for the remaining 1.9 percent of SNAP-eligible UPC spending.

## **1.4 Monthly food spending, food attributes, and food healthfulness**

For each household in our panel, we calculate the quarterly average of monthly total UPC spending, monthly kilocalories, and monthly micronutrients purchased in each calendar quarter. We exclude from our analysis the 10 households who spend more than \$5,000 on UPC products in a single month. We also compute several measures of the healthfulness of purchased foods.

We calculate the share of kilocalories purchased going to each of the TFP product categories, as described in section 2.4.1 of Hastings et al. (2019).

We calculate a nutrient density index for each of the nutrients that are generally required to appear on the Nutrition Facts label and for which the FDA recommends either increased or limited consumption, as described in section 2.4.2 of Hastings et al. (2019).

We calculate a nutrient density score (NDS) to summarize the nutrient density indexes, as described in section 2.4.2 of Hastings et al. (2019). We exclude from our analysis values of the NDS above the 99.9th percentile.

## **2 Methods and Estimates**

Our unit of analysis is the household-quarter. 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. The estimated coefficient on the SNAP quarter indicator is our summary of the longitudinal association between SNAP participation and healthfulness.

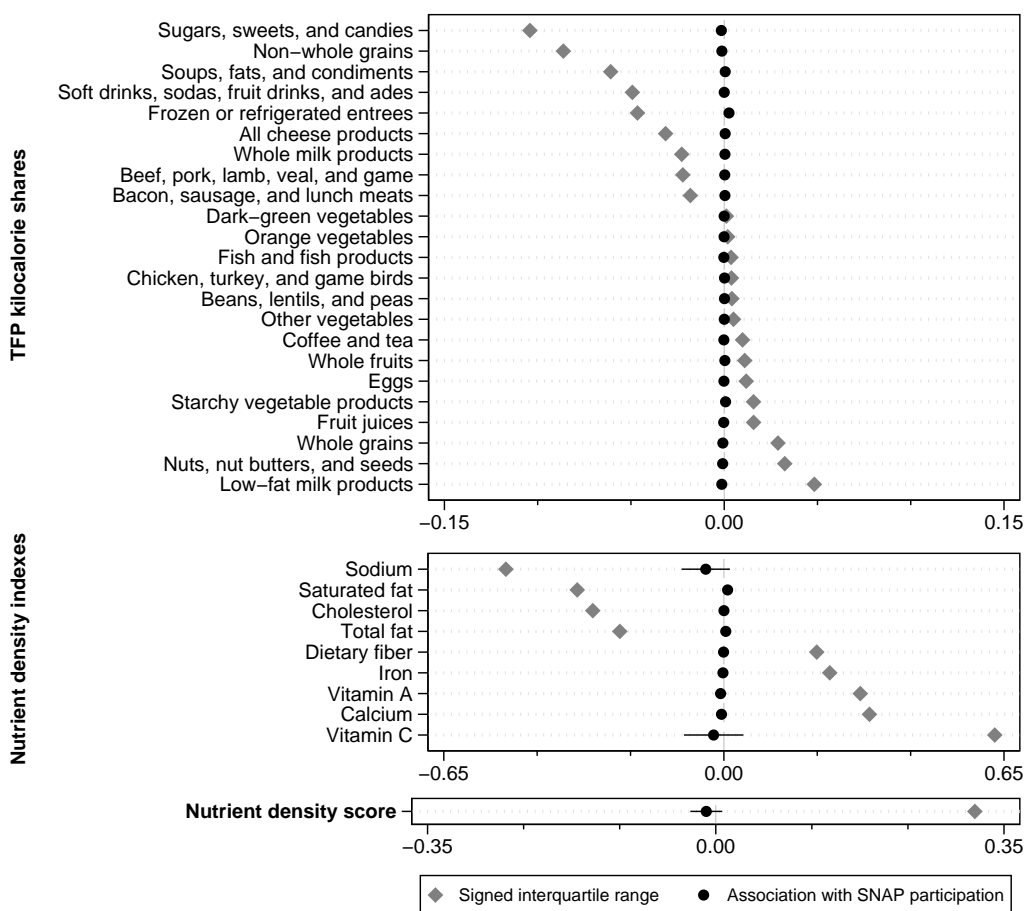
The figure presents the estimated coefficients for the full set of TFP kilocalorie shares, the full set of nutrient density indexes, and the NDS. For each outcome variable, we report the estimated coefficient, its confidence interval, and the cross-sectional interquartile range (IQR) of the average of the outcome variable, signed so that a positive IQR indicates that higher values of this outcome are associated with greater healthfulness.

The figure shows that, in most cases, the association between SNAP participation and healthfulness is small relative to the cross-sectional variation in measured healthfulness. For example, in the case of the NDS, the confidence interval on the estimated coefficient ranges from -0.0309 to 0.0079, or from -0.0982 to 0.0251 of an IQR. The figure also shows that, in most cases, the coefficient estimates are less precise than the corresponding estimate of the causal effect of SNAP reported in Figure 4 of Hastings et al. (2019). For example, for the NDS, the standard error on the estimated coefficient is 0.0099, which is larger than the standard error of 0.0039 on the estimated effect reported in Figure 4 of Hastings et al. (2019).

## References

- Bronnenberg, Bart J., Jean-Pierre Dube, Matthew Gentzkow, and Jesse M. Shapiro. 2015. Do pharmacists buy Bayer? Informed shoppers and the brand premium. *Quarterly Journal of Economics* 130(4): 1669-1726.
- Hastings, Justine, Ryan Kessler, and Jesse M. Shapiro. 2019. The effect of SNAP on the composition of purchased foods: Evidence and implications. *Brown University Working Paper*.
- Kilts Center for Marketing. 2016. Consumer panel dataset manual. The Kilts Center Archive of The Nielsen Company.
- Muth, Mary K., Peter H. Siegel, and Chen Zhen. 2007. Homescan data description. Excerpted from the *ERA Data Quality Study Design*. Accessed at <<http://qed.econ.queensu.ca/jae/2007-v22.7/hausman-leibtag/Homescan-data.pdf>> on October 19, 2017.

Figure: Association between SNAP participation and food healthfulness



Notes: Data come from the Nielsen Homescan Consumer Panel and the Homescan Panel Omnibus Survey. Each box presents the signed IQR of and the estimated association with SNAP participation on the given outcome(s). For the signed IQR, the sample is all households and the unit of observation is the household. For the estimated association with SNAP participation, the sample is the set of households that report being on SNAP in at least one survey response and report not being on SNAP in at least one survey response, 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 association with SNAP participation is estimated via a regression of the outcome on an indicator for whether the current quarter is a SNAP quarter, including fixed effects for household and calendar quarter. 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 of Hastings et al. (2019). 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 box, the outcome is the nutrient density score (NDS), and the IQR is signed to reflect the fact that the NDS is increasing in food healthfulness by construction. Error bars represent 95 percent confidence intervals based on asymptotic standard errors clustered by household.