The persistence of healthy behaviors in food purchasing*†

Marit Hinnosaar‡

June 2022

Abstract

When a policy gives temporary incentives for healthy behaviors, how long does the impact last? The U.S. Special Supplemental Nutrition Program for Women, Infants, and Children gives food vouchers to young children and their mothers. Using household-level scanner data, I study the reform of the program that introduced vouchers for healthier products. The difference-in-differences analysis shows that the reform makes purchases healthier during the program participation in the product categories most targeted by the reform (bread and milk). However, the effect is not always persistent. For bread, the effect decreases significantly within a couple of years after participants exit the program. Demand model estimates imply that price differences between healthy and unhealthy products play a large role in decreasing the program's impact. Therefore,

*Researcher(s) own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

†I thank Alex Mas, Andre Trindade, Andrea Pozzi, Aureo de Paula, Avi Goldfarb, Aviv Nevo, David Rivers, Erich Battistin, Garrett Johnson, Giovanni Mastrobuoni, Greg Veramendi, Isis Durrmeyer, Janet Currie, Jian-Da Zhu, Kim P. Huynh, Kosuke Uetake, Mario Pagliero, Michael Kummer, Toomas Hinnosaar, and seminar participants at the Collegio Carlo Alberto, Duke University, IFS, LSE, Royal Holloway, University of Bristol, University of East Anglia, University of Edinburgh, University of Essex, University of Houston, University of Nottingham, University of Pittsburgh, University of Tennessee, Risky Health Behaviors workshop at HCHE, Economics of Health Behaviours and Health Information workshop in London, IIOC, EARIE, JEI, Econometric Society European Winter Meeting, and Dondena Workshop on Public Policy for helpful comments and suggestions. Earlier versions of the paper were titled "How long do healthy habits last? The role of prices".

‡University of Nottingham and CEPR, e-mail: marit.hinnosaar@gmail.com.

some differences in the persistence can be explained by the relatively higher prices of healthy foods. Counterfactual analysis shows that a modest post-program subsidy might be a sustainable way to lengthen the program's impact and lead to long-term healthier purchases.

JEL codes: D12, I12, L66, I38

Keywords: Consumer behavior, dietary choices, long-term policy effects, WIC, food subsidy

1 Introduction

Obesity is a public health problem that leads to higher medical care costs and imposes negative externalities (Cawley, 2015). One cause of obesity is an unhealthy diet. While a diet might be an individual's optimal choice given budget and other constraints, the negative externalities of obesity provide a rationale for government intervention. Indeed, governments use various policies to improve diets, like fat and sugar taxes and regulation of nutrition labeling and advertising. To reduce costs, policies, such as vouchers and subsidies on healthy foods, often target specific products and give only temporary incentives. When designing and evaluating such policies, it is important to consider both their short- and long-term impacts.

This paper studies three questions. First, what is the impact of healthy food subsidies on consumption? Second, how long does the impact last after the subsidies end? Third, what role do prices have in affecting the persistence of the policy impact? Often, the healthy product is more expensive than its unhealthy substitute. A price-sensitive person, even when she otherwise prefers a healthy option, might not buy it because of the price difference. Therefore, temporary incentives might not change the long-run behavior even if these did change preferences.

Specifically, I study a U.S. federal program, the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). WIC gives vouchers for specific foods to children up to their fifth birthday and their mothers. The program provides a good setting for the study. The 2009 WIC policy reform changed the composition of food vouchers, introducing vouchers for healthier products. By providing free healthy products for a few years, the post-reform program gives temporary incentives to eat more healthily.

In the analysis, I use NielsenIQ household-level scanner data of grocery purchases. The dataset includes information on households' demographic characteristics and WIC status. I focus on the product categories included in the WIC food vouchers. To compute the product prices, I also use the NielsenIQ Retail Scanner Data.

To estimate the short- and long-term impacts of the healthy food subsidies, I use the WIC program reform that changed the composition of the food vouchers and introduced vouchers for healthier products. The advantage of the dataset is observing the same household in three states—before, while, and after receiving the WIC subsidies. I examine two groups of WIC households—those exposed to the old program and those exposed to the new program—and compare their within-household changes in purchases when the households start and stop receiving the vouchers. The main empirical strategy is a difference-in-differences analysis with household and time period fixed effects. A possible concern with the identification strategy is the existence of time-varying factors specific to the households that receive the post-reform WIC subsidies. While households starting to receive the subsidies either pre- or post-reform are likely to be subject to similar shocks (related to the child's birth), they might choose to stop receiving the vouchers because post-reform vouchers are different. To address the concern, I also examine changes in purchases when households become ineligible for the vouchers when the child becomes five years old. I perform a number of additional robustness checks, including providing the heterogeneity-robust difference-in-differences estimates proposed by de Chaisemartin and D'Haultfœuille (2020a,b) to correct for the possible bias of the two-way fixed effects estimator when treatment effects are heterogeneous, and propensity score matching estimators.

I find that the program reform makes purchases healthier during the program participation in the product categories most targeted by the reform. Specifically, it changed purchases of bread and milk, as the post-reform vouchers were restricted to whole wheat bread and low-fat milk. For bread, the effect decreases significantly within a couple of years after participants exit the program. For milk, the estimates are imprecise and don't allow ruling out that the effect is persistent. Typically, healthy products tend to be more expensive than unhealthy alternatives, but in the case of milk, in some regions, it is the opposite; that is, the healthy option is cheaper. Using exogenous regional differences in milk pricing, I show that when we restrict attention to states where the healthy option is not cheaper, the long-term impact is smaller than the short-term. The reform also modified the quantity of products in the vouchers. There are no statistically significant short- or long-term effects on the total quantities purchased of any of the products in the WIC vouchers: bread, milk, fruits and vegetables, juice, eggs, and cereals. But for several products, the estimates are imprecise and don't allow rejection of sizable positive or negative effects. Note that we don't necessarily expect sizable effects on total quantities because the vouchers are smaller than the typical consumption for several products.

To examine the long-term impact of healthy food subsidies on preferences and separate it from the role of prices, I estimate a demand model. I model demand using a discrete choice random coefficient multinomial logit model, allowing the long-term stock of past consumption to affect current tastes. Again, the policy reform provides an exogeneous variation in the past consumption, helping to identify its effect. Using the demand model estimates, I run a counterfactual analysis to study how to lengthen the program's impact. Demand model estimates imply that price differences between healthy and unhealthy products play a large role in decreasing the program's impact. Thus, some differences in the persistence of the impact can be explained by the relatively higher prices of healthy foods. Counterfactual analysis shows that a modest post-program subsidy might be a sustainable way to lengthen the program's impact and lead to long-term healthier purchases.

While the paper studies a specific program, many settings have similar concerns. For example, aid programs in developing countries provide free essential health products, which households should continue to obtain indefinitely, but programs' budget constraints prevent the perpetual supply of free goods. Combining temporarily free products with long-term subsidies might be a more sustainable way to lengthen the impact of the programs. Overall, the results imply that when a planner designs a public health policy or a firm chooses a product line, the trade-off between short-term and long-term effects should be considered.

The paper contributes to the literature in economics that studies nutritional choices. The literature has shown that prices matter (Dubois et al., 2014; Khan et al., 2016), but also that nutritional choices are persistent and difficult to change (Atkin, 2013, 2016; Ma et al., 2013; Oster, 2018; Allcott et al., 2019a).[1,2] The current paper contributes to the literature by analyzing the role of prices and the persistence of nutritional choices together when studying the impact of public policy.

The paper also adds a new finding to the literature on the long-term effects of temporary interventions. The evidence of long-term effects on health behaviors includes alcohol consumption (Kueng and Yakovlev, 2021), exercising (Charness and Gneezy, 2009; Royer et al., 2015), and smoking cessation (Giné et al., 2010). The growing experimental literature on the

[1]The finding of large persistence of nutritional choices is somewhat in contrast to the evidence from other choices, such as brand preferences and alcohol purchases. The past matters also for those (Eizenberg and Salvo, 2015; Sudhir and Tewari, 2015; Kueng and Yakovlev, 2021). But the extent is different, and there is evidence of forgetting (Mehta et al., 2004) and large adjustments to the new circumstances (Bronnenberg et al., 2012; Hinnosaar and Liu, 2021).

[2]There is also a large and growing literature on the impact of soda, sugar, and fat taxes including, among others, Fletcher et al. (2010); Wang (2015); Harding and Lovenheim (2017); Muller et al. (2017); Bollinger and Sexton (2018); Allcott et al. (2019b); Cawley et al. (2019); Dubois et al. (2020); Seiler et al. (2021). Other topics the nutritional choices literature has studied include the role of restaurants (Anderson and Matsa, 2011; Currie et al., 2010), nutritional information (Bollinger et al., 2011; Moorman et al., 2012; Puranam et al., 2017), bans on advertising (Dhar and Baylis, 2011; Dubois et al., 2018), and weight loss programs (Uetake and Yang, 2020).

long-term impacts of temporary incentives on nutritional choices has mixed results (including List and Samek, 2015; Belot et al., 2016; Loewenstein et al., 2016; Belot et al., 2018). The results in the paper are consistent with the work on informational treatments that has also found evidence of depreciation of the effects (e.g., Bronnenberg et al. (2020)). The current paper adds to the literature by illustrating the importance of prices in the persistence of temporary interventions.

The paper also contributes to the literature on the impact of the WIC program. Literature in economics studying the program has concentrated mostly on birth outcomes (Figlio et al., 2009; Hoynes et al., 2011), participation (Rossin-Slater, 2013), and the impact of WIC on market outcomes (Meckel, 2020).[3] Literature in nutrition science has studied the short-term effects of the 2009 change in the WIC food package on WIC participants' food consumption.[4] Griffith et al. (2018) provide evidence based on a similar program in the UK. Only a couple of studies address the persistence of WIC's impact on purchases. In particular, Khan et al. (2018) analyze the persistent impact focusing on brand loyalty in the breakfast cereals category. The current paper extends the analysis to food healthiness, which the literature has shown is more difficult to change than brand loyalty.[5] Frisvold et al. (2020) also use difference-in-differences analysis and the change in the WIC policy to estimate the impact of the WIC vouchers on purchases. In contrast to these papers, the identification in my analysis takes advantage of a longer sample period combined with a policy reform, which allows analysis of the same household before, while, and after receiving the WIC subsidies, comparing the changes in purchases of households receiving the subsidies before and after the policy reform. Furthermore, I provide evidence that the persistence of the program's impact depends on prices.

[3] Hoynes and Schanzenbach (2015) provide an overview of the literature on WIC.

[4] For overviews, see Schultz et al. (2015) and a report by the National Academies of Sciences, Engineering, and Medicine (NASEM, 2017).

[5] For example, Allcott et al. (2019a) find only small adjustments in food healthiness when people move, while Bronnenberg et al. (2012) find sizable adjustments in brand choices.

The rest of the paper is organized as follows. Section 2 describes the WIC program and data. Section 3 estimates the impact of healthy food subsidies and its persistence. Section 4 analyzes the role of prices. Section 5 concludes.

2 Background on WIC and data

2.1 Background on WIC

WIC is a U.S. in-kind food and nutrition assistance program targeted at mothers and children 0–4 years old. In 2016, eight million people participated in WIC per month, accounting for six percent of U.S. food and nutrition assistance expenditures (Oliveira, 2017). The average monthly WIC program cost for food per participant was 43 dollars (Oliveira, 2017).

The following categories of people are eligible to participate in WIC: pregnant, breastfeeding (up to the child's first birthday), and postpartum women (up to six months after birth); and children up to their fifth birthday.[6] The income eligibility requirement to participate in WIC is income not exceeding 185 percent of the federal poverty guidelines.[7] The income eligibility requirement is automatically satisfied if the individual or a family member is eligible for SNAP, Medicaid, or Temporary Assistance for Needy Families benefits (for more, see Oliveira, 2017). A large share of families with small children participate in WIC. According to the USDA report (Oliveira and Frazao, 2015), in 2012, 51 percent of infants (up to their first birthday) and almost 30 percent of pregnant and postpartum women and children aged 1–4 participated in WIC. Not everyone eligible participates in WIC. According to the USDA report (Johnson et al., 2014), in 2011, the rate at which eligible individuals participated in WIC was highest among infants (83%) and postpartum non-breastfeeding women (81%), and

[6] USDA. Food and Nutrition Service. "WIC Eligibility Requirements." Last published: August 3, 2016. https://www.fns.usda.gov/wic/wic-eligibility-requirements.

[7] States are allowed to decrease the eligible income cutoff, but in practice, in all states, it equals 185 percent of the federal poverty guidelines (Oliveira and Frazao, 2015).

lowest among children aged 1–4 (54%). Those who participate tend to have lower incomes.[8]

WIC provides quantity vouchers for specific food items (except for fruits and vegetables, which is a cash voucher). The set of foods in the WIC food package is small. Children aged 1–4 and women receive regular food items, which until 2009 included mainly milk, juice, cereal, eggs, legumes, and peanut butter. Infants, until their first birthday, receive infant formula, infant cereal, and infant juice.

In 2009, there was a major change in the types of foods provided by WIC. The goal of the change was to make WIC food packages consistent with new dietary guidelines. While WIC was established to fight malnutrition in low-income families, over time, obesity became a problem, and concerns were raised that WIC may contribute to child obesity (Oliveira and Frazao, 2009). The change added new foods: whole grain products, fruits, and vegetables. Specifically, it added whole wheat bread. The change also introduced restrictions on the fat percentage of milk. Specifically, whole milk was no longer allowed, except for one-year-old children. Altogether, the 2009 reform was cost neutral, as it reduced the quantities of some foods, in particular milk, juice, and eggs.[9] The WIC food packages, with the changes over the years, are shown in online appendix A.

The most controversial of the 2009 changes to the WIC food package was probably the restriction on the fat percentage of milk. For the past several decades, dietary guidelines have recommended decreasing dietary fat. But the medical and nutritional science literature has no consensus on whether that is beneficial (for examples of recent overviews of the debate, see Wang and Hu (2017); Ludwig et al. (2018)). However, the change in the WIC vouchers

[8] More than two-thirds of WIC participants have incomes below the federal poverty level (Oliveira and Frazao, 2015). Less than three percent of participants have incomes above the upper limit. Income can be above the limit because of automatic eligibility through other programs, which have higher income limits in some states.

[9] Additional small modifications to the food package were made by the 2014 change in the regulations. These changes further reduced milk fat percentage, allowing only low-fat (1%) or non-fat milk for children 2–4 years of age and women, and increased the cash voucher for fruits and vegetables. It also allowed states to authorize whole wheat macaroni (pasta) products as substitutes for whole wheat bread and allowed partial substitution of milk with yogurt.

was made based on the report commissioned from the Institute of Medicine (2006). The change was consistent with the recommendations by the American Academy of Pediatrics: whole milk for one-year-old children and 1% or non-fat for older children (Gidding et al., 2006). Therefore, despite the above-mentioned controversy, in this paper, to be consistent with the goal of the WIC reform, I call the restrictions to the milk fat percentage healthy.

2.2 Data

In the paper, the main data source is the NielsenIQ Consumer Panel. To compute the product prices, I also use the NielsenIQ Retail Scanner Data.[10] I use data that covers 11 years, 2006–2016. Nutritional information is obtained from the U.S. Department of Agriculture databases (USDA, 2018a,b).

The Consumer Panel is representative of the U.S. population. The participating households are asked to scan all their grocery purchases bought for personal at-home consumption. The dataset includes UPC-level information of the purchased quantities and expenditures. The reliability of the data was analyzed by Einav et al. (2010). In addition to purchases, the dataset has information on household demographic characteristics. Since 2006, the dataset has provided yearly information on households' self-reported WIC status.

I focus attention on categories that are in the standard WIC food packages for both mothers and children. These are milk, bread, fruits and vegetables, juice, eggs, and cereals.[11] The 2009 policy reform affected all these categories except cereals.[12] Specifically, the new

[10]Both datasets are from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. Information on availability and access to the data is available at https://www.chicagobooth.edu/research/kilts/datasets/nielsenIQ-nielsen.

[11]I don't analyze the impact on dry beans and peanut butter, which are also included in the food vouchers, but where a substitution is allowed between the products either by the household or in the state-level regulation.

[12]The 2009 policy reform included a minor change for cereals, which changed the choice set of available products. It required that at least half of the cereal brands available to WIC households must be whole grain. The reform did not require WIC households to purchase whole grain cereals.

WIC food packages added bread and fruits and vegetables and decreased the quantity of milk, juice, and eggs. Therefore, I will analyze the impact on the aggregate quantity of purchases in these categories (aggregating across brands, package sizes, and other product characteristics). Additionally, the reform required milk to be low-fat (except for one-year-old children) and bread to be whole wheat. Therefore, in the case of bread and milk, I also study the impact on the type of product purchased (fat percentage for milk and whether bread is whole wheat). Specifically, the analysis focuses on white dairy milk, loaves of bread, fruits and vegetables (including fresh, frozen, and canned), and juice (including concentrate and frozen). Milk is measured in gallons, bread in pounds, cereals in ounces, and eggs by count. To aggregate together the more heterogeneous products, different fruits and vegetables, and different types of juice (including concentrated), I measure the quantities of these categories in calories.[13]

Next, I outline the choices made in the construction of the datasets used in the estimation. Further details are in online appendix B.

Dataset for reduced form analysis. In section 3, the reduced form analysis focuses on households that report receiving WIC assistance and that were in the Consumer Panel the last year before, during, and the first and the second year after receiving the WIC vouchers. This restriction allows analysis of within-household changes when a household starts and stops receiving the WIC vouchers.

In 2006–2016, in the Consumer Panel, 3414 households self-reported currently receiving the WIC vouchers. In order to analyze within-household changes when households start receiving the vouchers and two years after they have stopped receiving the vouchers, one needs to observe the households for a minimum of four years (because WIC status in the

[13]In the analysis, I include only products with UPCs (barcodes). The products without UPCs are non-packaged items such as fruit, vegetables, meat, and baked goods. In the Nielsen dataset, only a small subset of households are asked to record products without UPCs. To use a bigger sample of households and make purchases of all households comparable, I exclude products without UPCs. Therefore, the measured quantity of fruits and vegetables is smaller than all fruit and vegetable purchases.

dataset is reported with yearly frequency). However, only 1336 of the households have been in the dataset for that long. Moreover, only 571 households are observed before, while, and after receiving the vouchers. Requiring that each household be observed the last year before and the first and second years after receiving the vouchers reduces the sample to 347 households. Table B1 in the online appendix summarizes how much each restriction decreases the sample.

How representative is the final sample of the WIC households in the Consumer Panel and the overall WIC population? Table B2 in the online appendix shows that the sample and the excluded households are similar regarding education, race, and marital status. But the households in the sample have a higher income (by 27%) and smaller size (by 13%), are less likely to be of Hispanic ethnicity (by 69%), and are less likely not to work full time (by 12%). It also shows that, while the Consumer Panel matches the administrative data of the WIC population rather well in terms of household size and race, it misses a large share of low-income and a majority of Hispanic households. As has been already noted in the literature (Khan et al., 2016), the average income of WIC households in the Consumer Panel is more than twice as high as in the administrative data. Also, over 40% of the WIC population is of Hispanic ethnicity, while the percentage of Hispanic households in the Consumer Panel is only about 12 and in the final sample only about four. Therefore, the results might not generalize to the whole WIC population, especially low-income and Hispanic households. However, the focus of the study isn't only on the WIC population but, more generally, learning about the impact of temporary incentives. The sample used in the analysis is somewhat more similar to the overall U.S. population than is a typical WIC participant.

Table 1 presents the demographic characteristics of the households in the sample, their regional locations, and summary statistics of their purchases. It compares the treatment group, that is, those that receive WIC assistance after the policy change (column 1), to those receiving it earlier (column 2), and to an additional control group of other households with children or lower income or both (column 3).[14] Specifically, the households in the additional control group, for at least one year, had a child (up to 17 years old) or had income not exceeding 200% of the federal poverty guidelines, or both. In the first year in the sample, 52% of households in the additional control group had children, 54% had lower income, and 20% had both children and lower income. As the reform coincides with the end of the recession, the question arises whether WIC households post-reform (treatment group) differ from households pre-reform (the control group in column 2). While the households are similar in income, household size, and marital status, the treatment group households are more likely to be college-educated (by 20%), white (by 13%), and not employed full time (by 25%). However, none of the differences are statistically significant at the 1% level in the Wilcoxon rank test.[15] In the analysis, the differences will be absorbed by household fixed effects. Figures B3–B6 in the online appendix present the distribution of purchases and changes over time.

Not all households purchase all the food products in WIC vouchers each quarter. When the outcome variable is the healthy bread or milk fat percentage, then the sample includes only households for whom the quarterly variable is defined (who purchase bread or milk, respectively) at least once per year and in total in at least two quarters. This restriction is imposed because we cannot use the household in the analysis if the percentage is not defined. When the outcome variable measures the quantity, then the sample includes all households that buy any food products with UPCs in a given year. The restriction excludes 104 household years by 57 households from the additional control group of households with children or lower income or both.[16] This restriction is imposed because when households don't

[14] Online appendix C shows that results remain similar if the non-WIC households are excluded.

[15] Also, none of the differences in demographic characteristics in table 1 between the treatment (column 1) and the second control group (column 3) are significant at the 1% level in the Wilcoxon rank test, except for household size.

[16] Sixteen of the 57 households never buy food products with UPCs. They are excluded from all the analysis, including table 1.

Table 1: Summary statistics

	Treatment group	Control groups	
	WIC post-reform	WIC pre-reform	Children or low-income
	(1)	(2)	(3)
Panel A: Demographic characteristics			
Household income (thousands)	52.9	53.3	54.1
Household size	3.4	3.5	3.0
Race: white	0.755	0.670	0.815
Ethnicity: Hispanic	0.046	0.047	0.073
College	0.407	0.340	0.474
Married	0.730	0.717	0.652
Not employed full time	0.705	0.566	0.701
Panel B: Region			
Northeast	10.8	17.9	16.5
Midwest	29.0	27.4	26.4
South	39.4	41.5	37.9
West	20.7	13.2	19.2
Panel C: Purchases (per household/quarter)			
Bread (lb)	10.53	10.79	8.10
Bread healthy %	9.25	6.86	9.99
Milk (gallons)	9.30	9.03	6.61
Milk fat %	1.99	2.02	1.83
Fruits & vegetables (kcal)	8207.02	7936.33	7328.92
Juice (kcal)	5323.15	4827.27	3713.31
Eggs (count)	66.11	63.89	50.92
Cereals (oz)	161.97	159.00	130.00
Number of households	241	106	96290

Note: Panel A presents the sample average demographic characteristics. Panel B shows the regional distribution of households. Panel C presents households' quarterly averages of purchases. All variables are measured during each household's first year in the sample. An observation is a household. Sample consists of households in the panel data fixed effects regressions. Column 1 includes households who received WIC assistance after the 2009 policy change. Column 2 includes households who received WIC assistance only before the 2009 policy change. Column 3 includes households who have children or lower income (income below 200 percent of the federal poverty guidelines) or both. Income is deflated to 2015 dollars using the consumer price index for urban consumers. *Not employed full time* indicates whether at least one household head is not employed full time. Purchases do not include Nielsen random weight products. *Bread healthy %* is the percentage of whole wheat bread in all bread purchases.

report any food purchases with UPCs, they are not suitable for studying the changes in WIC product purchases. These households either purchase only the food products without UPCs (such as unpackaged fruits and vegetables, meat, or baked goods), only eat in restaurants, or don't report food purchases.

Figures B7–B8 in the online appendix provide a preview of the findings in the next section.

The cohorts entering the program pre- versus post-reform are rather similar regarding the levels of purchases before receiving the vouchers. While receiving WIC post-reform vouchers, healthy bread share increased and milk fat percentage decreased, but the effect reversed when households stopped receiving the vouchers.

Dataset to estimate the demand model. In section 4, to keep the demand model estimation manageable, I restrict attention to the sample of households that all have reported receiving WIC assistance (households in columns 1–2 of table 1). The demand model is estimated using only the time periods before and after receiving WIC vouchers.[17] I exclude households that don't have a sufficient number of bread purchases before and after receiving the vouchers.

The demand model is estimated using data only on the bread category (for reasons that will be discussed in section 4). I aggregate all types of bread into four products. *White bread* is a non-whole-grain bread made from white refined flour. Following WIC regulation, I classify bread as *whole wheat bread* if it contains 100% whole wheat flour. I classify bread that is only partially made with whole wheat flour or that contains any other type of whole grain flour or rye flour *whole grain bread.* Finally, *other bread* aggregates together all other non-whole-grain bread types (like Cinnamon Raisin, Cinnamon Swirl, etc.).

Bread prices are constructed from the Retail Scanner Data for shopping trips to retail chains in the Retail Scanner dataset. To do that, prices are aggregated to the bread type, week, market (Nielsen scantrack market), and chain level. But more than half of the WIC households' bread purchases are at retailers (grocery stores and mass merchandisers) that are not in the Retail Scanner Data. For those shopping trips, the prices are constructed from the Consumer Panel data and prices are aggregated to the bread type, week, and market

[17] While households received WIC vouchers, it is unknown whether they paid with a voucher or cash on a given shopping trip. As such, the prices paid on a given shopping trip are unknown—whether these were the regular prices, or a price of zero because of the voucher.

(Nielsen scantrack market) level. In both cases, prices are deflated to 2015 dollars using the consumer price index for urban consumers. How different are the prices constructed from the two alternative datasets? We can analyze this for the subset of retailers that overlap in these datasets. Online appendix B.3 presents this comparison. For these overlapping retailers, bread prices constructed based on the two datasets line up rather closely; prices based on the Consumer Panel are about 10% lower for all four product types (column 5 in table B3). The magnitude of the difference is similar to what has been shown for aggregate grocery price indexes constructed based on the two datasets (for example, Seo (2019)). The mismeasurement of prices is unlikely to be a substantial concern because the difference is small, and importantly, in any given choice occasion (shopping trip), consumers face prices constructed in the same way. To alleviate the potential concerns, I analyze the robustness of results to the alternative price measures later in section 4, and the main estimates are similar to those obtained when using only prices from the Retail Scanner data.

Table 2 presents summary statistics of the prices of each product. The cheapest product is white bread, while whole wheat bread is about 60 cents more expensive.

Table 2: Summary statistics of bread prices

	Mean	SD	Perc. 25	Median	Perc. 75	N
	(1)	(2)	(3)	(4)	(5)	(6)
Other bread	1.76	0.33	1.53	1.67	1.95	28042
Whole grain bread	2.03	0.30	1.83	1.95	2.21	28042
Whole wheat bread	1.95	0.34	1.70	1.88	2.17	28042
White bread	1.36	0.28	1.17	1.30	1.49	28042

Note: The price of bread is measured in dollars per pound, deflated to 2015 dollars using the consumer price index for urban consumers. Observations are at the week, market, bread type, and (aggregate) retailer level.

3 The short- and long-term impacts of healthy food subsidies

3.1 Main empirical strategy

The goal of the empirical strategy is to identify the impact of the healthy food subsidies on WIC households, both in the short-term (while receiving the subsidies) and long-term (1–2 years after receiving the subsidies). The empirical strategy exploits the 2009 reform of the WIC program that changed the content of food vouchers. The analysis compares changes in purchases associated with starting and ending receipt of WIC vouchers in the two groups of households—those receiving the vouchers from the old program and those from the new program. These two groups of households are otherwise similar, but the vouchers are different. The new policy introduced healthy food subsidies—vouchers for whole wheat bread, low-fat milk, and fruits and vegetables. It reduced the subsidized quantity of milk, juice, and eggs while keeping the subsidized quantity of cereals unchanged.

In the main analysis, using household-level quarterly data, I estimate the following difference-in-differences regression:[18]

$$
\begin{aligned}
Y_{it} = {} & \alpha_1 WIC_{it} + \alpha_2 AfterWICYears1,2_{it} \\
& + \beta_1 ReformedWIC_{it} + \beta_2 AfterReformedWICYears1,2_{it} \\
& + \alpha_3 AfterWICYear3^+_{it} + \beta_3 AfterReformedWICYear3^+_{it} \\
& + \sigma \cdot PostReform_{it} + X_{it}\zeta + \delta_i + \gamma_t + \varepsilon_{it}.
\end{aligned} \tag{1}
$$

The outcome variable Y_{it} is a measure of purchases (e.g., average fat percentage of milk)

[18] To address concerns raised by the recent econometrics literature regarding staggered adoption difference-in-differences designs, I also calculate the heterogeneity-robust difference-in-differences estimators proposed by de Chaisemartin and D'Haultfœuille (2020a,b). As a robustness check, I also provide the propensity score matching estimator.

of household i in time period t. Variable WIC_{it} is an indicator for household i receiving WIC vouchers in period t. Variable $AfterWICYears1,2_{it}$ indicates that household i finished receiving WIC vouchers 1–2 years earlier. Variable $PostReform_{it}$ is an indicator for time periods after the WIC reform. The variable is not collinear with time period fixed effects because the 2009 reform took place in different months in different states. Therefore, the variable is also indexed by i.

The first variable of interest $ReformedWIC_{it}$ is an indicator of receiving the vouchers post reform and therefore equals $ReformedWIC_{it} = PostReform_{it} \times WIC_{it}$. The second variable of interest $AfterReformedWICYear1,2_{it}$ indicates that household i finished receiving reformed WIC vouchers 1–2 years earlier.[19]

The coefficients of interest β_1 and β_2 measure the short- and long-term effects of healthy food subsidies (introduced by the policy change) on WIC households. Therefore, the coefficients measure the average treatment effect on treated of the policy change.

In the balanced panel (up to two years after vouchers) of WIC households, only a few remain in the Nielsen panel for more than two years after finishing receiving the vouchers. As is standard in the literature, to increase statistical power, I don't drop the observations three or more years after receiving the vouchers. Instead, I include separate dummies for these periods. As such, $AfterWICYears3^+_{it}$ and $AfterReformedWICYear3^+_{it}$ indicate receiving WIC and reformed WIC vouchers three or more years earlier. I don't report the coefficient estimates because the sample is balanced only up to two years after receiving the vouchers. Therefore, these are biased estimates for these effects and cannot be interpreted.

The regressions include time-varying household characteristics X_{it}: logarithm of income, household size, and indicator variables for the year before the child is born, and children aged 0, 1, ..., 5, 6–12, and 13–17. For milk, the regression also includes interaction terms for

[19]The first two years after receiving the vouchers are combined to increase statistical power. However, online appendix C.1 shows that the results remain qualitatively similar when the years are separated.

one-year-old child and WIC and reformed WIC to take into account that, for one-year-old children, the reformed WIC food vouchers make an exception and provide whole milk.[20] The regression also includes household fixed effects δ_i and time period fixed effects γ_t.

Each WIC household in the analysis is observed before, while, and the first and second years after receiving WIC vouchers. In the main analysis, I also include an additional control group that don't receive WIC vouchers (households in column 3 of table 1) to control for possible changes in preferences over time. However, the results remain similar when restricting the sample to households that receive pre- or post-reform vouchers.

Identification. The identification assumption in the difference-in-differences analysis is that the purchases of the households who received post-reform WIC subsidies would have otherwise followed the same trend as the purchases of the households in the control group (who either received the pre-reform subsidies or did not receive any subsidies). Due to household fixed effects, the specification allows the levels of purchases of the households receiving post-reform subsidies to differ systematically from those in the control group, although figures B7–B8 show that the levels are rather similar for most products for households receiving pre- and post-reform vouchers.

A possible concern with the identification strategy is that the selection in and out of the WIC program could be correlated with the demand for healthy food products. The largest incentive to enroll in the WIC program is provided by infant formula, as infant formula is the most expensive item in the WIC food vouchers, resulting in the retail price of food vouchers for infants to be $123 per month compared to only $39 for older children (Kline et al., 2018). This is also reflected in the WIC enrollment rates: enrollment among the eligible population is highest among infants (Johnson et al., 2014). Therefore, when enrolling in the program,

[20]Because the data provide only yearly information about age, I generate two interaction terms: one with the last two quarters of the calendar year when the child becomes one year old, and the other with the first two quarters of the calendar year when the child becomes two years old.

the selection based on food products is unlikely. However, we might be more worried that participants dropped out of the program because they did not like the healthy products in the new vouchers. To address the concern, after presenting the main estimates, I analyze changes in purchases when households become ineligible for the vouchers, and the conclusions remain the same.

To provide additional support for the identifying assumption, I use summary statistics to examine the differential trends before treatment (online appendix C.2). There are slight differences in changes between treatment and control groups for healthy bread purchases and milk fat percentage from two to one years before treatment. However, the magnitudes of the differential changes before treatment are much smaller than the change in the treatment group at the time of treatment (from a year before treatment to the year of treatment). Considering the differences in the magnitude of the effects, it is unlikely that the small differential trends before treatment considerably bias the estimates. While for the quantity of bread and cereals, the changes before treatment for treatment and control groups are rather similar, there are slight differential trends also for the quantity of milk, fruits and vegetables, juice, and eggs. I will re-examine the concern about time-varying factors specific to post-reform WIC households after presenting the main estimates: allowing for separate trends and testing the parallel trends assumption when calculating the heterogeneity-robust difference-in-differences estimators.

3.2 Results: Impact of healthy food subsidies

Table 3 presents the main results of the short- and long-term impacts of healthy food subsidies introduced by the policy reform. The estimate of β_1 measures the short-term effect (while receiving the subsidies), and β_2 measures the long-term effect (1–2 years after receiving the subsidies).

Table 3 shows that the change in the WIC program in the short term increases the healthy

Table 3: The impact of healthy food subsidies introduced by the policy reform

	Bread healthy %	Milk fat %	Bread log. quantity	Milk log. quantity	Fruits&veg. log. quantity	Juice log. quantity	Eggs log. quantity	Cereals log. quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
β_1: Reformed WIC	6.922***	-0.135***	-0.013	0.022	-0.002	0.275	-0.064	0.036
	(2.213)	(0.050)	(0.053)	(0.054)	(0.087)	(0.206)	(0.095)	(0.085)
β_2: After Reformed WIC Years 1-2	1.380	-0.063	0.002	0.004	-0.044	0.158	-0.116	0.189
	(1.944)	(0.066)	(0.070)	(0.064)	(0.112)	(0.269)	(0.116)	(0.118)
α_1: WIC	-1.799	0.003	0.040	0.051	0.037	0.457***	0.073	0.132*
	(1.333)	(0.036)	(0.041)	(0.046)	(0.066)	(0.166)	(0.079)	(0.069)
α_2: After WIC Years 1-2	1.696	-0.009	0.008	0.031	0.037	0.221	0.021	-0.113
	(1.483)	(0.059)	(0.057)	(0.052)	(0.093)	(0.235)	(0.098)	(0.102)
σ: Post-Reform	0.207	-0.014**	-0.001	0.008	0.002	0.111***	0.021	-0.032*
	(0.270)	(0.007)	(0.009)	(0.008)	(0.018)	(0.037)	(0.015)	(0.017)
Year-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald test, $\beta_1 = \beta_2$, p-value	0.006	0.256	0.822	0.784	0.713	0.660	0.664	0.191
WIC households	301	301	347	347	347	347	347	347
Households	85421	80721	96637	96637	96637	96637	96637	96637
Household-quarters	1244237	1192570	1602252	1602252	1602252	1602252	1602252	1602252

Note: Each column presents estimates of a separate panel data fixed effects regression (equation (1)). A unit of observation is a household-quarter pair. Bread healthy (whole wheat) % and milk fat % are in the range of 0–100; quantity is measured in pounds for bread, gallons for milk, kilocalories for fruits and vegetables and juice, count for eggs, and ounces for cereals. All regressions include household size, logarithm of income, dummies for the year before the child is born, and children aged $0, 1, \ldots, 5$, 6–12, and 13–17. In column 2, regressions include interaction terms for WIC (and Reformed WIC) and the last two quarters of the calendar year when the child turns one year old, and the other with the first two quarters of the calendar year when the child turns two years old. Regressions include household and time period fixed effects. The sample includes households in columns 1–3 in table 1. In columns 1 and 2, the sample includes only households for whom the quarterly variable is defined (who purchase bread or milk, respectively) at least once per year and in total at least two quarters. In columns 3–8, the sample includes all households that buy any food products with UPCs in a given year. Standard errors (in parentheses) are clustered at the household level. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

(whole wheat) bread purchases (as a share from all bread purchases) and decreases milk fat percentage. For bread, the magnitude of the seven-percentage-point increase is large, corresponding to an increase of about 75% compared to healthy bread share of nine percent in the first year in the sample (table 1). For milk, the magnitude of the change is much smaller, corresponding to a change of only about seven percent. For bread, the effect decreases already in the first two years after the end of receiving the subsidies. For milk, the estimates are imprecise and don't allow ruling out large long-term effects. Note that healthy products are typically more expensive than unhealthy alternatives, making healthy products less affordable for low-income households. But in the case of milk, in some regions, it is the opposite; low-fat milk is cheaper than whole milk. This makes it more affordable to continue consuming the healthy product when the program ends. I will explore the role of prices in later parts of the paper.

The reform of the WIC program has no statistically significant short- or long-term effects on the total quantities purchased (columns 3–8). But for several products, the estimates are imprecise and don't allow rejection of sizable positive or negative effects. While the reform changed the quantities in food vouchers, this doesn't necessarily have to affect purchases. Economic theory suggests that vouchers increase consumption (beyond income effect) only if one consumes less without vouchers. But table 1 shows that before receiving the vouchers, for several products, including fruits and vegetables, the typical consumption is larger than what is provided by vouchers (described in table A1).[21]

The estimates regarding the short-term impact of the healthy food subsidies are largely consistent with the previous literature. The literature has found that the change in the food vouchers increased whole wheat bread consumption and lowered milk fat percentage while receiving the vouchers (Schultz et al., 2015; NASEM, 2017). Previous results regarding fruits and vegetables have been mixed (Zhang et al., 2020).

Robustness. Below, I summarize the analysis meant to explore the sensitivity of the results; more details are presented in online appendix C.3. The results are robust to restricting the sample only to the WIC households. Also, results remain similar for the healthy bread share and milk fat percentage when allowing separate trends for the treatment group. The key results remain the same when using the heterogeneity-robust difference-in-differences

[21]To alleviate the concern that the WIC fruits and vegetables voucher is too small compared to the recommended fruits and vegetables intakes, in 2021, the size of the voucher was more than tripled. Source: WIC Policy Memorandum, 2021-3, FNS-GD-2021-0029.

estimates proposed by de Chaisemartin and D'Haultfœuille (2020a,b) to correct for the possible bias of the two-way fixed effects estimator when treatment effects are heterogeneous. The conclusions regarding the healthy bread share and milk fat percentage remain unchanged when using the propensity score estimates. To address the concern that households choose to drop out of the new WIC program because they don't like the healthy foods, I analyze changes in purchases when households become ineligible for the vouchers when the child becomes five years old. I find that households who received the post-reform vouchers when their youngest child was four, in the following years (with the child aged five and six), changed their behavior in similar ways as those who stopped receiving the vouchers in table 3.

To account for the possible unobservable changes when households receive WIC vouchers, in this paper, the analysis compares households participating in the old and new WIC programs. As a result, the estimates measure the incremental effect of the new WIC program relative to the old. If, instead, we were to exclude from the sample the households in the old WIC program, the estimates would be slightly different, as these would measure the combined effect of the new vouchers and everything else that changes when households receive the vouchers (online appendix C.4).

Prices and the program's impact. Khan et al. (2016) show that there are exogenous regional differences in milk pricing: states where milk prices increase along with fat percentage and others where milk prices are flat. I use these regional differences to analyze the heterogeneity in the program's impact. Specifically, I calculate state-level median milk price differences of different fat percentages (whole milk, 2%, 1%, and skim milk) across stores and time periods using data on private label one-gallon milk.[22] Based on these price differences, I categorize states as those with flat prices where the price of whole milk is not higher than others, those where whole milk is more expensive than others, and those with mixed price

[22] Private label accounts for about 80% of the milk sold, and 1 gallon is the most popular package size.

rankings. Figure C6 in the online appendix presents a map describing the regional differences in prices. Table 4 presents estimates of the short- and long-term impacts of the post-reform WIC subsidies estimated separately in states where prices are flat and in those where prices increase along with the fat percentage.[23] In states where prices are flat (healthy product is not cheaper), the long-term effect is not as large as the short-term effect. The Wald test in column 1 shows that with 95% probability, the estimates allow rejection of the long-term impact (β_2) of the same size as the short-term effect (β_1).

Table 4: The short- and long-term impact and prices. Dependent variable: milk fat percentage

	Prices flat (1)	Prices increasing (2)
β_1: Reformed WIC	-0.262**	-0.281***
	(0.105)	(0.092)
β_2: After Reformed WIC Years 1-2	-0.009	-0.301
	(0.111)	(0.198)
α_1: WIC	0.058	0.151**
	(0.071)	(0.060)
α_2: After WIC Years 1-2	-0.017	0.193
	(0.098)	(0.183)
σ: Post-Reform	-0.013**	-0.010**
	(0.005)	(0.005)
Year FE	Yes	Yes
Quarter FE	Yes	Yes
Household FE	Yes	Yes
Wald test, $\beta_1 = \beta_2$, p-value	0.031	0.906
WIC households	94	60
Households	27769	18707
Household-quarters	404521	278787

Note: The table presents results from re-estimating the regression in column 2 of table 3. In column 1, the sample is restricted to households in states where milk prices are flat (whole milk is not more expensive than other types of milk). In column 2, the sample is restricted to households in states where milk prices increase along with fat percentage (whole milk is more expensive than other types of milk). Standard errors (in parentheses) are clustered at the household level. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

Externalities to other product categories. While the WIC vouchers include only a small set of products, there could be nutritional externalities to other purchases. The 2009

[23]Similar estimates using only the WIC sample are in the online appendix (table C10).

changes in WIC vouchers aimed to increase fiber consumption in bread and reduce saturated fat consumption in milk. Consumers could compensate in other product categories, for example, consume more saturated fat in other dairy products. Alternatively, there could be positive externalities if consumers learn from the WIC program and increase fiber and reduce saturated fat consumption in other product categories. Online appendix C.7 analyzes the externalities. The results show that the reform decreased saturated fat consumption in milk and increased fiber consumption in bread. For non-WIC product categories, the point estimates are small, and although the estimates are imprecise, they allow ruling out large externalities.

Discussion and possible limitations. The analysis might raise the question of whether the impact of the new WIC program is driven by the income effect. The answer is no. While indeed the WIC vouchers increase income, which might lead the households to switch to healthier foods, but the change in WIC vouchers kept the value of the vouchers the same. Therefore, in the above regressions, controlling for receiving WIC vouchers (α_1 and α_2 in equation 1) captures the WIC-related effects, including the income effect. Therefore, the parameters of interest (β_1 and β_2 in equation 1) measure only the incremental effect of the new WIC program, identifying the impact of the food vouchers being restricted to healthy foods, separating it from the overall impact of the vouchers (including the income effect).

The above analysis focuses on measuring the impact of the new WIC vouchers; it does not address the mechanism of why purchases changed in some categories while receiving the new WIC vouchers. It does not measure the importance of the substitution effect (change in relative prices) and the information effect (learning about recommended foods), and behavioral heuristics such as mental budgeting (Hastings and Shapiro, 2018). The limited persistence in the case of healthy bread is more consistent with substitution effect and mental budgeting than with learning.

Finally, as is always the case with scanner data, it includes information only on purchases, not consumption or food waste. If households were more likely to discard food that they received for free, then the short-run impact of the program would be overestimated.

4 The role of prices

In this section, I estimate a demand model to examine the long-term impact of healthy food subsidies on preferences and separate it from the role of prices. I focus only on bread for two reasons. First, bread was a product category with a high short-term impact of the subsidy. Second, it is a typical product category in terms of the pricing of healthy and unhealthy products—the healthy version is more expensive than the unhealthy alternative (as opposed to milk). In this way, it provides a good example of typical choices between healthy and unhealthy foods. I use the demand model estimates for counterfactual analysis, which answers the question of how long does it take for the purchases to return to the pre-WIC level with different prices.

4.1 Demand model

I model demand using a standard discrete choice multinomial logit model with random coefficients. Households choose between four bread products: white, whole grain, whole wheat, and other. The last option, other, is the composite of all other bread products. Household i, in shopping trip t, chooses a product with the highest indirect utility, where the indirect utility from product $j \in \{White, WG, WW, Other\}$ equals:

$$u_{ijt} = \alpha_i p_{jt} + \gamma_i \cdot \mathbf{1}[s_{i,t-1} = j] + (\beta \kappa_{WW,it} + \theta_i) \cdot WW_j + \delta_i WG_j + \lambda_i White_j + \varepsilon_{ijt} \qquad (2)$$

where p_{jt} is price and coefficient α_i measures the household's price sensitivity. Variable $s_{i,t-1} \in \{White, WG, WW, Other\}$ indicates the product purchased on the previous trip. The corresponding coefficient (γ_i) measures the consumer's short-term state dependence. The indicator variables $White_j$, WG_j, and WW_j indicate whether the alternative j is white bread, whole grain, or whole wheat, respectively. The corresponding coefficients $(\theta_i, \delta_i, \lambda_i)$ measure time-invariant tastes for those products.[24]

I model consumer heterogeneity with a multivariate normal distribution across price, tastes for products, and short-term state dependence coefficients $(\alpha_i, \gamma_i, \theta_i, \delta_i, \lambda_i)$. I estimate the parameters of the distribution of the random coefficients, specifically, means $(\bar{\alpha}, \bar{\gamma}, \bar{\theta}, \bar{\delta}, \bar{\lambda})$, standard deviations $(\sigma_p, \sigma_s, \sigma_{WW}, \sigma_{WG}, \sigma_{White})$, and the corresponding covariances $\boldsymbol{\rho}$, allowing correlation in preferences across these attributes.

The variable $\kappa_{WW,it}$ equals the last year's (last 52 weeks') share of whole wheat bread in bread purchases. The corresponding coefficient β measures long-term taste inertia. One could think of $\kappa_{WW,it}$ as a stock of taste for whole wheat bread, analogous to the taste stock defined in Atkin (2013). The inclusion of long-term state dependence in the model is motivated by nutrition literature, which provides evidence that repeated exposure is required to change nutritional preferences (Sullivan and Birch, 1990; Birch, 1999; Maier et al., 2007; Lakkakula et al., 2010; Anzman-Frasca et al., 2012; Hausner et al., 2012).

Finally, the taste shock ε_{ijt} is assumed to be independent across households, products, and shopping trips, distributed according to a type 1 extreme value distribution. Let's denote the first five terms on the right-hand side of equation (2) by $v_{ijt} = \alpha_i p_{jt} + \gamma_i \cdot \mathbf{1}[s_{i,t-1} = j] + (\beta \kappa_{WW,it} + \theta_i) \cdot WW_j + \delta_i WG_j + \lambda_i White_j$. Then, the probability that household i buys product j on trip t equals:

$$Prob_{ijt} = \frac{\exp(v_{ijt})}{\sum_{k=1}^{J} \exp(v_{ikt})} \tag{3}$$

[24] I don't include advertising because advertising typically varies by brand and not by product characteristics (by the type of bread).

Identification. While the formal proof of identification is beyond the scope of this paper, let us discuss the intuition of how the available data variation identifies the parameters of the model.

Let us first look at the data variation that allows us to pin down the long-term state dependence parameter β (assumed to be the same across households), separating it from the household-specific constant-over-time preference for whole wheat bread θ_i. We observe every household making purchases both before and after receiving the WIC vouchers. The post-reform vouchers generated an exogeneous jump in the state variable: the last year's share of whole wheat bread $\kappa_{WW,it}$. The within-household change in the state variable from before to after receiving the vouchers allows us to pin down the long-term state dependence parameter because, both before and after vouchers, a household is characterized by the same time-invariant taste for whole wheat θ_i.[25] Figure 1 describes this jump in the state variable. It presents the within-household comparison of whole wheat bread share before versus after receiving the vouchers; it is shown separately for households receiving pre- versus post-reform WIC vouchers. The first grey bar is the whole wheat share's average across post-reform WIC households before receiving the vouchers. The first blue bar is the average across the same households after receiving the vouchers. The comparison of the first grey and blue bars describes the within-household increase (statistically significant at the 1% level) in whole wheat bread share due to the post-reform WIC vouchers.

Next, let us look at the data variation that helps pin down the household-specific short-term state dependence γ_i. The setting allows to identify short-term state dependence because bread is often bought, and even more importantly, households often switch between different types of bread. In the sample, the median household buys bread 100 times, and all households switch at least once. As noted by Simonov et al. (2020), these switches help to identify the short-term state dependence because they allow observing each household in both states:

[25] Using a major change in conditions to identify inertia is analogous to Handel (2013).

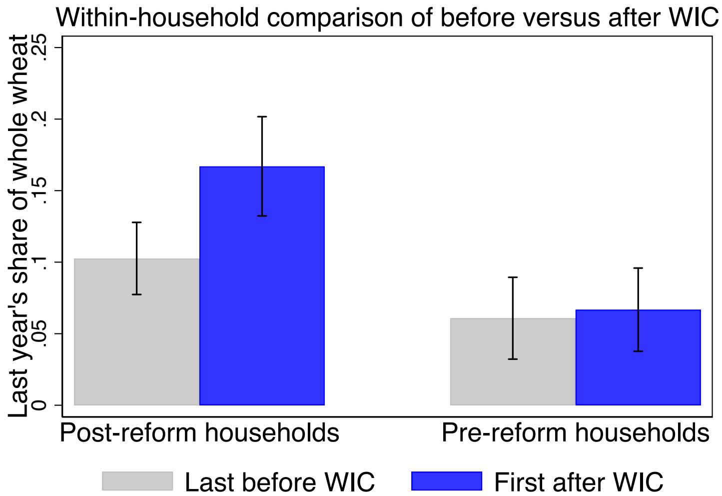


Figure 1: Within-household comparison of the last year's share of whole wheat bread: before WIC versus after WIC; presented separately for pre- versus post-reform WIC households

Note: Each bar presents the average across households, and the capped spikes are the 95% confidence intervals. The unit of observation is a household. The first two bars are each calculated as the average across 178 households receiving post-reform WIC vouchers. The last two bars each average over 74 households receiving pre-reform vouchers. Grey bars are averages during the previous trip before receiving the vouchers, and blue bars during the first trip after receiving the vouchers. According to the t-test, for post-reform households, the means before versus after vouchers differ at the 1% significance level.

having bought the product or not on the previous trip. For each household, we can compare the probability of purchasing the product in two different states: one where they bought the product the last shopping trip and the other where they didn't. Importantly, the probabilities are calculated separately for each household. Figure D1 in the online appendix presents the average probabilities and 95% confidence intervals by state (whether bought or not on the previous trip) and product. The figure shows that the average probabilities in the two states are statistically significantly different from each other, suggesting that, on average, there is short-term state dependence. We can also calculate the difference between these conditional probabilities for each household and product pair. Again, as the probabilities are calculated separately for each household, the difference in probabilities is also household-specific. By and large, the difference in probabilities in the two states captures the household's state dependence. A household is more state dependent (the difference close to one) if it is very likely to buy the product when it bought the last trip and not at all likely to buy when it didn't buy the last trip. The histogram of the differences between the probabilities (figure D2 in the online appendix) shows that households are heterogeneous in state dependence. A few households are variety-seeking (negative difference), but the majority of households are more likely to buy the same product as last time (positive difference).

Finally, as Simonov et al. (2020) draw attention to, when analyzing state dependence, the literature (including Keane (1997); Seetharaman et al. (1999); Shum (2004); Dubé et al. (2010)) has typically relied on assumptions that ignore the initial conditions problem—the issue that the initial observed state is endogeneous. In the main specification of the model, the current paper follows the literature and assumes that the initial state is exogeneous. In the current setting, it is likely that the possible bias from the initial conditions problem is small because the panel is very long, the median household makes exactly 100 purchases, and all households switch between products. To formally address the concern about the initial conditions bias, after estimating the demand model, I will use the method proposed by Simonov et al. (2020) to bound the true short-term state dependence value.

Below, I will discuss the identification of the price coefficient.

Control function. To address the concern of price endogeneity, I also estimate the model using a control function. Following Petrin and Train (2010), I assume the pricing function takes the following form where the price of product j in week[26] t in market m in (aggregate) retailer r equals

$$p_{jtmr} = \mathbf{z'}_{jtmr}\boldsymbol{\eta} + \xi_{jtmr} \tag{4}$$

where $\mathbf{z}_{jtmr}$ is the vector of exogenous instruments and ξ_{jtmr} is the unobserved price shock. The pricing function is estimated at the week, market, and (aggregate) retailer level because

[26]With slight abuse of notation, here week, instead of a shopping trip, is indexed by t.

price variables are constructed so that these vary only by market and (aggregate) retailer, not by the consumer. Typically, when estimating a demand model, we are concerned that prices might depend on the unobserved factors that directly affect demand. In the control function approach, the identifying assumption is that prices and taste shocks, ε, are independent conditional on the unobserved factors, ξ, affecting prices.

For the instrument, I use the average price of the same (aggregate) retailer in other markets (following Hausman (1996); Nevo (2001); DellaVigna and Gentzkow (2019)). I also include product fixed effects and market fixed effects. The instrument has a strong first-stage (table D1 in the online appendix).

Estimation. The demand model is estimated using simulated maximum likelihood. I estimate the model with the control function in two steps. First, I estimate the reduced form pricing regression (equation 4) with ordinary least squares to recover the residuals ξ_{jtm}. Then, I include these residuals as an additional regressor (control function) in the indirect utility (equation 2) and estimate the demand model using simulated maximum likelihood. In the model with the control function, standard errors are calculated using bootstrap (over two steps of the estimation, estimating first the control function and then the demand model) with 100 bootstrap samples.

The demand model is estimated using purchases before and after the household received the WIC vouchers. The reason why I don't use time periods when households receive WIC vouchers is that I don't have information about whether, on a given shopping trip, they are paying with the voucher or cash. Therefore, I wouldn't know the price they faced on a given shopping trip (zero because of the voucher, or the regular price). In the household's first year in the sample, I exclude the first 10 shopping trips to initialize the long-term state dependence variable $\kappa_{WW,it}$ and use the latest excluded shopping trip to initialize the short-term state dependence variable $s_{i,t-1}$.

4.2 Estimation results

The estimates of the demand model are in table 5. Specifically, it includes the estimates of the means of the random coefficients and estimates of fixed coefficients; estimates of standard deviations and covariances of the random coefficients are in table D2 in the online appendix. Columns 1 and 2 present estimates without and with a control function, respectively. As expected, the estimated coefficient on price is larger in absolute value with the control function. The estimates show the importance of both the short-term state dependence and last year's share of whole wheat bread affecting current purchases. Online appendix D.3 shows that the key results are robust to alternative constructions of prices. Using the method proposed by Simonov et al. (2020), it also shows that the results are robust to alternative initial conditions for state dependence.

Table 5: Bread demand model estimates

	No control function		Control function	
	(1)		(2)	
	Estim.	SE	Estim.	SE
Mean				
$\bar{\alpha}$: Price	-1.029***	(0.088)	-1.316***	(0.287)
$\bar{\gamma}$: Short-term state dependence	0.693***	(0.025)	0.614***	(0.067)
β: Last year's share of whole wheat	3.872***	(0.121)	3.784***	(0.254)
$\bar{\theta}$: Whole wheat bread	-1.534***	(0.051)	-1.549***	(0.147)
$\bar{\delta}$: Whole grain bread	-0.271***	(0.059)	-0.606***	(0.195)
$\bar{\lambda}$: White bread	-0.403***	(0.077)	-0.809**	(0.362)
Residual			0.539*	(0.277)
Log-likelihood	-22843		-22821	
Number of choices	112168		112168	
Number of households	252		252	

Note: The table presents estimates of two random coefficient logit models. The table presents the estimates of the means of the random coefficients and estimates of fixed coefficients; estimates of standard deviations and covariances of the random coefficients are in table D2 in the online appendix. For each model, the first column presents parameter estimates and the second column standard errors. The base type of bread is *other* bread. The sample includes households from columns 1–2 in table 1 with a sufficient number of bread purchases; for each household the sample excludes the years when the household received WIC vouchers. Standard errors are clustered at the household level. In column 2, standard errors are obtained by bootstrap (over 2 steps of the estimation: estimating first the control function and then the demand model) with 100 bootstrap samples. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

Counterfactual analysis. I run a counterfactual analysis to study how long it takes for purchases to return to the pre-WIC level with alternative prices. Specifically, the counterfactual policy is a 10-cent subsidy on whole wheat bread; the subsidy is given to the households after their eligibility for the current WIC program runs out when the child turns five years old.

This subsidy is easily implementable via coupons that the program could give households after the current WIC policy ends when the child turns five years old. The WIC program already works with retailers to ensure that stores accept the WIC vouchers as a payment method. The contracts with retailers could be extended so that stores would also accept the WIC coupons for a discount. Nowadays, WIC vouchers are electronic, and so could be the coupons, which makes the coupons easy to use and limits the paperwork.

The counterfactual analysis uses estimates from column 2 in table 5 and average prices (column 1 in table 2).[27] The 10-cent subsidy decreases the price gap with the cheapest (white) bread by less than 20%, but it has a much larger impact on price gaps with more similarly priced products. The counterfactual analysis is done in the following steps. In step 1, 1000 synthetic households are simulated based on the estimated distribution of preferences. To focus on the role of preferences, I assume that the households are homogeneous in the shopping frequency and the length of time spent in the WIC program. For each household, given the household's preferences and prices, the pre-WIC equilibrium purchase probabilities and state variables κ_{WW} and s_{t-1} are simulated using 1000 iterations. In step 2, to approximate receiving the WIC vouchers, I assume that for all households for 36 purchase occasions, the price of whole wheat bread equals zero on every third shopping occasion. (This reflects that, typically, households buy much more bread than is covered by the voucher.) From step 2, for each household, we obtain the predicted purchase probabilities, κ_{WW}, and s_{t-1}, when

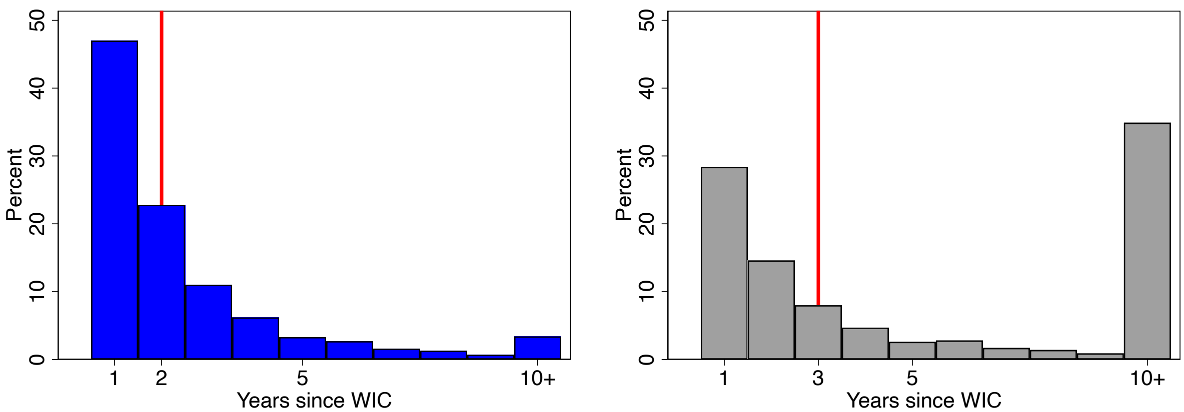
[27]The counterfactuals are calculated under the assumption that the estimated control function on prices does not change.

the household exits the WIC program (after 36 shopping occasions). In step 3, post-WIC purchase histories are predicted. I assume all households make 18 shopping trips with bread purchases per year.[28] I predict purchase histories for 20 years (360 purchase occasions) using either the same average prices or the average prices with a 10-cent whole wheat bread price discount. Then, for each household, it is calculated how many years it takes to return to the equilibrium whole wheat bread purchase probability calculated in step 1. Online appendix D.4 presents more details about the counterfactuals. It shows that, as expected, the purchase probability of whole wheat bread is an increasing function of the state variable κ_{WW}. Receiving the vouchers shifts the distribution upwards, increasing both the state variable and the purchase probability. It also presents an example of a counterfactual household and describes counterfactual households' heterogeneity.

The results of the counterfactual analysis are in figure 2. With the current prices (figure 2a), 70% of the households return to pre-WIC consumption levels in the first two years after exiting the program. If the price of whole wheat bread were 10 cents lower (figure 2b), the return to pre-WIC levels would be considerably slower: only 43% return to pre-WIC levels in the first two years. The median returning time increases by one year, from two years in the baseline to three years with the discount.

What do we learn from the counterfactual analysis? The counterfactual analysis shows that if we would like the program to have a long-term effect beyond the period during which vouchers are received, then that might be possible via a rather modest subsidy. While the paper studies a specific program and a specific product, many settings have similar concerns. For example, aid programs in developing countries provide free essential health products, such as insecticide-treated bednets, deworming drugs, and home water purification solutions, which households should continue to obtain and use indefinitely. But the programs face budget constraints that limit the ability to support everyone indefinitely. Combining temporarily free

[28] The median household in the first year post-WIC makes 18 shopping trips with bread purchases.



(a) Base (b) 10-cent whole wheat bread price discount

Figure 2: Number of years it takes to for purchases to return to pre-WIC levels

Note: Each figure presents the histogram of the number of years it takes to return to pre-WIC levels, based on counterfactual analysis of 1000 randomly drawn households. The red lines show the median number of years it takes to return to pre-WIC levels, which is two years in the baseline (figure 2a) and three years in the counterfactual (figure 2b). The counterfactuals are obtained using the estimates from column 2 in table 5 and average prices (over all households and time periods). In step 1, 1000 households are drawn from the estimated distribution, and equilibrium purchase shares (and κ_{WW} and s_{t-1}) for each household are simulated given the prices. In step 2, to approximate receiving the WIC vouchers for two years, it is assumed that for 36 purchase occasions, on every third shopping occasion, the price of whole wheat bread equals zero. This gives for each household the predicted purchase shares, κ_{WW}, and s_{t-1}, when the household exits the WIC program. In step 3, subsequent purchase histories are predicted using either the same average prices (figure 2a) or the average prices with a 10-cent whole wheat bread price discount (figure 2b). Then, for each household, it is calculated how many years it takes to return to the equilibrium calculated in step 1.

products with long-term subsidies might be a more sustainable way to lengthen the impact of the programs.

5 Conclusion

In this paper, I study the short- and long-term impacts of healthy food subsidies. I find that the subsidies increase the purchases of subsidized products in the case of whole wheat bread and low-fat milk. But for bread, the effect decreases already in the first two years after the end of receiving the subsidies. For milk, the estimates are imprecise and don't allow ruling out that the short- and long-term effects are of similar size. Demand model estimates show that the program has some long-term impact on preferences. The estimates imply that price differences between healthy and unhealthy products play a role in the decrease in the program's long-term impact.

References

Allcott, H., R. Diamond, J.-P. Dubé, J. Handbury, I. Rahkovsky, and M. Schnell (2019a): "Food Deserts and the Causes of Nutritional Inequality," *The Quarterly Journal of Economics*, 134, 1793–1844.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019b): "Regressive Sin Taxes, with an Application to the Optimal Soda Tax*," *The Quarterly Journal of Economics*, 134, 1557–1626.

Anderson, M. L. and D. A. Matsa (2011): "Are Restaurants Really Supersizing America?" *American Economic Journal: Applied Economics*, 3, 152–188.

Anzman-Frasca, S., J. S. Savage, M. E. Marini, J. O. Fisher, and L. L. Birch (2012): "Repeated exposure and associative conditioning promote preschool children's liking of vegetables," *Appetite*, 58, 543–553.

Atkin, D. (2013): "Trade, Tastes, and Nutrition in India," *American Economic Review*, 103, 1629–1663.

——— (2016): "The Caloric Costs of Culture: Evidence from Indian Migrants," *American Economic Review*, 106, 1144–1181.

Belot, M., N. Berlin, J. James, and V. Skafida (2018): "The formation and malleability of dietary habits: A field experiment with low income families," Tech. Rep. 11317, Institute for the Study of Labor (IZA).

Belot, M., J. James, and P. Nolen (2016): "Incentives and children's dietary choices: A field experiment in primary schools," *Journal of Health Economics*, 50, 213–229.

Birch, L. L. (1999): "Development of food preferences," *Annual Review of Nutrition*, 19, 41–62.

Bollinger, B., P. Leslie, and A. Sorensen (2011): "Calorie Posting in Chain Restaurants," *American Economic Journal: Economic Policy*, 3, 91–128.

Bollinger, B. and S. Sexton (2018): "Local Excise Taxes, Sticky Prices, and Spillovers: Evidence from Berkeley's Soda Tax," SSRN Scholarly Paper ID 3087966, Social Science Research Network, Rochester, NY.

Bronnenberg, B. J., J.-P. H. Dube, and M. Gentzkow (2012): "The Evolution of Brand Preferences: Evidence from Consumer Migration," *American Economic Review*, 102, 2472–2508.

Bronnenberg, B. J., J.-P. Dubé, and R. E. Sanders (2020): "Consumer Misinformation and the Brand Premium: A Private Label Blind Taste Test," *Marketing Science*, 39, 382–406.

Cawley, J. (2015): "An Economy of Scales: A Selective Review of Obesity's Economic Causes, Consequences, and Solutions," *Journal of Health Economics*, 43, 244–268.

Cawley, J., D. Frisvold, A. Hill, and D. Jones (2019): "The impact of the Philadelphia beverage tax on purchases and consumption by adults and children," *Journal of Health Economics*, 67, 102225.

Charness, G. and U. Gneezy (2009): "Incentives to Exercise," *Econometrica*, 77, 909–931.

CURRIE, J., S. DELLAVIGNA, E. MORETTI, AND V. PATHANIA (2010): "The Effect of Fast Food Restaurants on Obesity and Weight Gain," *American Economic Journal: Economic Policy*, 2, 32–63.

DE CHAISEMARTIN, C. AND X. D'HAULTFŒUILLE (2020a): "Difference-in-Differences Estimators of Intertemporal Treatment Effects," SSRN Scholarly Paper ID 3731856, Social Science Research Network, Rochester, NY.

——— (2020b): "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110, 2964–2996.

DELLAVIGNA, S. AND M. GENTZKOW (2019): "Uniform Pricing in U.S. Retail Chains," *The Quarterly Journal of Economics*, 134, 2011–2084.

DHAR, T. AND K. BAYLIS (2011): "Fast-Food Consumption and the Ban on Advertising Targeting Children: The Quebec Experience," *Journal of Marketing Research*, 48, 799–813.

DUBOIS, P., R. GRIFFITH, AND A. NEVO (2014): "Do Prices and Attributes Explain International Differences in Food Purchases?" *American Economic Review*, 104, 832–867.

DUBOIS, P., R. GRIFFITH, AND M. O'CONNELL (2020): "How Well Targeted Are Soda Taxes?" *American Economic Review*, 110, 3661–3704.

DUBOIS, P., R. GRIFFITH, AND M. O'CONNELL (2018): "The Effects of Banning Advertising in Junk Food Markets," *The Review of Economic Studies*, 85, 396–436.

DUBÉ, J.-P., G. J. HITSCH, AND P. E. ROSSI (2010): "State Dependence and Alternative Explanations for Consumer Inertia," *The RAND Journal of Economics*, 41, 417–445.

EINAV, L., E. LEIBTAG, AND A. NEVO (2010): "Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?" *Quantitative Marketing and Economics*, 8, 207–239.

Eizenberg, A. and A. Salvo (2015): "The Rise of Fringe Competitors in the Wake of an Emerging Middle Class: An Empirical Analysis," *American Economic Journal: Applied Economics*, 7, 85–122.

Figlio, D., S. Hamersma, and J. Roth (2009): "Does Prenatal WIC Participation Improve Birth Outcomes? New Evidence From Florida," *Journal of Public Economics*, 93, 235–245.

Fletcher, J. M., D. E. Frisvold, and N. Tefft (2010): "The effects of soft drink taxes on child and adolescent consumption and weight outcomes," *Journal of Public Economics*, 94, 967–974.

Frisvold, D., E. Leslie, and J. P. Price (2020): "Do Targeted Vouchers Instill Habits? Evidence from Women, Infants, and Children," *Contemporary Economic Policy*, 38, 67–80.

Gidding, S. S., B. A. Dennison, L. L. Birch, S. R. Daniels, M. W. Gilman, A. H. Lichtenstein, K. T. Rattay, J. Steinberger, N. Stettler, and L. V. Horn (2006): "Dietary Recommendations for Children and Adolescents: A Guide for Practitioners," *Pediatrics*, 117, 544–559.

Giné, X., D. Karlan, and J. Zinman (2010): "Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation," *American Economic Journal: Applied Economics*, 2, 213–235.

Griffith, R., S. von Hinke, and S. Smith (2018): "Getting a Healthy Start: The Effectiveness of Targeted Benefits for Improving Dietary Choices," *Journal of Health Economics*, 58, 176–187.

Handel, B. R. (2013): "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *American Economic Review*, 103, 2643–2682.

Harding, M. and M. Lovenheim (2017): "The effect of prices on nutrition: Comparing the impact of product- and nutrient-specific taxes," *Journal of Health Economics*, 53, 53–71.

Hastings, J. and J. M. Shapiro (2018): "How Are SNAP Benefits Spent? Evidence from a Retail Panel," *American Economic Review*, 108, 3493–3540.

Hausman, J. A. (1996): "Valuation of New Goods under Perfect and Imperfect Competition," in *The Economics of New Goods*, ed. by T. F. Bresnahan and R. J. Gordon, University of Chicago Press, 207–248.

Hausner, H., D. L. Hartvig, H. C. Reinbach, K. Wendin, and W. L. P. Bredie (2012): "Effects of repeated exposure on acceptance of initially disliked and liked Nordic snack bars in 9–11 year-old children," *Clinical Nutrition*, 31, 137–143.

Hinnosaar, M. and E. Liu (2021): "Malleability of Alcohol Consumption: Evidence from Migrants," SSRN Scholarly Paper ID 3569603.

Hoynes, H., M. Page, and A. H. Stevens (2011): "Can Targeted Transfers Improve Birth Outcomes?: Evidence From the Introduction of the WIC Program," *Journal of Public Economics*, 95, 813–827.

Hoynes, H. W. and D. W. Schanzenbach (2015): "U.S. Food and Nutrition Programs," Working Paper 21057, National Bureau of Economic Research.

Institute of Medicine (2006): *WIC Food Packages: Time for a Change*, Washington, D.C.: The National Academies Press.

Johnson, P., L. Giannarelli, E. Huber, and D. Betson (2014): "National and State-Level Estimates of Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) Eligibles and Program Reach, 2011," Report WIC-14-ELIG, U.S. Department of Agriculture, Food and Nutrition Service, Office of Research and Analysis.

Keane, M. P. (1997): "Modeling Heterogeneity and State Dependence in Consumer Choice Behavior," *Journal of Business & Economic Statistics*, 15, 310–327.

Khan, R., K. Misra, and V. Singh (2016): "Will a Fat Tax Work?" *Marketing Science*, 35, 10–26.

Khan, R., T. Zhu, and S. Dhar (2018): "The effect of the WIC program on consumption patterns in the cereal category," *Quantitative Marketing and Economics*, 16, 79–109.

Kline, N., C. Warner-Griffin, E. Wilcox-Cook, and B. Thorn (2018): "Fiscal year 2014 WIC food package costs: Final report," Tech. rep., U.S. Department of Agriculture, Food and Nutrition Service, Office of Policy Support, Alexandria, VA.

Kueng, L. and E. Yakovlev (2021): "The Long-Run Effects of a Public Policy on Alcohol Tastes and Mortality," *American Economic Journal: Economic Policy*, 13, 294–328.

Lakkakula, A., J. Geaghan, M. Zanovec, S. Pierce, and G. Tuuri (2010): "Repeated taste exposure increases liking for vegetables by low-income elementary school children," *Appetite*, 55, 226–231.

List, J. A. and A. S. Samek (2015): "The Behavioralist as Nutritionist: Leveraging Behavioral Economics to Improve Child Food Choice and Consumption," *Journal of Health Economics*, 39, 135–146.

Loewenstein, G., J. Price, and K. Volpp (2016): "Habit Formation in Children: Evidence From Incentives for Healthy Eating," *Journal of Health Economics*, 45, 47–54.

Ludwig, D. S., W. C. Willett, J. S. Volek, and M. L. Neuhouser (2018): "Dietary Fat: From Foe to Friend?" *Science*, 362, 764–770.

Ma, Y., K. L. Ailawadi, and D. Grewal (2013): "Soda Versus Cereal and Sugar Versus Fat: Drivers of Healthful Food Intake and the Impact of Diabetes Diagnosis," *Journal of Marketing*, 77, 101–120.

Maier, A., C. Chabanet, B. Schaal, S. Issanchou, and P. Leathwood (2007): "Effects of repeated exposure on acceptance of initially disliked vegetables in 7-month old infants," *Food Quality and Preference*, 18, 1023–1032.

Meckel, K. (2020): "Is the Cure Worse Than the Disease? Unintended Effects of Payment Reform in a Quantity-Based Transfer Program," *American Economic Review*, 110, 1821–1865.

Mehta, N., S. Rajiv, and K. Srinivasan (2004): "Role of Forgetting in Memory-Based Choice Decisions: A Structural Model," *Quantitative Marketing and Economics*, 2, 107–140.

Moorman, C., R. Ferraro, and J. Huber (2012): "Unintended Nutrition Consequences: Firm Responses to the Nutrition Labeling and Education Act," *Marketing Science*, 31, 717–737.

Muller, L., A. Lacroix, J. L. Lusk, and B. Ruffieux (2017): "Distributional Impacts of Fat Taxes and Thin Subsidies," *The Economic Journal*, 127, 2066–2092.

NASEM (2017): *Review of WIC Food Packages: Improving Balance and Choice: Final Report*, Washington, D.C.: National Academies Press.

Nevo, A. (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69, 307–342.

Oliveira, V. (2017): "The Food Assistance Landscape: FY 2016 Annual Report," Economic Information Bulletin EIB-169, U.S. Department of Agriculture, Economic Research Service.

Oliveira, V. and E. Frazao (2009): "The WIC Program: Background, Trends, and Economic Issues, 2009 Edition," Economic Research Report 73, United States Department of Agriculture, Economic Research Service.

——— (2015): "The WIC Program: Background, Trends, and Economic Issues, 2015 Edition," Economic Information Bulletin 134, U.S. Department of Agriculture, Economic Research Service.

Oster, E. (2018): "Diabetes and Diet: Purchasing Behavior Change in Response to Health Information," *American Economic Journal: Applied Economics*, 10, 308–348.

Petrin, A. and K. Train (2010): "A Control Function Approach to Endogeneity in Consumer Choice Models," *Journal of Marketing Research*, 47, 3–13.

Puranam, D., V. Narayan, and V. Kadiyali (2017): "The Effect of Calorie Posting Regulation on Consumer Opinion: A Flexible Latent Dirichlet Allocation Model with Informative Priors," *Marketing Science*, 36, 726–746.

Rossin-Slater, M. (2013): "WIC in Your Neighborhood: New Evidence on the Impacts of Geographic Access to Clinics," *Journal of Public Economics*, 102, 51–69.

Royer, H., M. Stehr, and J. Sydnor (2015): "Incentives, Commitments, and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company," *American Economic Journal: Applied Economics*, 7, 51–84.

Schultz, D. J., C. Byker Shanks, and B. Houghtaling (2015): "The Impact of the 2009 Special Supplemental Nutrition Program for Women, Infants, and Children Food Package Revisions on Participants: A Systematic Review," *Journal of the Academy of Nutrition and Dietetics*, 115, 1832–1846.

Seetharaman, P., A. Ainslie, and P. K. Chintagunta (1999): "Investigating Household State Dependence Effects across Categories," *Journal of Marketing Research*, 36, 488–500.

Seiler, S., A. Tuchman, and S. Yao (2021): "The Impact of Soda Taxes: Pass-Through, Tax Avoidance, and Nutritional Effects," *Journal of Marketing Research*, 58, 22–49, publisher: SAGE Publications Inc.

Seo, B. (2019): "Firm Scope and the Value of One-Stop Shopping in Washington State's Deregulated Liquor Market," SSRN Scholarly Paper ID 2863094.

Shum, M. (2004): "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereals Market," *Journal of Economics & Management Strategy*, 13, 241–272.

Simonov, A., J.-P. Dubé, G. Hitsch, and P. Rossi (2020): "State-Dependent Demand Estimation with Initial Conditions Correction," *Journal of Marketing Research*, 57, 789–809, publisher: SAGE Publications Inc.

Sudhir, K. and I. Tewari (2015): "Long Term Effects of 'Prosperity in Youth' on Consumption: Evidence from China," SSRN Scholarly Paper ID 2681182.

Sullivan, S. A. and L. L. Birch (1990): "Pass the sugar, pass the salt: Experience dictates preference." *Developmental psychology*, 26, 546.

Thorn, B., N. Kline, C. Tadler, E. Budge, E. Wilcox-Cook, J. Michaels, M. Mendelson, K. L. Patlan, and V. Tran (2018): "WIC Participant and Program Characteristics 2016," Tech. rep., U.S. Department of Agriculture, Food and Nutrition Service.

Uetake, K. and N. Yang (2020): "Inspiration from the "Biggest Loser": Social Interactions in a Weight Loss Program," *Marketing Science*, 39, 487–499.

USDA (2018a): *USDA Branded Food Products Database*, U.S. Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory.

——— (2018b): *USDA National Nutrient Database for Standard Reference*, U.S. Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory.

Wang, D. D. and F. B. Hu (2017): "Dietary Fat and Risk of Cardiovascular Disease: Recent Controversies and Advances," *Annual Review of Nutrition*, 37, 423–446.

Wang, E. Y. (2015): "The impact of soda taxes on consumer welfare: implications of storability and taste heterogeneity," *The RAND Journal of Economics*, 46, 409–441.

Zhang, Q., M. A. Alsuliman, M. Wright, Y. Wang, and X. Cheng (2020): "Fruit and Vegetable Purchases and Consumption among WIC Participants after the 2009 WIC Food Package Revision: A Systematic Review," *Advances in Nutrition*, 11, 1646–1662.

Online Appendix: Not For Publication

A Online Appendix: WIC food package

Table A1: WIC food package, per month

	Children	Women		
	1–4 years	Pregnant, partially breast-feeding	Postpartum	Breastfeeding
Panel A: 2006-2009[1]				
Milk (gallons)[2]	6	7	6	7
Whole wheat/grain bread (lb)	-	-	-	-
Fresh fruits and vegetables ($)	-	-	-	-
Juice (fluid oz)	2.25	2.25	1.5	2.625
Eggs	24-30	24-30	24-30	24-30
Breakfast cereal (oz)	36	36	36	36
Canned fish (oz)	-	-	-	26
Carrots (lb)	-	-	-	2
Cheese (lb)	-	-	-	1
Legumes (lb)	1	1	-	1
And/or peanut butter (oz)	Or 18	Or 18	-	And 18
Panel B: 2009-2016[3]				
Milk (gallons)[4]	4	5.5	4	6
Whole wheat/grain bread (lb)[5]	2	1	-	1
Fresh fruits and vegetables ($)[6]	6 (8)	8 (10)	8 (10)	10
Juice (gallons)	1	1.125	0.75	1.125
Eggs	12	12	12	24
Breakfast cereal (oz)	36	36	36	36
Canned fish (oz)	-	-	-	30
Carrots (lb)	-	-	-	-
Cheese (lb)	-	-	-	1
Legumes (lb)	1	1	1	1
And/or peanut butter (oz)	Or 18	And 18	Or 18	And 18

Notes: 1: In 2009, the exact date of the introduction of the new food package (described in panel B) varied by state. 2: Milk is allowed to be partially substituted by cheese. 3. In 2014, federal regulation required small modifications to the food package, the implementation date (in 2014–2015) of the modifications varied by state and product. 4: For children 2-4 years of age and women reduced fat (2%), low-fat (1%) or nonfat milk is the standard until 2014; afterwards only low-fat (1%) or non-fat milk is the standard. Whole milk is the standard milk for 1-year-old children. Milk is allowed to be partially substituted by cheese, and since 2015, states are allowed to authorize milk to be partially substituted by yogurt (only low-fat and non-fat yogurts can be authorized for children over 2 years of age and women). 5: As substitutes to bread, states are allowed to authorize on an equal weight basis brown rice, bulgur, oatmeal, whole-grain barley, or soft corn or whole wheat tortillas, and since 2014 also whole wheat macaroni (pasta). 6: Fresh fruits and vegetables allowance is the larger amount in parenthesis for for children since 2014 and for women since 2009–2010 (depending on a state). Sources: For panel A: Institute of Medicine (2006), pp. 212–215. For panel B: Federal Register, 2007, Vol. 72, No. 234, pp. 68990–68991. Federal Register, 2009, Vol. 74, No. 250, p 69243. Federal Register, 2014, Vol. 79, No. 42, p. 12294.

B Online Appendix: Dataset construction and summary statistics

In this appendix, I describe the construction of the dataset from the NielsenIQ Consumer Panel and Retail Scanner data from years 2006–2016. The Retail Scanner data is used to construct price variables and the Consumer Panel data is used for both purchases and prices.

B.1 Dataset construction

This part describes the construction of the sample and variables, except prices which are discussed in appendix B.3.

Purchases of subsidized products. In the case of bread, I concentrate on loaves of bread. I exclude organic products, multi-unit products, and unusual package sizes. That is, I restrict the sample to four most common package sizes: 16, 20, 22, and 24 oz. Note that in the majority of states, only 16 oz package bread is allowed as a WIC product, some states also allow 24 oz packages.

In the case of milk, I concentrate on dairy white fluid refrigerated milk. I exclude goat milk, lactose free milk, and organic milk. I also drop multi-unit products, and unusual package sizes. That is, the package sizes which I keep are: one gallon, half a gallon, and a quart (0.25 gallon). Note that in the majority of states, only one gallon package size is allowed as a WIC product, a few states also allow half a gallon and a quart. I keep only packages which are either plastic or carton, and drop glass bottles and other unusual packages.

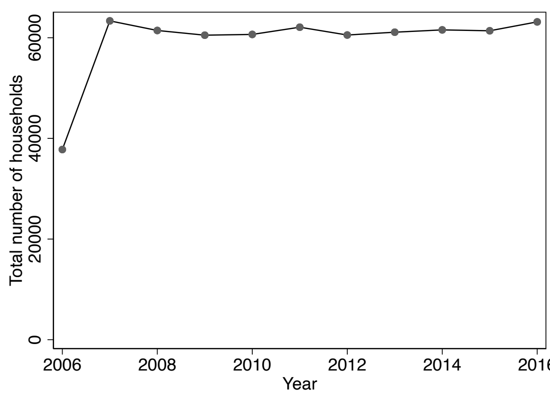
Fruits and vegetables include fresh, frozen, and canned products. Fruits and vegetables include all products with barcodes (UPCs), except potatoes. Juice includes regular, frozen, and concentrated juice, but not fruit drinks. Cereals include breakfast cereals, hot cereals (including grits and cream of wheat), and granola.

Nutritional characteristics. The information about calories, saturated fat, and fiber content is obtained from the USDA databases (USDA, 2018a,b). However, the databases do not include nutritional information for all the products in the Nielsen dataset. I use a sequential matching procedure similar to what is used by Dubois et al. (2014) and Oster (2018).

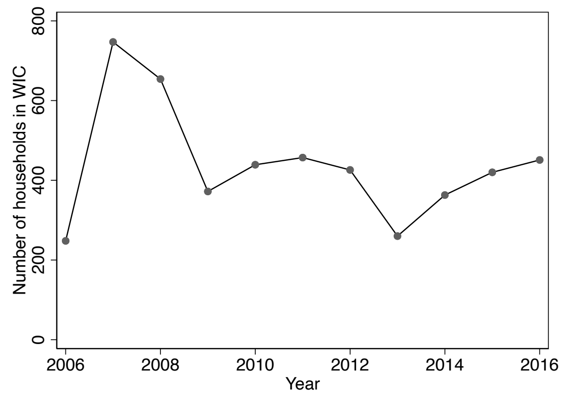
Product definition and aggregation of purchases in reduced form analysis (section 3). In section 3, purchases of each product are summed up to quarterly level. Following WIC regulations, I classify a bread as whole wheat bread if it is made using 100% whole wheat flour. A bread that only partially is made using whole wheat flour is classified as non-whole-wheat. I call the whole wheat bread healthy bread and I don't include whole grain bread in healthy bread because majority of whole grain breads don't satisfy the WIC requirement of having whole grain as the primary ingredient by weight.

Sample for reduced form analysis (section 3). In section 3, the reduced form analysis focuses on households that report receiving WIC assistance and who are in the Consumer Panel data before, during, and at least two years after receiving the WIC vouchers. Households are not necessarily in the Consumer Panel continuously, for all years in a row. I require that each household is in the dataset continuously at least at the time when the household starts to receive the vouchers (the last year before receiving the vouchers and the first year while receiving the vouchers) and when he stops receiving the vouchers (the last year while receiving the vouchers and the first year after receiving the vouchers). This allows analyzing within-household changes when a household starts and stops receiving the WIC vouchers.

In 2006–2016, 3414 households report currently receiving WIC vouchers. Figure B1 presents for each year, the total number of households and households that report currently receiving WIC vouchers. In order to analyze within-household changes when households start receiving the vouchers and two years after they stop receiving the vouchers, one would need to observe the households for a minimum of four years. However, only 1336 of the households are in the dataset for that long. Moreover, only 571 households are observed before, while, and after receiving the vouchers. Requiring that each household is observed the last year before and the first and the second year after receiving the vouchers reduces the sample to 347 households. How much each restriction decreases the sample is described in table B1.



(a) Total



(b) Reporting receiving WIC voucher

Figure B1: Number of households

Note: The increase in the number of households reporting receiving WIC vouchers from year 2006 to 2007 is partly due to the increase in the NielsenIQ Consumer Panel from about 40,000 to about 60,000 households.

Table B1: Sample construction

	Number of households
NielsenIQ Consumer Panel dataset 2006-2016	158004
.. report receiving WIC vouchers	3414
.. & in the dataset at least 4 years	1336
.. & in the dataset before report receiving WIC	695
.. & in the dataset after report receiving WIC	571
.. & in the dataset in the 1st and 2nd year after report receiving WIC	374
.. & in the dataset the latest year before report receiving WIC	347

Table B2 presents the comparison of households included versus excluded from the sample. It shows that the sample and the excluded households are similar in terms of education, race, and marital status, but the households in the sample have higher income (by 27%), smaller size (by 13%), are less likely to be Hispanic ethnicity (by 69%), and are less likely not to work full time (by 12%). It also shows that while the consumer panel matches the overall WIC population rather well in terms of household size and race, it misses a large share of low-income and a majority of Hispanic households. As has been already noted in the literature (Khan et al., 2016), the average income of WIC households in the consumer panel is more than twice as high as in the administrative data. Also, over 40% of the WIC population is of Hispanic ethnicity, while in the consumer panel only about 12% and in the final sample only about 4%.

Table B2: Comparison of WIC households in the sample, excluded from the sample WIC households in the NielsenIQ Consumer Panel, and administrative data

	NielsenIQ WIC households 2006-2016		Wilcoxon rank test	Administrative data WIC households 2016
	Sample	Excluded	p-value	
	(1)	(2)	(3)	(4)
		Panel A: Demographic characteristics		
Household income (thousands)	52.2	41.2	0.000	18.4
Household size	3.5	4.0	0.000	4.1
Race: white	0.723	0.739	0.517	0.586
Ethnicity: Hispanic	0.040	0.128	0.000	0.418
College	0.395	0.404	0.733	.
Married	0.726	0.741	0.540	.
Not employed full time	0.709	0.805	0.000	.
		Panel B: Region		
Northeast	13.0	15.7	.	14.3
Midwest	28.8	26.5	.	18.3
South	40.3	39.0	.	41.7
West	17.9	18.8	.	25.7
Number of households	347	3067		

Notes: Source for the administrative data in column 4 is the USDA report (Thorn et al., 2018). The average income in the report is calculated to 2015 dollars using the consumer price index for urban consumers. Column 1–2 present unweighted data, that is, NielsenIQ household weights are not used.

The characteristics of the households in the sample by treatment and control group are presented in columns 1–2 of table 1.

I use other households with children or lower income or both as an additional control group (column 3 in table 1). Specifically, these households at least for one year had a child up to 17 years old or had income not exceeding 200% of the federal poverty guidelines or both.

Product definition and aggregation of purchases in demand model estimation (section 4). In section 4, I group bread into four products: white, whole wheat, whole grain, and other. White bread is a non-whole-grain bread made from white refined flour. Following WIC regulation, I classify bread as whole wheat bread if it contains 100% whole wheat flour. I classify bread that is only partially made with whole wheat flour or contains any other type of whole grain flour or rye flour whole grain bread. Finally, other bread aggregates together all other non-whole-grain bread types (like Cinnamon Raisin, Cinnamon Swirl, etc.).

Purchases of each product are aggregated to week and (aggregate) retailer level. The aggregation is done in the following way. First, I aggregate purchases to week, bread type, and retail chain level. If a household bought more than one type of bread in a retail chain in a given week, I keep the largest purchase by weight. If a household bought bread in more than one retail chain per week, then I keep the largest purchase by weight in a chain in the Retail Scanner Dataset (if any exists) and the largest purchase in a chain not in the Retail Scanner Dataset (if any exists).

Sample for demand model (section 4). To keep the demand model estimation manageable, I restrict attention to the sample of households who all have reported receiving WIC assistance (households in columns 1–2 of table 1). The demand model is estimated using only the time periods before and after receiving WIC vouchers. While households received WIC vouchers, I don't know whether they paid with a voucher or cash on a given shopping trip. Hence, I don't know the price they faced on a given shopping trip: whether the regular price or price zero because of the voucher.

I use only purchases in grocery stores and mass merchandisers (discount stores). Drug stores and others are excluded because those are unlikely to sell all four types of bread.

In the first year of each household in the dataset, I don't use the first ten shopping trips in the estimation. But I use these purchases to construct the stock variable of whole wheat bread past purchases. Ideally, I would exclude the whole first year from the estimation, but that would result in excluding all the households who are in the dataset only one year before receiving the WIC vouchers.

I exclude households that don't have a sufficient number of bread purchases in grocery and discount stores. Each purchase should also have enough information about the past purchases to construct the stock variable of past year's whole wheat bread purchases. Specifically, I exclude households who don't have purchases before receiving the vouchers. I also require that a household has purchases in the first and second year after receiving the vouchers, and altogether at least ten purchases in the first two years after receiving the vouchers. In order to use the impact of WIC policy reform for identification, I also require that a household purchased bread in the grocery and discount stores during the last year on WIC.

Construction of demographic variables. I generate income variable as the mid-point of the reported income interval. Except for the highest income group, for which the mid-point cannot be calculated as no highest level is reported. For that group I assume that their household annual income equals 115,000 dollars, which is consistent with the current income distribution. Income is deflated to 2015 dollars using the consumer price index for urban consumers.

Underreporting of WIC status. In the survey data, household's WIC status seems to be underreported. The information on WIC status is collected via a survey question, which households are not required to answer. Therefore, as expected, in the dataset, the reported share of households receiving WIC assistance is lower than in the administrative data. Figure B2 shows the percentage of income eligible households who report receiving WIC assistance, by the age of the youngest child. The figure illustrates two aspects of the data. First, the magnitude of underreporting. Back-of-the-envelope calculations suggest that about forty percent of WIC participants report it in the survey.[29] Second, the households' reported WIC participation mirrors well the pattern in the administrative data. Namely, WIC participation is the highest during the calendar year when the child is born and decreases after that up to the child's fifth birthday. In the analysis, I focus on households who report receiving WIC assistance.

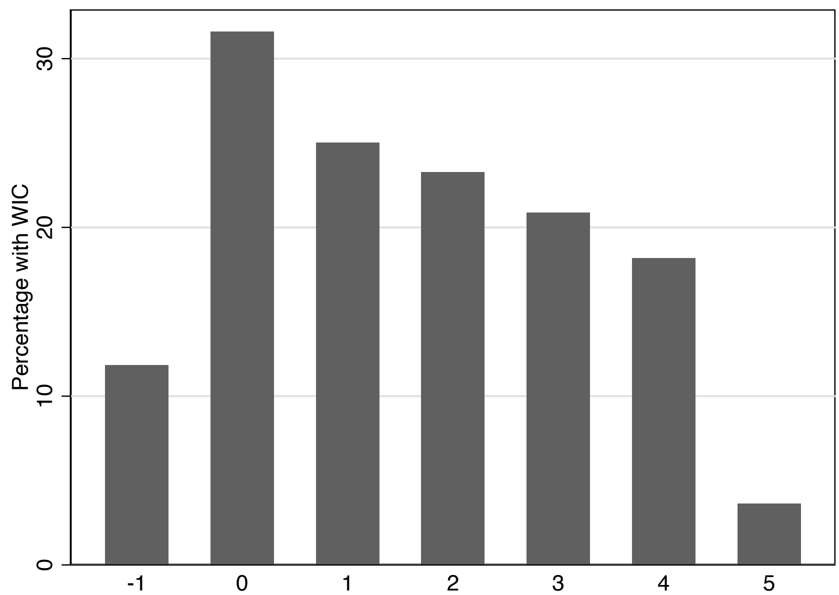


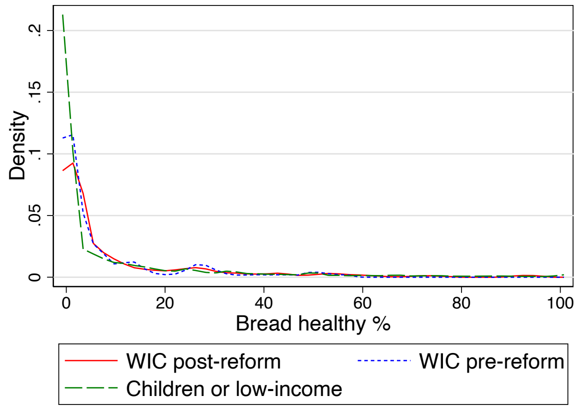
Figure B2: Percentage of households reporting receiving WIC assistance, by the years since the youngest child is born

Note: Sample is restricted to WIC income eligible households. Calculated as an average over household-years and weighted by Nielsen projection weights. Year 0 indicates that the youngest child is born during the current panel year. Income eligibility is calculated using the mid-point value of Nielsen reported household income interval.

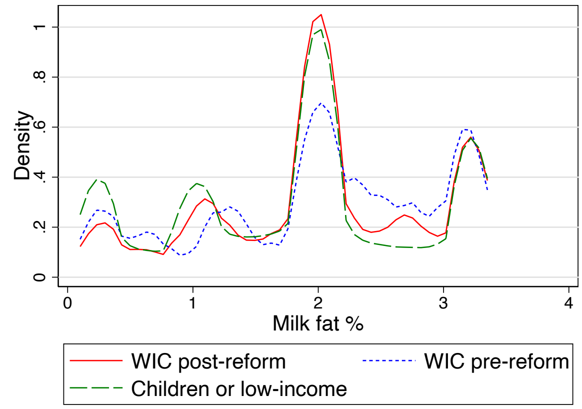
[29]According to the USDA report (Johnson et al., 2014), around 80% of eligible infants (up to their first birthday) receive WIC. However, in this dataset, the reported WIC participation of the corresponding households is slightly above 30%. According to the administrative data, participation of children aged 1–4 is about 50%, but in this dataset, only around 20%.

How could the underreporting bias the analysis? First, if non-WIC households in the control group also received the WIC vouchers, then the short-term effect would be underestimated. It is reassuring that later in the analysis, when non-WIC households are excluded, the estimates remain similar. Second, if households selectively report the years when they receive the WIC vouchers, specifically, if underreporting increases with the number of years in the sample, then a household would be mistakenly assigned to the post-WIC period while it is still receiving the vouchers. In that case, the persistence of WIC would be overestimated. It is somewhat reassuring that in section 3, we will see that the estimated persistence is rather low already.

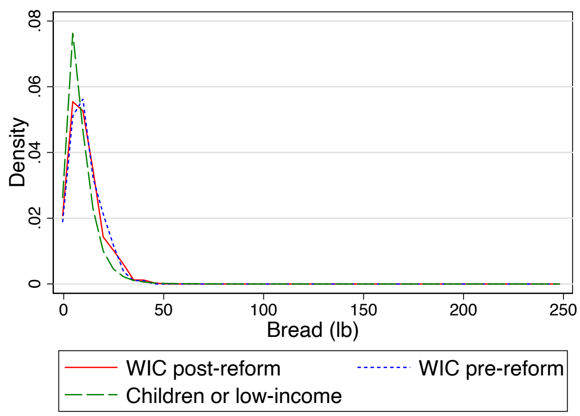
B.2 Additional summary statistics of the dataset for reduced form analysis



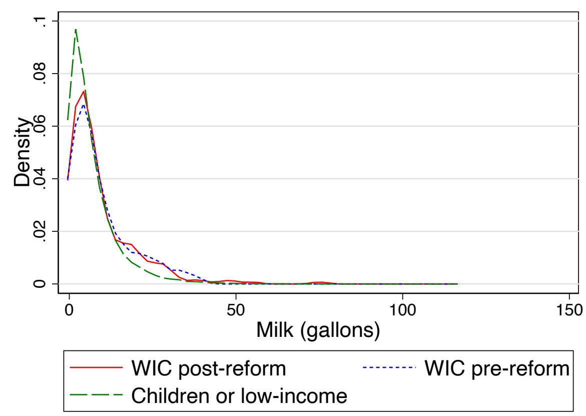
(a) Bread healthy %



(b) Milk fat %



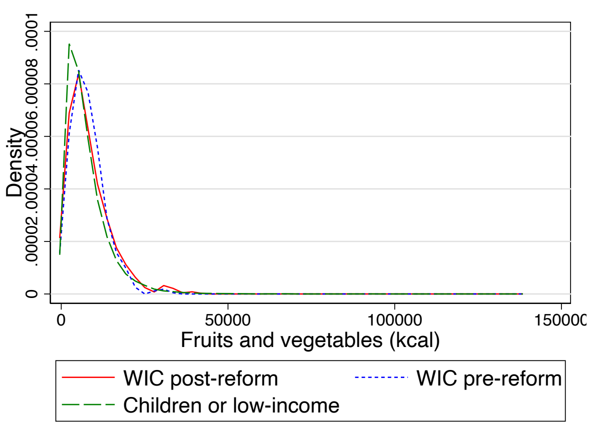
(c) Quantity of bread



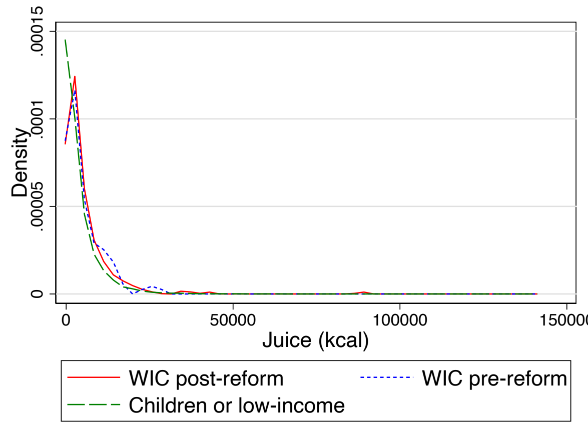
(d) Quantity of milk

Figure B3: Kernel density estimates of the distribution of purchases of bread and milk

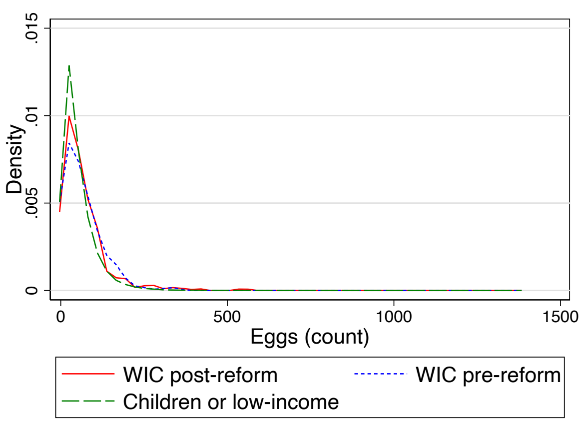
Note: The kernel density estimates of the distribution of quarterly average household purchases during first year in the sample, separately for the three groups in table 1. Estimates are obtained using Epanechnikov kernel and the bandwidth that would minimize the mean integrated squared error if the data were Gaussian and a Gaussian kernel were used. Except for milk fat percentage which uses bandwidth 0.1 because with the multimodal distribution the above bandwidth would oversmooth the density.



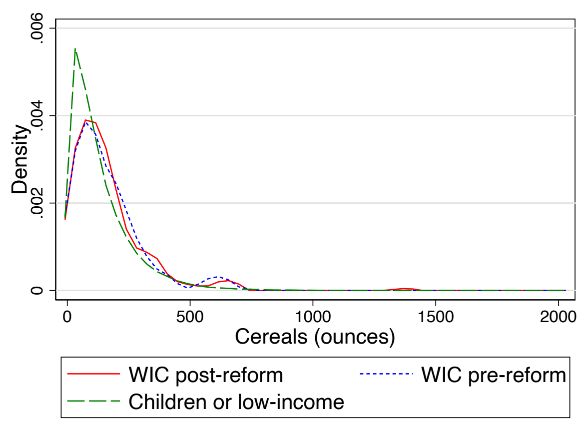
(a) Quantity of fruits and vegetables



(b) Quantity of juice



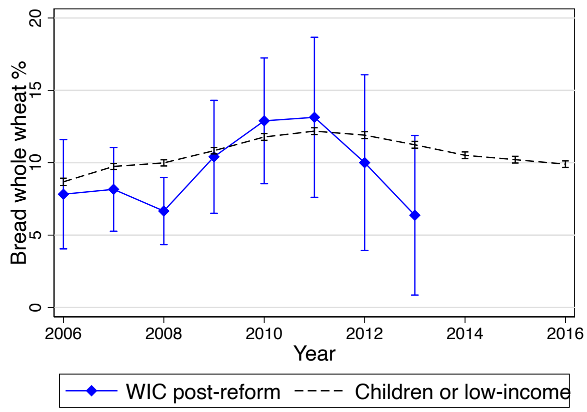
(c) Quantity of eggs



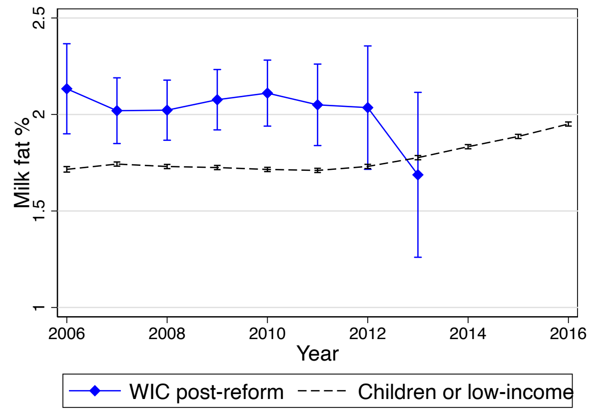
(d) Quantity of cereals

Figure B4: Kernel density estimates of the distribution of purchases of other products

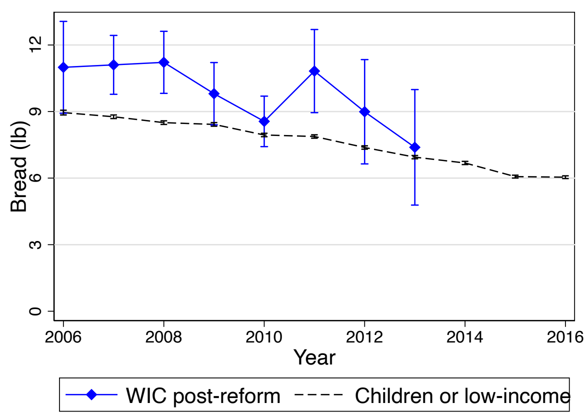
Note: The kernel density estimates of the distribution of quarterly average household purchases during first year in the sample, separately for the three groups in table 1. Estimates are obtained using Epanechnikov kernel and the bandwidth that would minimize the mean integrated squared error if the data were Gaussian and a Gaussian kernel were used.



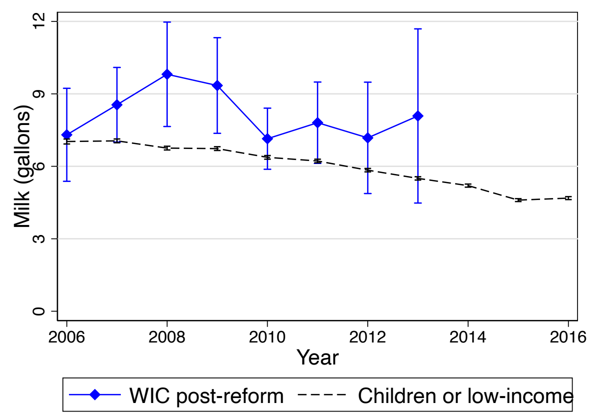
(a) Bread healthy %



(b) Milk fat %



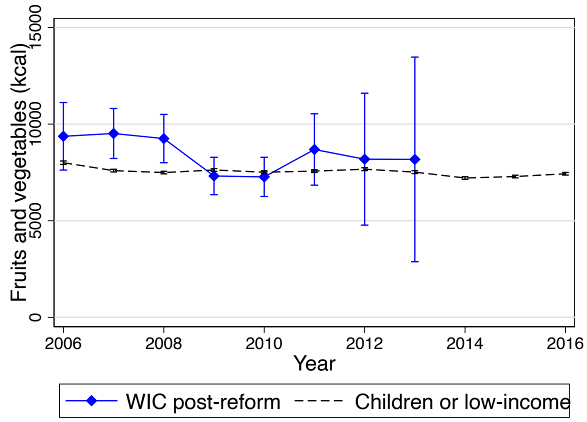
(c) Quantity of bread



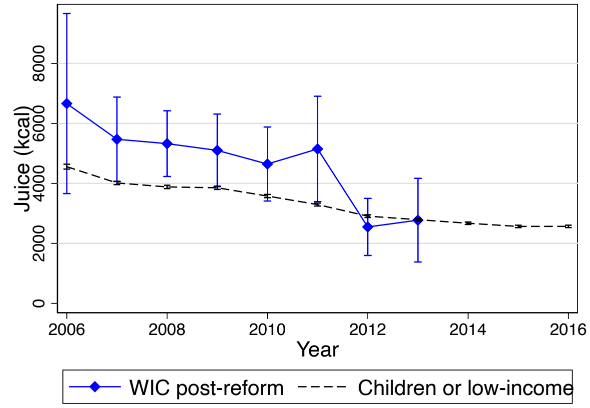
(d) Quantity of milk

Figure B5: Average quarterly purchases over time, bread and milk

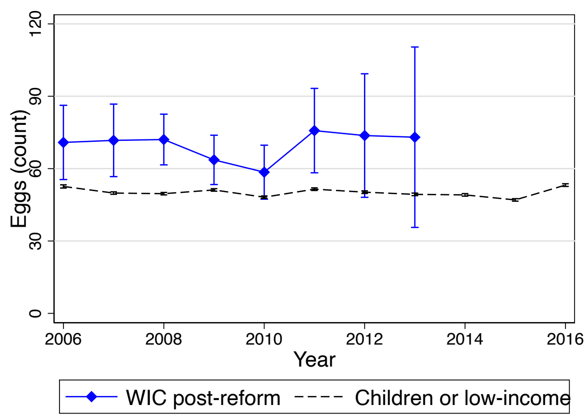
Note: Averages and 95% confidence intervals of households' yearly average quarterly purchases. When calculating the average and confidence intervals, for each year, the unit of observation is a household. The blue line presents the average purchases for the WIC post-reform households (column 1 in table 1) and the black line for the additional control group of Children or low-income households (column 3 of table 1). For the WIC post-reform households only years before receiving WIC vouchers are included. It is not possible to look at the trend for WIC pre-reform households using years before vouchers, because that would be limited to only the pre-reform years. Depending on a year, the averages are calculated based 22–139 households for the WIC post-reform and 22,267–38,565 households for the Children or low-income group.



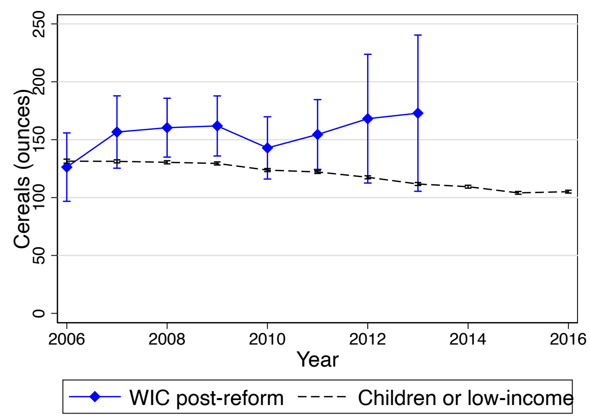
(a) Quantity of fruits and vegetables



(b) Quantity of juice



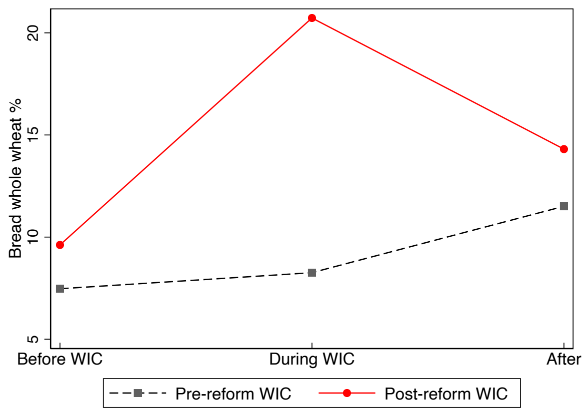
(c) Quantity of eggs



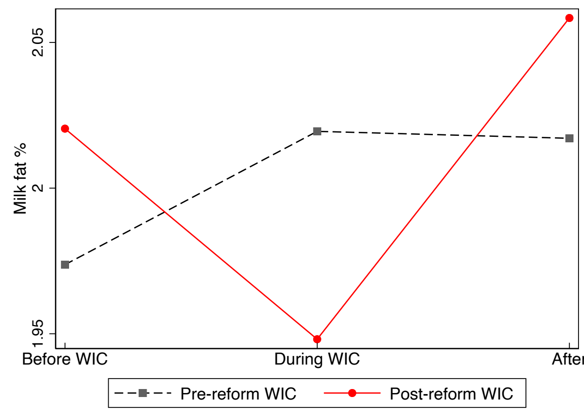
(d) Quantity of cereals

Figure B6: Average quarterly purchases over time, other products

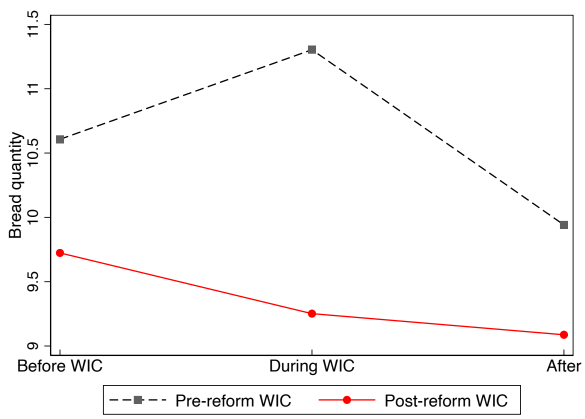
Note: Averages and 95% confidence intervals of households' yearly average quarterly purchases. When calculating the average and confidence intervals, for each year, the unit of observation is a household. The blue line presents the average purchases for the WIC post-reform households (column 1 in table 1) and the black line for the additional control group of Children or low-income households (column 3 of table 1). For the WIC post-reform households only years before receiving WIC vouchers are included. It is not possible to look at the trend for WIC pre-reform households using years before vouchers, because that would be limited to only the pre-reform years. Depending on a year, the averages are calculated based 22–139 households for the WIC post-reform and 22,267–38,565 households for the Children or low-income group.



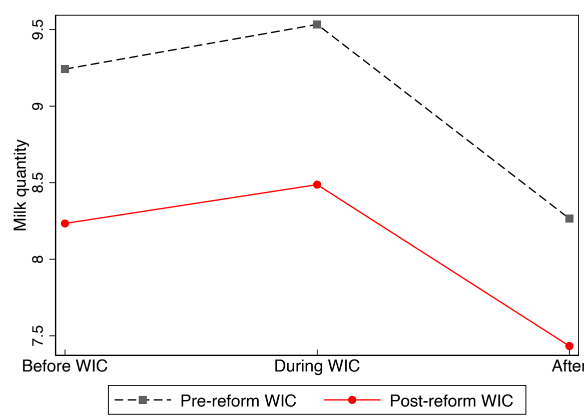
(a) Bread healthy %



(b) Milk fat %



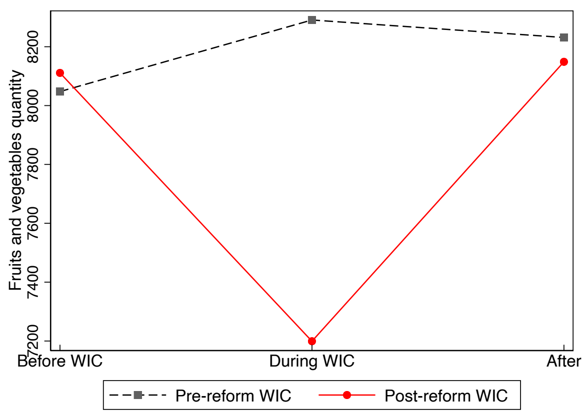
(c) Quantity of bread



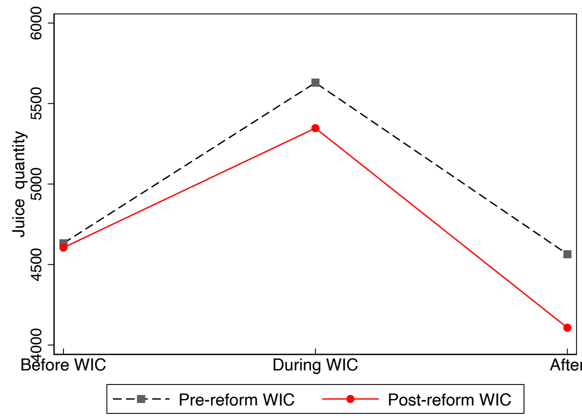
(d) Quantity of milk

Figure B7: Comparison of changes in purchases before, while, and after receiving WIC vouchers, pre-reform WIC versus post-reform WIC households, bread and milk

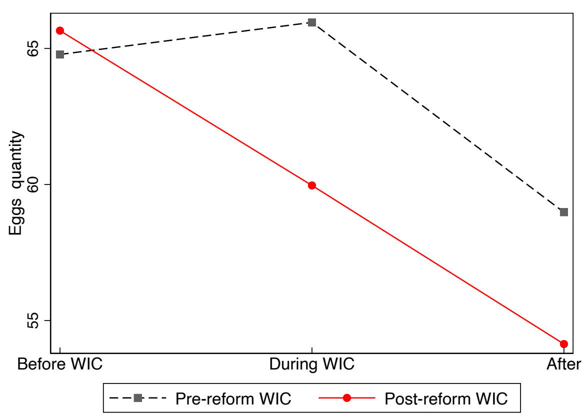
Note: The sample of *Pre-reform WIC* includes households in column 2 of table 1 who started to receive WIC vouchers in 2007–2008. The sample of *Post-reform WIC* includes households in column 1 of table 1 who started to receive WIC vouchers in 2010–2011. The sample excludes time periods earlier than four years before receiving the vouchers and later than two years after receiving the vouchers.



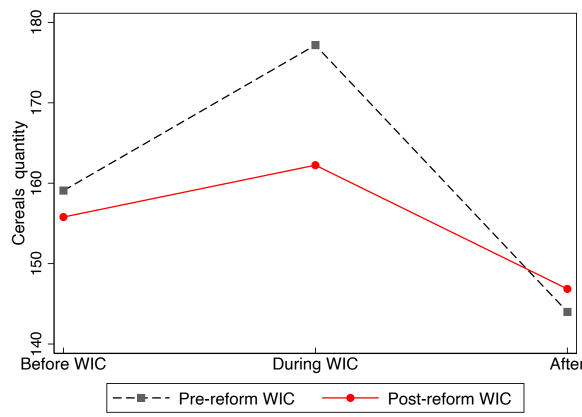
(a) Quantity of fruits and vegetables



(b) Quantity of juice



(c) Quantity of eggs



(d) Quantity of cereals

Figure B8: Comparison of changes in purchases before, while, and after receiving WIC vouchers, pre-reform WIC versus post-reform WIC households, other products

Note: The sample of *Pre-reform WIC* includes households in column 2 of table 1 who started to receive WIC vouchers in 2007–2008. The sample of *Post-reform WIC* includes households in column 1 of table 1 who started to receive WIC vouchers in 2010–2011. The sample excludes time periods earlier than four years before receiving the vouchers and later than two years after receiving the vouchers.

B.3 Prices

Bread prices are constructed from the Retail Scanner Data for shopping trips to retail chains in the Retail Scanner dataset. To do that, prices are aggregated to the bread type, week, market (Nielsen scantrack market) and chain level. But more than half of the WIC households' bread purchases are at retailers (grocery stores and mass merchandisers) that are not in the Retail Scanner Data. For those shopping trips the prices are constructed from the Consumer Panel Data and prices are aggregated to the bread type, week, and market (Nielsen scantrack market) level. In both cases, prices are deflated to 2015 dollars using the consumer price index for urban consumers.

I use data on sales and purchases in grocery stores and mass merchandisers (discount stores). Drug stores and others are excluded.

Retail Scanner Data. Using the Retail Scanner Data, I define a chain by the retailer code. If a chain is present in only one market, then to be able to calculate instruments (average price in other markets), I follow DellaVigna and Gentzkow (2019) and split the market. To do that, I assign stores by 3 digit zip code into one of the two split markets, by first ranking 3 digit zip codes by the number of stores and then using a greedy algorithm to assign zip codes to split markets with a goal of making the split markets as equal as possible by the number of stores.

To aggregate UPC level prices to product prices, first I convert these to price per ounce. Then, for each of the products (white, whole wheat, whole grain, and other bread), I calculate the weighted average prices. The weight of the UPC is equal to its share of volume of that product in a market in a given year in a retail chain.

Consumer Panel Data. Using the Consumer Panel Data, I ignore consumers' possible use of coupons, and use the price without subtracting the value of coupons. I convert all

UPC level prices to price per ounce to be able to aggregate to product level prices. As is standard in the literature that constructs prices from the Consumer Panel data, I exclude outliers. Specifically, I exclude UPCs that are bought by less than 25 households in the entire dataset. I also exclude purchases if the price is more than four times the median price or less than one-fourth of the median of all purchases in the entire dataset for that UPC. Finally, for each bread type, market, year, and aggregate retailer combination, I replace prices lower than the first percentile with the first percentile and similarly, replace prices higher than the 99th percentile with the 99th percentile.

The Consumer Panel Data has price information about a UPC in a given week in a given market only when there is a household in the sample who buys the product. We don't observe the price of a UPC if noone in the sample bought it. I use the following method to fill in the missing price information. First for each week-market-chain triplet, I calculate UPCs average price (using the household sample weights included in the Consumer Panel Data). If in a given week the price of the UPC is missing in a market in a chain, then I use the price of the same UPC in the same chain in the same market within the previous and following 4 weeks. If there is still no price information available for the UPC, then I exclude the UPC for that week and market and chain.

Using these UPC-chain level prices, I calculate the weighted average product (white, whole wheat, whole grain, and other bread) prices for two aggregate retailers. The two aggregate retailers are first, those in the Retail Scanner Data, and second, those not in the Retail Scanner Data. The weight of the UPC and chain pair is equal to its share of volume of that product in a market in a given year in an aggregate retailer. I use sample weights provided in the dataset to calculate the weighted average.

Price comparison. How different are the prices constructed from the two alternative datasets? We can look at that for the subset of retailers that is overlapping in these datasets.

Table B3 presents summary statistics of prices calculated from Consumer Panel and Retail Scanner data. In columns 2–3 prices are calculated from the Consumer Panel data. These prices are calculated aggregating across all the retailers in the Retail Scanner data (column 3) or across all the retailers not in the Retail Scanner (column 2). Column 4 presents prices calculated from the Retail Scanner data, these prices are calculated at the retailer level. Column 1 combines prices from columns 2 and 4, these are the prices used in demand model estimation. Column 5 compares prices for the retailers in the Retail Scanner data calculated using Consumer Panel (column 3) and Retail Scanner data (column 4). That is, column 5 presents the percentage difference (divided by 100) between the prices in columns 4 and 3.

For these overlapping retailers, bread prices constructed based on the two datasets line up rather closely, prices based on the consumer panel data are on about 10% lower for all four product types (column 5 in table B3). The magnitude of the difference is very similar to what has been shown for aggregate grocery price indexes constructed based on the two datasets (for example, Seo (2019)). The mis-measurement of prices is unlikely to be a substantial concern since the difference is small, and importantly, in any given choice occasion (shopping trip), consumers face prices constructed in the same way.

Table 2 presents summary statistics of prices used in the demand model estimation.

Table B4 shows that majority of the households in the sample make bread purchases both in retailers that are only in the consumer panel and in those in the retail scanner data. The market shares of different types of bread purchases in those two types of retailers are also rather similar.

Table B3: Summary statistics of bread prices calculated using alternative methods

	Consumer Panel & Retail Scanner	Stores only in	Overlapping stores		
		Consumer Panel	Consumer Panel	Retail Scanner	%Δ
	(1)	(2)	(3)	(4)	(5)
			Panel A: All bread products		
Mean	1.77	1.64	1.79	1.98	-0.10
Sd	0.41	0.33	0.35	0.42	-0.17
Min	0.61	0.61	0.84	0.78	0.07
25th perc.	1.49	1.40	1.51	1.68	-0.10
Median	1.75	1.65	1.82	2.00	-0.09
75th perc.	2.03	1.87	2.05	2.29	-0.10
Max	4.27	2.90	3.03	4.27	-0.29
Number of obs.	112168	68608	43560	43560	43560
			Panel B: Whole wheat bread		
Mean	1.95	1.77	2.07	2.23	-0.07
Sd	0.34	0.22	0.23	0.30	-0.25
Median	1.88	1.74	2.07	2.22	-0.07
Number of obs.	28042	17152	10890	10890	10890
			Panel C: Whole grain bread		
Mean	2.03	1.92	1.97	2.20	-0.10
Sd	0.30	0.23	0.28	0.33	-0.16
Median	1.95	1.90	1.99	2.17	-0.09
Number of obs.	28042	17152	10890	10890	10890
			Panel D: White bread		
Mean	1.36	1.24	1.38	1.54	-0.10
Sd	0.28	0.19	0.19	0.30	-0.37
Median	1.30	1.22	1.35	1.49	-0.09
Number of obs.	28042	17152	10890	10890	10890
			Panel E: Other bread		
Mean	1.76	1.62	1.74	1.97	-0.12
Sd	0.33	0.22	0.23	0.36	-0.35
Median	1.67	1.58	1.70	1.94	-0.12
Number of obs.	28042	17152	10890	10890	10890

Note: Price of bread is measured in dollars per pound, deflated to 2015 dollars using the consumer price index for urban consumers. Observations are at the week, market, bread type, and (aggregate) retailer level. Columns 2–3 present prices calculated from the Consumer Panel data. These prices are calculated aggregating across all the retailers in the Retail Scanner data (column 3) or across all the retailers not in the Retail Scanner (column 2). Column 4 presents prices calculated from the Retail Scanner data, these prices are calculated at the retailer level. Column 1 combines prices from columns 2 and 4. Column 5 presents the percentage difference (divided by 100) between the prices in columns 5 and 4.

Table B4: Descriptive statistics of bread purchases

	Retailers in Only in Consumer Panel (1)	Retail Scanner (2)
Share of households purchasing in retailers	0.98	0.87
Number of retailers household purchases (median)	3	2
Number of households	252	252
	Bread type market shares	
Other bread	0.32	0.33
Whole grain bread	0.13	0.20
Whole wheat bread	0.12	0.10
White bread	0.42	0.37
Number of shopping trips	17152	10890

C Online Appendix: Additional analysis for the short- and long-term impact of healthy food subsidies

This appendix reports additional analysis that are mentioned in the main text of section 3.

C.1 Long-term effect separately for years 1 and 2

Table C1 presents the estimates of the long-term effect separately for years 1 and 2 after receiving the WIC vouchers. It estimates the following regression similar to the main specification (equation (1)):

$$\begin{aligned}
Y_{it} = {} & \alpha_1 WIC_{it} + \alpha_{2,1} AfterWICYear1_{it} + \alpha_{2,2} AfterWICYear2_{it} \\
& + \beta_1 ReformedWIC_{it} + \beta_{2,1} AfterReformedWICYear1_{it} + \beta_{2,2} AfterReformedWICYear2_{it} \\
& + \alpha_3 AfterWICYear3^{+}_{it} + \beta_3 AfterReformedWICYear3^{+}_{it} \\
& + \sigma \cdot PostReform_{it} + X_{it}\zeta + \delta_i + \gamma_t + \varepsilon_{it} \qquad (5)
\end{aligned}$$

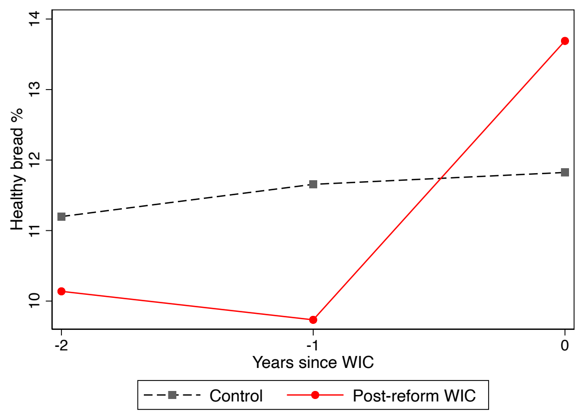
Compared to the estimates of the long-term effect β_2 in the main specification (table 3), the corresponding estimates $\beta_{2,1}, \beta_{2,2}$ are similar, but less precise.

Table C1: The impact of post-reform WIC on subsidized products; long-term effect separately for years 1 and 2 after receiving the WIC vouchers

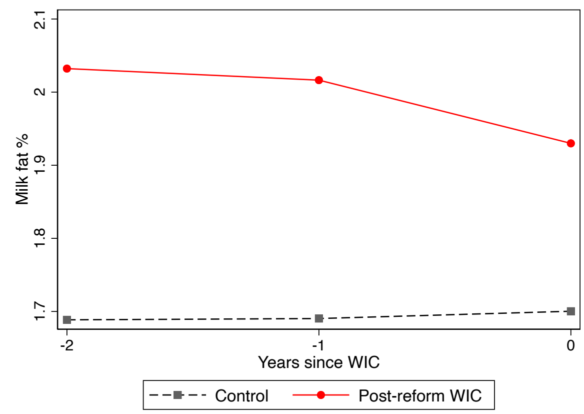
	Bread healthy % (1)	Milk fat % (2)	Bread log. quantity (3)	Milk log. quantity (4)	Fruits&veg. log. quantity (5)	Juice log. quantity (6)	Eggs log. quantity (7)	Cereals log. quantity (8)
β_1: Reformed WIC	6.925***	-0.135***	-0.013	0.022	-0.002	0.275	-0.064	0.036
	(2.212)	(0.050)	(0.053)	(0.054)	(0.087)	(0.206)	(0.095)	(0.085)
$\beta_{2,1}$: After Reformed WIC Year 1	4.033*	-0.046	-0.056	0.030	-0.065	0.604**	-0.026	0.228*
	(2.074)	(0.069)	(0.071)	(0.071)	(0.111)	(0.299)	(0.120)	(0.123)
$\beta_{2,2}$: After Reformed WIC Years 2	-1.255	-0.082	0.060	-0.022	-0.023	-0.288	-0.206	0.150
	(2.157)	(0.075)	(0.082)	(0.076)	(0.142)	(0.305)	(0.133)	(0.146)
α_1: WIC	-1.799	0.003	0.040	0.051	0.037	0.456***	0.073	0.132*
	(1.332)	(0.035)	(0.041)	(0.046)	(0.066)	(0.166)	(0.079)	(0.069)
$\alpha_{2,1}$: After WIC Year 1	0.672	-0.044	0.037	-0.003	0.084	-0.161	-0.038	-0.126
	(1.424)	(0.059)	(0.056)	(0.058)	(0.089)	(0.265)	(0.100)	(0.105)
$\alpha_{2,2}$: After WIC Year 2	2.708	0.028	-0.022	0.064	-0.009	0.604**	0.079	-0.101
	(1.758)	(0.067)	(0.069)	(0.065)	(0.120)	(0.261)	(0.112)	(0.128)
σ: Post-Reform	0.207	-0.014**	-0.001	0.008	0.002	0.111***	0.021	-0.032*
	(0.270)	(0.007)	(0.009)	(0.008)	(0.018)	(0.037)	(0.015)	(0.017)
Year-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald test, $\beta_1 = \beta_{2,1}$, p-value	0.136	0.153	0.495	0.893	0.560	0.265	0.746	0.105
Wald test, $\beta_1 = \beta_{2,2}$, p-value	0.001	0.485	0.363	0.590	0.887	0.062	0.317	0.440
WIC households	301	301	347	347	347	347	347	347
Households	85421	80721	96637	96637	96637	96637	96637	96637
Household-quarters	1244237	1192570	1602252	1602252	1602252	1602252	1602252	1602252

Note: Each column presents estimates from a separate panel data fixed effects regression (equation (5)). A unit of observation is a household-quarter pair. Bread healthy (whole wheat) % and milk fat % are in the range of 0–100, quantity is measured in pounds for bread, gallons for milk, kilocalories for fruits and vegetables and juice, counted for eggs, and ounces for cereals. All regressions include household size, logarithm of income, dummies for the year before child is born, and children aged $0, 1, \ldots, 5$, 6-12, and 13-17. In column 2, regressions include an interaction terms for WIC (and Reformed WIC) and the last two quarters of the calendar year when the child gets one years old and the other with the first two quarters of the calendar year when the child gets two years old. The sample includes households in columns 1–3 in table 1. Regressions include household and time period fixed effects. Standard errors (in parentheses) are clustered at the household level. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

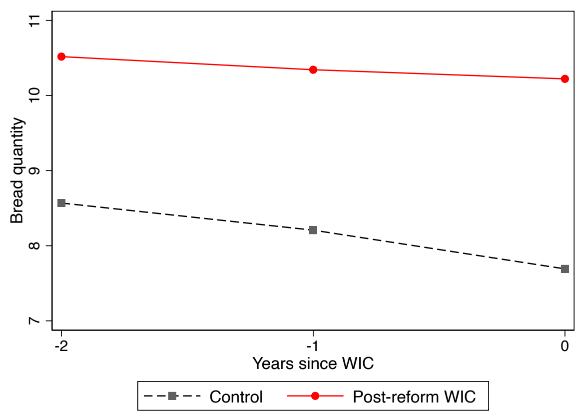
C.2 Pre-trends



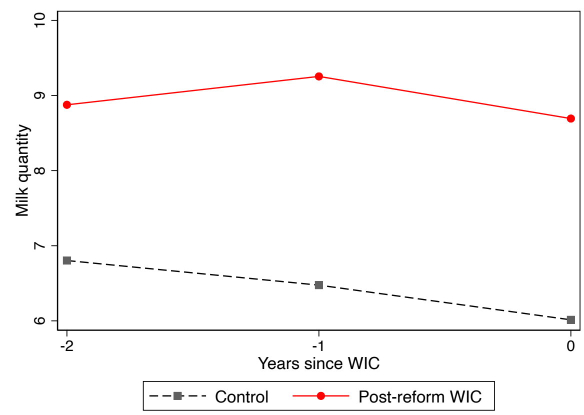
(a) Bread healthy %



(b) Milk fat %



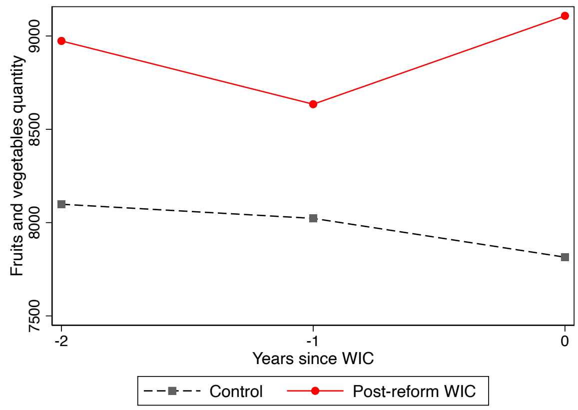
(c) Quantity of bread



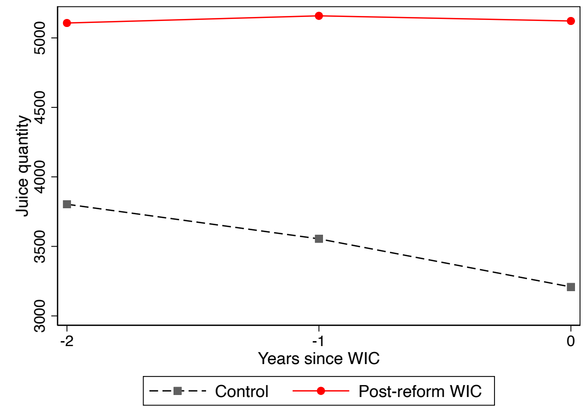
(d) Quantity of milk

Figure C1: Pre-trends comparison households in post-reform WIC versus control group of households not receiving WIC vouchers, bread and milk

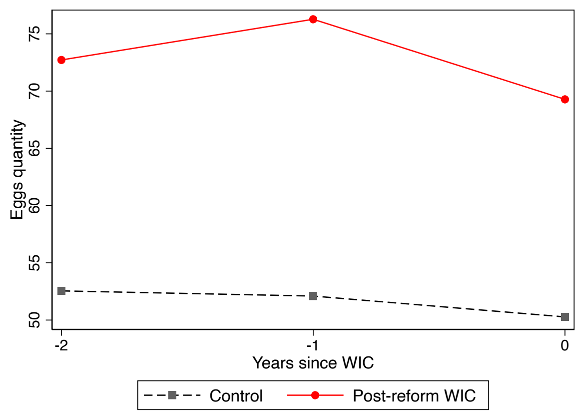
Note: The sample of *Post-reform WIC* (treatment group) includes households in column 1 of table 1 who are in the dataset at least two years before starting to receive the post-reform WIC vouchers. The sample of *Control* includes households in column 3 of table 1, it does not include households in column 2 because only very few of those are in the dataset two years before receiving the vouchers. For milk fat percentage, the sample excludes observations in the treatment group, when a household has a one year old child because for one-year-old children, the reformed WIC food vouchers make an exception and provide whole milk. The figure presents averages of purchases from two years before receiving the post-reform WIC vouchers (-2) to the first year of receiving the post-reform WIC vouchers (0), for the treatment group (the households receiving the post-reform vouchers) and the corresponding control group. The averages are calculated in the following steps. First, each treatment group cohort (households that started receiving post-reform vouchers in year t) is matched to the corresponding control group: households who are in the dataset continuously for years $t-2$ to t. Second, the cohort averages are calculated of both the treatment and control group. Third, the weighted average over the cohorts ($t = 2009, ..., 2014$) is calculated, where the weights are equal to the treated cohort's share in all the treated observations.



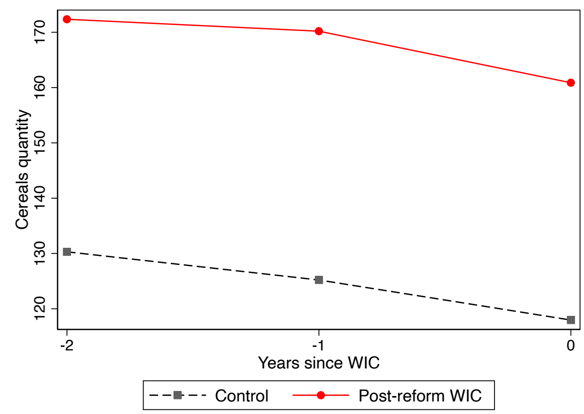
(a) Quantity of fruits and vegetables



(b) Quantity of juice



(c) Quantity of eggs



(d) Quantity of cereals

Figure C2: Pre-trends comparison households in post-reform WIC versus control group of households not receiving WIC vouchers, other products

Note: The sample of *Post-reform WIC* (treatment group) includes households in column 1 of table 1 who are in the dataset at least two years before starting to receive the post-reform WIC vouchers. The sample of *Control* includes households in column 3 of table 1, it does not include households in column 2 because only very few of those are in the dataset two years before receiving the vouchers. The figure presents averages of purchases from two years before receiving the post-reform WIC vouchers (-2) to the first year of receiving the post-reform WIC vouchers (0), for the treatment group (the households receiving the post-reform vouchers) and the corresponding control group. The averages are calculated in the following steps. First, each treatment group cohort (households that started receiving post-reform vouchers in year t) is matched to the corresponding control group: households who are in the dataset continuously for years $t-2$ to t. Second, the cohort averages are calculated of both the treatment and control group. Third, the weighted average over the cohorts ($t = 2009, ..., 2014$) is calculated, where the weights are equal to the treated cohort's share in all the treated observations.

C.3 Robustness analysis of the main reduced form estimates

C.3.1 Analysis based only on households that received the WIC vouchers

To alleviate the concern, that households start receiving post-reform WIC vouchers due to a shock that changes food purchases, I compare only the pre- and post-reform WIC households who arguably are all subject to similar shocks (related to child's birth) when starting to receive WIC vouchers. Table C2 presents estimates of the main specification (equation (1)), except that time period fixed effects are excluded, as these would not be identified. The short-term effects' estimates, using this sample of only pre- and post-reform WIC households, are similar to those from the main specification; long-term effects' estimates are qualitatively similar. The difference between the short- and long-term effect is statistically significant at the 1% level for the healthy bread and 5% level for the milk fat percentage.

Table C2: The impact of post-reform WIC on subsidized products; sample restricted to WIC households

	Bread healthy %	Milk fat %	Bread log. quantity	Milk log. quantity	Fruits&veg. log. quantity	Juice log. quantity	Eggs log. quantity	Cereals log. quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
β_1: Reformed WIC	4.583**	-0.143***	-0.017	-0.035	-0.044	-0.132	-0.135	-0.012
	(2.117)	(0.052)	(0.053)	(0.056)	(0.091)	(0.211)	(0.099)	(0.097)
β_2: After Reformed WIC Years 1-2	-0.456	-0.015	-0.052	-0.067	-0.024	-0.183	-0.109	0.064
	(1.962)	(0.067)	(0.071)	(0.066)	(0.112)	(0.269)	(0.117)	(0.120)
α_1: WIC	-1.184	0.029	-0.005	0.010	-0.005	0.294*	0.053	0.069
	(1.379)	(0.037)	(0.042)	(0.045)	(0.065)	(0.167)	(0.076)	(0.069)
α_2: After WIC Years 1-2	1.464	0.003	-0.084	-0.092*	-0.113	-0.253	-0.134	-0.218**
	(1.584)	(0.059)	(0.060)	(0.053)	(0.099)	(0.246)	(0.102)	(0.106)
σ: Post-Reform	3.606***	0.020	-0.098**	-0.002	0.004	0.151	0.050	-0.144*
	(1.156)	(0.036)	(0.039)	(0.036)	(0.061)	(0.153)	(0.070)	(0.076)
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald test, $\beta_1 = \beta_2$, p-value	0.006	0.040	0.587	0.604	0.865	0.848	0.827	0.529
WIC households	301	301	347	347	347	347	347	347
Households	301	301	347	347	347	347	347	347
Household-quarters	9133	9064	11224	11224	11224	11224	11224	11224

Note: The table presents estimates from the same regressions as in table 3, but the sample is restricted to households who received WIC vouchers (households in columns 1–2 in table 1) and the regressions do not include time period fixed effects. Standard errors (in parentheses) are clustered at the household level. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

C.3.2 Trends

I re-estimate the main specification allowing separate trends for the treatment group, and the estimates for the healthy bread and milk fat percentage remain similar. Additional tests for parallel trends are provided using the heterogeneity-robust difference-in-differences estimates, which are explained below.

Table C3: The impact of post-reform WIC on subsidized products, including trends for treatment group

	Bread healthy %	Milk fat %	Bread log. quantity	Milk log. quantity	Fruits&veg. log. quantity	Juice log. quantity	Eggs log. quantity	Cereals log. quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
β_1: Reformed WIC	6.742***	-0.147***	0.019	-0.021	-0.084	0.074	-0.159*	0.002
	(2.212)	(0.050)	(0.054)	(0.054)	(0.087)	(0.208)	(0.096)	(0.086)
β_2: After Reformed WIC Years 1-2	1.007	-0.092	0.037	-0.079	-0.192*	-0.286	-0.318***	0.084
	(1.944)	(0.066)	(0.072)	(0.064)	(0.112)	(0.270)	(0.116)	(0.119)
α_1: WIC	-1.872	-0.005	0.011	0.042	0.032	0.340**	0.027	0.073
	(1.332)	(0.036)	(0.043)	(0.046)	(0.066)	(0.168)	(0.080)	(0.069)
α_2: After WIC Years 1-2	1.660	-0.013	-0.006	0.026	0.036	0.165	-0.001	-0.142
	(1.482)	(0.059)	(0.057)	(0.053)	(0.093)	(0.236)	(0.098)	(0.102)
σ: Post-Reform	0.017	0.000	0.000	0.000	0.003	0.004	0.001	0.001
	(0.270)	(0.007)	(0.009)	(0.008)	(0.018)	(0.037)	(0.015)	(0.017)
Year-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald test, $\beta_1 = \beta_2$, p-value	0.004	0.379	0.788	0.365	0.341	0.175	0.184	0.483
WIC households	301	301	347	347	347	347	347	347
Households	85421	80721	96637	96637	96637	96637	96637	96637
Household-quarters	1244237	1192570	1602252	1602252	1602252	1602252	1602252	1602252

Note: The table presents estimates from the same regressions as in table 3, except that it includes trends separately for the cohorts that start receiving post-reform WIC vouchers in 2009–2011 and those that start in 2012-2014. (There are no cohorts starting later as the sample is restricted to WIC households who are observed at least 2 years after receiving the vouchers.) Standard errors (in parentheses) are clustered at the household level. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

C.3.3 Assessing the importance of negative weights in the two-way fixed effects estimator

Recent econometrics literature has raised concerns about the two-way fixed effects estimator when the timing of treatment is staggered and treatment effects are heterogeneous either over time within units or across groups of units treated at different times (Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2020b). The issue is that the two-way fixed effects estimator is a weighted average of the heterogeneous treatment effects where some weights potentially could be negative, which would lead to a bias in the estimates. Due to the negative weights, the estimate may, for example, be negative even when all the average treatment effects are positive.

To assess the concern, first, I calculate the share of negative weights (table C4).[30] Throughout the appendix, I concentrate on analyzing only the short-term effect, because estimating the long-term effects would require allowing switching back from treatment, which the considered methods don't accommodate. Reassuringly, the share of negative weights is either zero or very small which implies that it is unlikely to lead to a substantial bias. To further study whether the treatment effect heterogeneity is potentially problematic for β_1 estimate, we can look at the two diagnostic measures at the bottom of table C4. The first measure $\underline{\hat{\sigma}}_{fe}$ corresponds to the minimum amount of treatment effect heterogeneity (the minimal value of the standard deviation of the treatment effect across the treated groups and time periods) under which $\hat{\beta}_1$ and the true average treatment effect on the treated (ATT) could be of opposite signs. The second measure $\underline{\underline{\hat{\sigma}}}_{fe}$ corresponds to the minimum amount of treatment effect heterogeneity (the minimal value of the standard deviation of the treatment effect across the treated groups and time periods) under which $\hat{\beta}_1$ could be of a different sign than the treatment effect in all the treated group and time periods. Note that the

[30] I use the `twowayfeweights` Stata package developed by the authors to accompany the work in de Chaisemartin and D'Haultfœuille (2020b), and is available from the STATA repository.

second measure is defined only if there are any negative weights. It is reassuring that in these regressions, $\underline{\hat{\sigma}}_{fe}$-s are larger than $\hat{\beta}_1$-s, which implies that to make the estimated $\hat{\beta}_1$-s invalid, we would need large treatment effect heterogeneity.

Table C4: Negative weights

	Bread healthy %	Milk fat %	Bread log. quantity	Milk log. quantity	Fruits&veg. log. quantity	Juice log. quantity	Eggs log. quantity	Cereals log. quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
β_1: Reformed WIC	6.848***	-0.139***	-0.019	0.001	-0.040	0.165	-0.072	0.075
	(2.196)	(0.053)	(0.050)	(0.054)	(0.086)	(0.205)	(0.094)	(0.088)
Year-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
WIC households	301	301	347	347	347	347	347	347
Households	85421	80721	96637	96637	96637	96637	96637	96637
Household-quarters	1241128	1189478	1598440	1598440	1598440	1598440	1598440	1598440
Sum of negative weights	.	-0.022	.	.	.	.	.	.
$\underline{\hat{\sigma}}_{fe}$	14.504	0.230	0.041	0.003	0.086	0.356	0.154	0.161
$\underline{\underline{\hat{\sigma}}}_{fe}$	.	1.403	.	.	.	.	.	.

Note: Table presents the estimates from the same regression as in the main specification (table 3) except that it excludes time periods when the household that received post-reform vouchers is not receiving WIC vouchers anymore. Standard errors (in parentheses) are clustered at the household level. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level. The bottom part of table presents the sum of negative weights and other diagnostic measures calculated using the `twowayfeweights` Stata package, developed by the authors to accompany the work in de Chaisemartin and D'Haultfœuille (2020b), and is available from the STATA repository.

Second, I calculate the heterogeneity-robust difference-in-differences estimators proposed by de Chaisemartin and D'Haultfœuille (2020a,b) of the short-term effects of the post-reform subsidies.[31] The estimators are valid even when the treatment effect is heterogeneous over time and across groups. The analysis uses yearly data (instead of quarterly) to make it easier to bootstrap standard errors. The estimates in table C5 confirm the main results on the healthy bread and milk fat percentage. Figures C3–C4 allow further assessing the plausibility of the common trends assumption by presenting both the heterogeneity-robust difference-in-differences estimators and placebo estimators for two years before treatment.

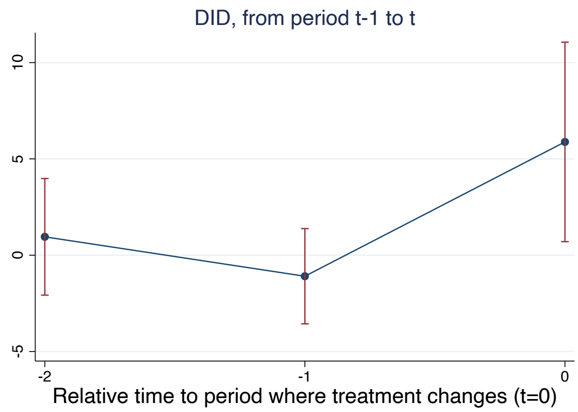
[31] I use the `did_multiplegt` Stata package developed by the authors to accompany the work in de Chaisemartin and D'Haultfœuille (2020b,a), and is available from the STATA repository.

For all products, there are no differences in pre-trends, which provides additional support for the identification strategy.

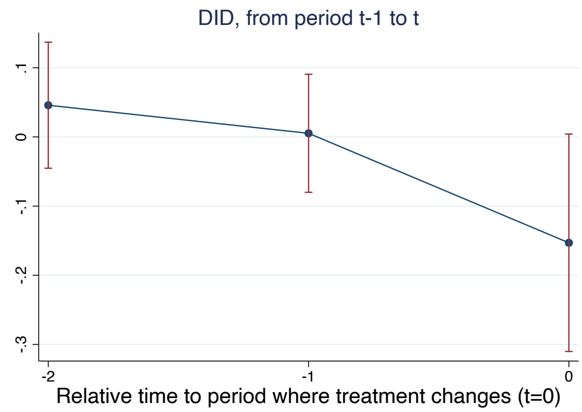
Table C5: Heterogeneity-robust difference-in-differences estimates

	Bread healthy %	Milk fat %	Bread log. quantity	Milk log. quantity	Fruits&veg. log. quantity	Juice log. quantity	Eggs log. quantity	Cereals log. quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reformed WIC	6.405***	-0.128**	0.078	0.020	0.029	-0.106	-0.109	0.135*
	(2.168)	(0.057)	(0.058)	(0.053)	(0.063)	(0.229)	(0.090)	(0.073)
Treated households	174	176	202	202	202	202	202	202
WIC households	301	301	347	347	347	347	347	347
Households	58573	54830	67124	67124	67124	67124	67124	67124

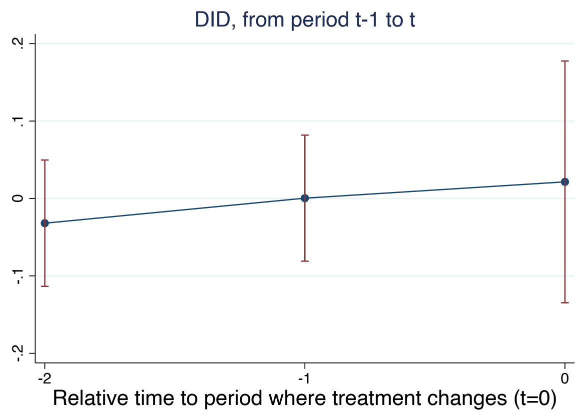
Note: The reported coefficients are heterogeneity-robust difference-in-differences estimates of average dynamic treatment effects. The coefficients are estimated using the method developed by de Chaisemartin and D'Haultfœuille (2020a) and their `did_multiplegt` Stata package, which is available from the STATA repository. Observations are at the household and year pair level. The sample includes all WIC households (columns 1–2 in table 1). It also includes households in column 3 in table 1 who are in the sample for at least 2 years. The sample excludes time periods when a household that received post-reform vouchers is not receiving vouchers anymore. For milk fat percentage, the sample excludes observations in the treatment group, when a household has a one year old child because for one-year-old children, the reformed WIC food vouchers make an exception and provide whole milk. Standard errors (in parentheses) obtained by 50 bootstrap replications are clustered at the household level. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.



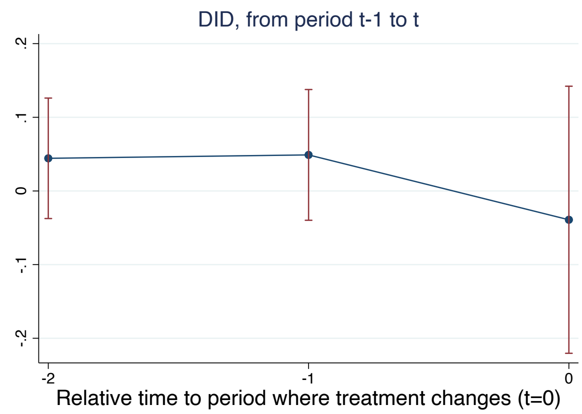
(a) Bread healthy %



(b) Milk fat %



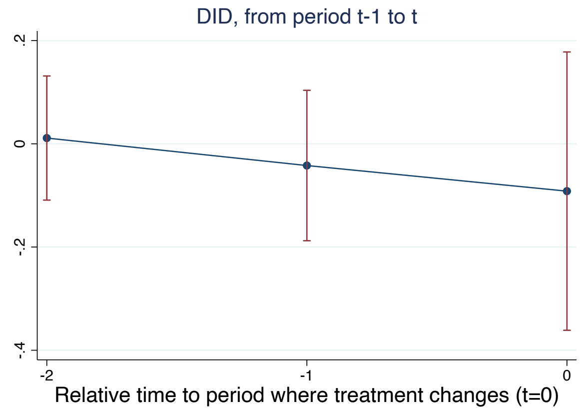
(c) Log. quantity of bread



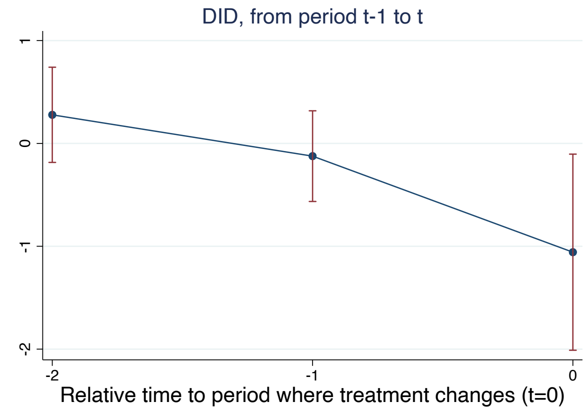
(d) Log. quantity of milk

Figure C3: Heterogeneity-robust difference-in-differences estimates and placebo, bread and milk

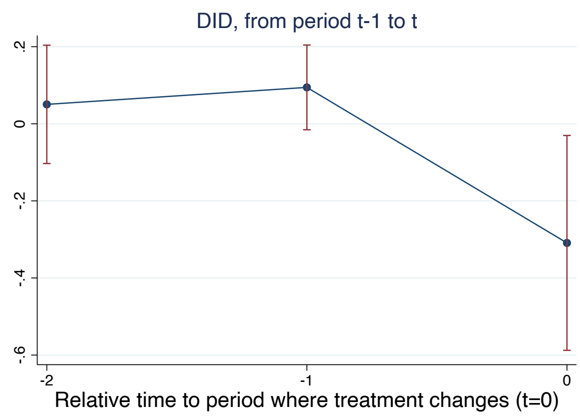
Note: Figures present heterogeneity-robust difference-in-differences estimates of the treatment effect and placebos. The coefficients are estimated using the method developed by de Chaisemartin and D'Haultfœuille (2020b) and their `did_multiplegt` Stata package, which is available from the STATA repository. Observations are at the household and year pair level. The sample includes WIC households (columns 1–2 in table 1) who are in the dataset at least two years before receiving WIC vouchers. From the households receiving post-reform vouchers (column 1 in table 1), the sample excludes those who started to receive WIC vouchers before 2010. The sample also includes households in column 3 in table 1 who are in the sample for at least three years. For milk fat percentage, the sample excludes observations in the treatment group, when a household has a one year old child because for one-year-old children, the reformed WIC food vouchers make an exception and provide whole milk. The figures present point estimates and 95% confidence intervals using standard errors, clustered at the household level, obtained by 50 bootstrap replications.



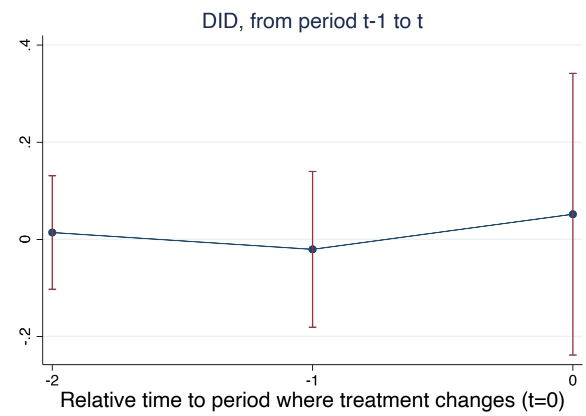
(a) Log. quantity of fruits and vegetables



(b) Log. quantity of juice



(c) Log. quantity of eggs



(d) Log. quantity of cereals

Figure C4: Heterogeneity-robust difference-in-differences estimates and placebo, other products

Note: Figures present heterogeneity-robust difference-in-differences estimates of the treatment effect and placebos. The coefficients are estimated using the method developed by de Chaisemartin and D'Haultfœuille (2020b) and their `did_multiplegt` Stata package, which is available from the STATA repository. Observations are at the household and year pair level. The sample includes WIC households (columns 1–2 in table 1) who are in the dataset at least two years before receiving WIC vouchers. From the households receiving post-reform vouchers (column 1 in table 1), the sample excludes those who started to receive WIC vouchers before 2010. The sample also includes households in column 3 in table 1 who are in the sample for at least three years. The figures present point estimates and 95% confidence intervals using standard errors, clustered at the household level, obtained by 50 bootstrap replications.

C.3.4 Propensity score matching estimators

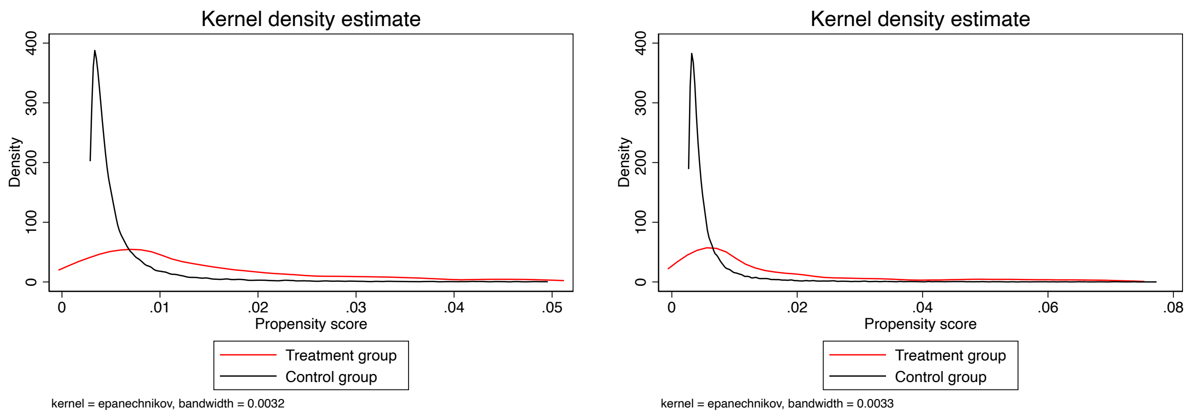
This appendix presents propensity score matching estimators. To obtain the estimators, first, I collapse the data to two separate cross-sections: one by calculating the differences in purchases before receiving and while receiving the WIC vouchers, and the other by calculating the differences in purchases before receiving and in year 1–2 after receiving the WIC vouchers. For the share of healthy bread and percentage of low-fat milk, it is a simple difference in the averages. For the quantities, it is the difference between the logarithms of averages ($\log(y_t/y_{t-1})$), where y_t is the average quantity purchased while or after receiving WIC vouchers and y_{t-1} is the average quantity before receiving WIC vouchers. For each of the households in the control group who don't receive WIC vouchers, I draw a random time period when they started to receive WIC vouchers and another random time period when they end receiving the vouchers. Table C6 presents the estimates of propensity score from a logit model that uses household demographic characteristics to predict WIC participation after the policy reform. The regressions include indicators for children up to four years old at the time when households starts to receive the vouchers (in column 1) or after having received the vouchers (in column 2). I don't use the information on children before starting to receive the vouchers because it would be less informative as then most households don't have children. Figure C5 presents the kernel density estimates of the distribution of the propensity scores separately for the treatment and control group. The sample on the figure is a trimmed sample, which excludes the households with propensity scores smaller than the 10th percentile and larger than the 90th percentile of the treatment group. In the second step, I use these predicted probabilities to estimate the average treatment effect on the treated using propensity-score matching based on 10 nearest neighbors.

Table C7, columns 2–4 and 5–6 present the propensity score matching estimates of the average treatment effect on the treated. For comparison, columns 1 and 4 present estimates from a simple cross-sectional OLS regression. Columns 1–3 present estimates of the impact

Table C6: Predicting WIC participation after the policy reform, logit

	(1)	(2)
Household size	0.374***	0.382***
	(0.045)	(0.043)
Log. income	-0.446***	-0.430***
	(0.102)	(0.102)
College	-0.442***	-0.416***
	(0.142)	(0.142)
Non-white	0.232	0.217
	(0.146)	(0.146)
N	41256	41256

Note: A unit of observation is a household. The sample includes households in columns 1–3 in table 1. The regressions also include indicators for children up to 4 years old. Standard errors are in parentheses. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.



(a) WIC

(b) After WIC

Figure C5: Distribution of propensity scores

of the reformed WIC and columns 4–6 the long-term impact of 1–2 years after the reformed WIC. In columns 3 and 4, the sample is trimmed by excluding households with small and large values of the propensity score (below and above the 10th percentile of the treatment group). The results for the healthy bread share and milk fat percentage are similar to those in the main analysis. However, for juice and eggs, the estimates imply that in the short-term the reformed WIC decreases purchases. This reflects the reformed program decreasing the quantity of juice and eggs in the vouchers. The estimates with the trimmed sample imply that the reformed WIC in the short-term decreases also bread and cereals purchases. In the main specification, the estimates for the product quantities were noisy and did not allow to rule out negative effects of similar size.

Table C7: Robustness of the short- and long-term impact of post-reform WIC, propensity score matching

Dependent variable	Reformed WIC			After Reformed WIC Years 1-2		
	OLS	PSM	PSM Trimmed	OLS	PSM	PSM Trimmed
	(1)	(2)	(3)	(4)	(5)	(6)
Bread, healthy %	5.104***	5.463***	4.640***	4.296***	4.380***	2.934*
	(1.116)	(1.560)	(1.706)	(1.324)	(1.469)	(1.598)
N	34925	34925	24262	34420	34420	26109
Milk, fat %	-0.095***	-0.099**	-0.076*	-0.066*	-0.055	-0.065
	(0.027)	(0.039)	(0.043)	(0.037)	(0.038)	(0.042)
N	32674	32674	22770	32285	32285	24549
Bread, log. quantity	-0.058*	-0.059	-0.082*	0.052	0.067	0.041
	(0.034)	(0.041)	(0.047)	(0.044)	(0.045)	(0.050)
N	41256	41256	27661	41256	41256	30318
Milk, log. quantity	0.020	0.032	0.012	0.081**	0.110**	0.067
	(0.029)	(0.041)	(0.044)	(0.040)	(0.047)	(0.051)
N	41256	41256	27661	41256	41256	30318
Fruits & veg., log. quantity	-0.047	-0.030	-0.075	0.094	0.063	0.160***
	(0.058)	(0.059)	(0.071)	(0.076)	(0.059)	(0.062)
N	41256	41256	27661	41256	41256	30318
Juice, log. quantity	-0.289*	-0.314**	-0.438***	0.020	0.129	0.041
	(0.156)	(0.136)	(0.153)	(0.182)	(0.126)	(0.144)
N	41256	41256	27661	41256	41256	30318
Eggs, log. quantity	-0.118**	-0.107*	-0.153**	-0.026	-0.025	0.055
	(0.054)	(0.061)	(0.069)	(0.070)	(0.065)	(0.070)
N	41256	41256	27661	41256	41256	30318
Cereals, log. quantity	-0.061	-0.067	-0.130*	0.102	0.092	0.112
	(0.061)	(0.064)	(0.076)	(0.077)	(0.063)	(0.068)
N	41256	41256	27661	41256	41256	30318

Note: A unit of observation is a household. Columns 1 and 4 present estimates from OLS regressions, the remaining columns present propensity score matching estimates of the average treatment effect on the treated. Dependent variable is the difference before and while receiving WIC vouchers in the share of healthy bread, fat percentage of milk or the difference between the logarithms of quantities ($\log(y_t/y_{t-1})$). Columns 1–3 present estimates of the impact of the reformed WIC and columns 4–6 of 1–2 years after the reformed WIC. The sample includes households in columns 1–3 in table 1. In columns 3 and 4, the sample is trimmed by excluding households with small and large values of the propensity score (below and above the 10th percentile of the treatment group). Standard errors (in parentheses) are adjusted to take into account that the propensity score is estimated. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

C.3.5 Switching out of WIC due to becoming ineligible

Here, I focus on the time period when households become ineligible for the WIC vouchers. Specifically, I focus on households that received WIC vouchers when the child was 4 years old and become ineligible for the vouchers when the child gets 5 years old. I analyze the changes in their purchases when the child is 5 and 6 compared to the base period when the child was 4 and they received the vouchers. Note that here as opposed to the main sample, the comparison with the time period before receiving WIC vouchers cannot be made because we simply don't observe the households for over so many years (especially, the households who became ineligible after receiving only the pre-reform WIC vouchers).

Again, the empirical strategy exploits the 2009 reform of the WIC program that changed the content of food vouchers. I compare two groups of households—those that participated in the program before the policy reform and those who participated after the policy reform. I compare the changes in their purchases when they become ineligible for the program. These two groups of households should be otherwise similar, both become ineligible when their child becomes five years old, but the vouchers were different.

The sample includes only the WIC households that received WIC vouchers when the child was four and who are in the Nielsen panel both one and two years later when the child is five and six years old. It is a balanced panel of households over these three years. I exclude from the sample any time periods before receiving the vouchers. I also exclude from the sample households with a four year old child in 2009, because the reform took place in the middle of 2009, and hence, these households only for a short time period received the post-reform WIC vouchers. As in the main specification, as a control group the households from column 3 in table 1 are also included.

Using a yearly household panel I estimate the following difference-in-differences regression:

$$
\begin{aligned}
Y_{it} &= \alpha_{2,5} Age5_{it} \times ReceivedWIC_i + \alpha_{2,6} Age6_{it} \times ReceivedWIC_i \\
&= \beta_{2,5} Age5_{it} \times ReceivedWIC_i \times Reformed_i + \beta_{2,6} Age6_{it} \times ReceivedWIC_i \times Reformed_i \\
&+ \alpha_0 Age3^-_{it} \times ReceivedWIC_i + \alpha_3 Age7^+_{it} \times ReceivedWIC_i \\
&+ \beta_0 Age3^-_{it} \times ReceivedWIC_i \times Reformed_i + \beta_3 Age7^+_{it} \times ReceivedWIC_i \times Reformed_i \\
&+ X_{it}\zeta + \delta_i + \gamma_t + \varepsilon_{it} \qquad (6)
\end{aligned}
$$

where $ReceivedWIC_i$ indicates whether household i received WIC vouchers when the child was four years old and $Reformed_i$ indicates whether the vouchers received were post-reform WIC vouchers. Variables $Age5_{it}$ and $Age6_{it}$ are indicators for the child being 5 and 6 years old.[32] The omitted age (the base group) is 4 years old. Variables $ReceivedWIC_i$ and $Reformed_i$ don't vary with time, and hence, these are captured by the household fixed effects. Therefore, we only estimate their interactions with child age.

A few WIC households are in the Nielsen sample also in years before the child is four and after the child is 6. In order to increase statistical power, I don't drop the time periods. Instead, I include separate dummies for these time periods $\alpha_0, \alpha_3, \beta_0, \beta_3$, but I don't report these estimates because due to the unbalanced sample these are biased estimates for these effects, and hence cannot be interpreted.

The regressions include time-varying household characteristics X_{it}: logarithm of income, household size, and indicator variables for the period before the child is born, and children aged 0, 1, ..., 5, 6-12, and 13-17. The regression also includes household fixed effects δ_i and year fixed effects γ_t.

The coefficients of interest are $\beta_{2,5}$ and $\beta_{2,6}$ which estimate the impact of the WIC reform

[32]The dataset has only yearly information about age. I say that the child is T is years old in the calendar year of his/her Tth birthday.

on the change in purchases when the WIC household becomes ineligible. Note that the household becomes ineligible during the year when the child gets 5 years old. Hence, the household is eligible for only part of the calendar year when the child has the 5th birthday.[33] Hence, if the long-term effect of WIC is decreasing over time (as in the main specification table 3) we would expect $\beta_{2,6}$ to be larger than $\beta_{2,5}$ in absolute value.

The results in table C8 show that indeed, for the healthy bread and milk fat percentage, the long-term effect ($\beta_{2,5}$ and $\beta_{2,6}$) estimates are in the same direction than in the main specification (β_2 in table 3). This shows that the limited persistence of the program's impact is not specific to the households in the main specification, where some households might voluntarily choose to exit the WIC program. But instead, the limited persistence characterizes also the households who exit the program when they become ineligible due to the child's age.

It is somewhat surprising that for the healthy bread and milk fat percentage the magnitude of $\beta_{2,6}$ coefficient estimate in table C8 is much larger than β_2 in table 3. If the short-term effect were similar in both samples, then table C8 would imply that the post-reform WIC in the long-term has the opposite of the intended effect. But the magnitudes of $\beta_{2,6}$ and β_2 cannot directly be compared because the samples are different. In table 3, majority of the households were exiting the program when the child was much younger than five. It is also likely that they were receiving the vouchers for a shorter period of time. Unfortunately, when focusing on households exiting the program due to ineligibility as child gets 5 years old, due to the data limitations we cannot compare the purchases to the time period before receiving WIC vouchers. Hence, $\beta_{2,6}$ in table C8 does not measure the long-term effect of the program, it only measures the change compared to the year when the child was 4 and the household was receiving the vouchers.

[33] The majority of the households that received the vouchers when the child was 4, do not report receiving vouchers when the child is 5. This could reflect that they report the WIC status in the end of the calendar year.

Table C8: The impact of post-reform WIC on subsidized products; switching out of WIC due to becoming ineligibile when the child gets 5 years old

	Bread healthy %	Milk fat %	Bread log. quantity	Milk log. quantity	Fruits&veg. log. quantity	Juice log. quantity	Eggs log. quantity	Cereals log. quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\beta_{2,5}$: Age 5 × Received WIC × Reformed	-4.118*	0.137***	-0.076	0.072	-0.021	-0.062	-0.058	0.154
	(2.472)	(0.052)	(0.088)	(0.098)	(0.086)	(0.199)	(0.182)	(0.123)
$\beta_{2,6}$: Age 6 × Received WIC × Reformed	-11.557***	0.250***	0.058	0.215	-0.083	0.162	-0.299	-0.014
	(3.091)	(0.075)	(0.147)	(0.144)	(0.104)	(0.368)	(0.206)	(0.159)
$\alpha_{2,5}$: Age 5 × Received WIC	-0.606	-0.035	0.005	-0.019	0.001	-0.131	0.025	-0.110
	(1.349)	(0.032)	(0.075)	(0.077)	(0.070)	(0.102)	(0.164)	(0.103)
$\alpha_{2,6}$: Age 6 × Received WIC	0.381	-0.030	-0.085	-0.102	0.041	-0.783***	0.166	-0.085
	(1.743)	(0.045)	(0.122)	(0.113)	(0.081)	(0.283)	(0.166)	(0.118)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
WIC households	159	156	173	173	173	173	173	173
Households	85279	80576	96463	96463	96463	96463	96463	96463
Household-years	339364	316299	398786	398786	398786	398786	398786	398786

Note: Each column presents estimates of a separate panel data fixed effects regression (equation (6)) that measure the changes in purchases when the child is 5 and 6 compared to the base period when the child was 4 and the household was receiving the vouchers. A unit of observation is a household-year pair. Bread healthy (whole wheat) % and milk fat % are in the range of 0–100, quantity is measured in pounds for bread, gallons for milk, kilocalories for fruits and vegetables and juice, counted for eggs, and ounces for cereals. All regressions include household size, logarithm of income, dummies for the year before child is born, and children aged $0, 1, \ldots, 5$, 6-12, and 13-17. Regressions include household and time period fixed effects. Standard errors (in parentheses) are clustered at the household level. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

C.4 Excluding pre-reform WIC households from the sample

When households start to receive WIC vouchers, there are other changes that could affect purchases. Some of which are observed, such as a new child, but some may be unobservable. Comparing households that enter into old and new WIC programs allows to account for these changes. If instead, we were to exclude from the sample the households in the old WIC program, the estimates would measure the combined effect of the new vouchers and everything else that changes when households receive the vouchers.

This appendix analyzes the differences from the main specification if the pre-reform WIC households had been excluded from the control group. Specifically, the sample includes only the WIC households who started receiving vouchers in 2010 or later; and the control group from column 3 of table 1. Using this sample, I estimate the following difference-in-differences regression, which is a simplified version of the main specification given by equation (1):

$$\begin{aligned} Y_{it} = &\tilde{\beta}_1 WIC_{it} + \tilde{\beta}_2 AfterWICYears1,2_{it} + \tilde{\beta}_3 AfterWICYear3^{+}_{it} \\ &+ X_{it}\zeta + \delta_i + \gamma_t + \varepsilon_{it} \end{aligned} \tag{7}$$

The results in table C9 illustrate that without the pre-reform WIC control group, the estimated effects of the WIC vouchers ($\tilde{\beta}$) are the combinations of α and β estimates from table 3. For example, the estimated short-term effect of subsidies on the healthy bread share is smaller (equals 4.2) than β_1 estimate (which equals 6.9) in table 3, because it captures also the negative α_1 estimate (which equals -1.8). The estimates of β_1 and β_2 in table 3 measure the incremental effect of the new voucher program relative to the old. While the estimates in table C9 measure the effect of the new vouchers and anything else that tends to change when WIC is adopted.

Table C9: The impact of post-reform WIC on subsidized products, the sample excludes households who received pre-reform WIC vouchers

	Bread healthy %	Milk fat %	Bread log. quantity	Milk log. quantity	Fruits&veg. log. quantity	Juice log. quantity	Eggs log. quantity	Cereals log. quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\tilde{\beta}_1$: WIC	4.193**	-0.103*	0.023	0.040	-0.009	0.639***	-0.030	0.127
	(1.847)	(0.053)	(0.050)	(0.050)	(0.078)	(0.189)	(0.085)	(0.080)
$\tilde{\beta}_2$: After WIC Years 1-2	3.271*	-0.064	0.009	-0.001	-0.055	0.198	-0.097	0.045
	(1.762)	(0.048)	(0.056)	(0.053)	(0.089)	(0.186)	(0.084)	(0.084)
Year-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald test, $\tilde{\beta}_1 = \tilde{\beta}_2$, p-value	0.588	0.338	0.775	0.384	0.610	0.025	0.427	0.278
WIC households	149	154	175	175	175	175	175	175
Households	85269	80574	96465	96465	96465	96465	96465	96465
Household-quarters	1239541	1188115	1596596	1596596	1596596	1596596	1596596	1596596

Note: Each column presents estimates from a separate panel data fixed effects regression (equation (7)). A unit of observation is a household-quarter pair. Bread healthy (whole wheat) % and milk fat % are in the range of 0–100, quantity is measured in pounds for bread, gallons for milk, kilocalories for fruits and vegetables and juice, counted for eggs, and ounces for cereals. All regressions include household size, logarithm of income, dummies for period before child is born, and children aged $0, 1, \ldots, , 5$, 6-12, and 13-17. In column 2, regression includes interaction terms for WIC and the last two quarters of the calendar year when the child gets one years old and the other with the first two quarters of the calendar year when the child gets two years old. Regressions include household and time period fixed effects. Standard errors (in parentheses) are clustered at the household level. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

C.5 Map describing regional differences in milk prices

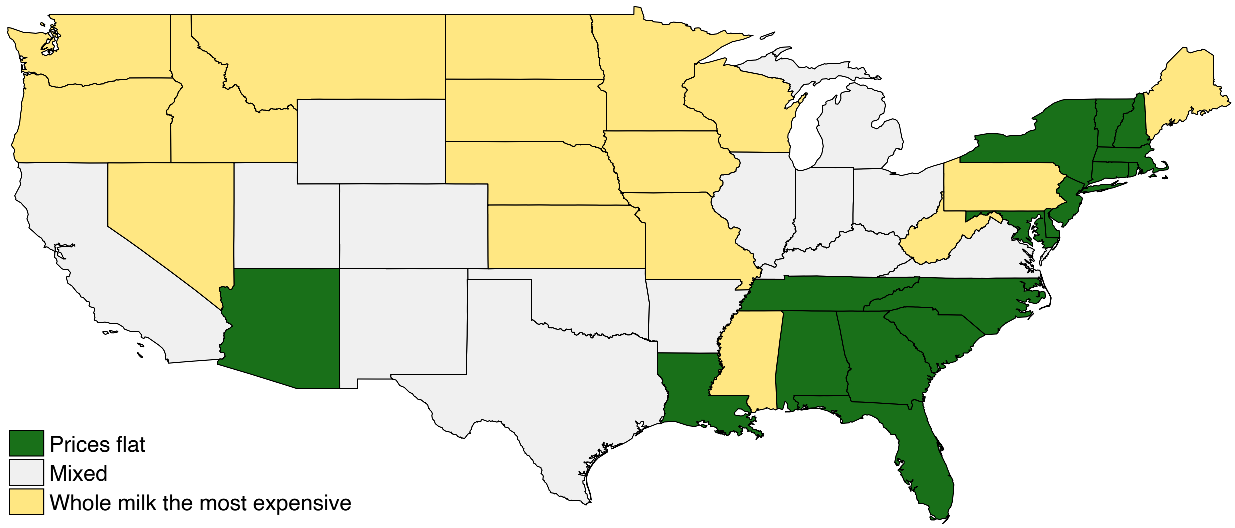


Figure C6: Milk prices

Note: States are categorized to those with flat prices where the price of whole milk is not higher than others, those where whole milk is more expensive than others (more than a cent more expensive), and those with mixed price rankings. States are categorized based on state-level median milk price differences of different fat percentages (whole milk, 2%, 1%, and skim milk) across stores and time periods using data on private label one-gallon milk.

C.6 Analysis of prices and persistence based only on WIC households

Table C10: The short- and long-term impact and prices. Sample restricted to WIC households. Dependent variable: milk fat percentage

	Prices flat (1)	Prices increasing (2)
β_1: Reformed WIC	-0.288***	-0.315***
	(0.100)	(0.105)
β_2: After Reformed WIC Years 1-2	0.074	-0.310
	(0.115)	(0.215)
α_1: WIC	0.099	0.203***
	(0.073)	(0.070)
α_2: After WIC Years 1-2	-0.047	0.237
	(0.103)	(0.189)
σ: Post-Reform	0.079	0.023
	(0.065)	(0.093)
Household FE	Yes	Yes
Wald test, $\beta_1 = \beta_2$, p-value	0.003	0.976
WIC households	94	60
Households	94	60
Household-quarters	2802	1831

Note: The table presents estimates from the same regressions as in table 4, but the sample is restricted to households who received WIC vouchers (households in columns 1–2 in table 1) and the regressions do not include year and quarter fixed effects. Standard errors (in parentheses) are clustered at the household level. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

C.7 Externalities to other products

The online appendix analyzes the externalities. It estimates regression (1) to measure the impact of the program on fiber consumption first in bread which is included in WIC vouchers and then in product categories not included in the WIC vouchers. Similarly, it estimates the impact on saturated fat consumption first in milk and then in product categories not included in the WIC vouchers. To increase precision and make the estimates comparable, the outcome variables are in the form of percentile rank. The estimates have a simple interpretation: how much the reform shifted household's position in the national distribution of the nutrition measure.

The change in the WIC vouchers in the short-term increased household's rank of fiber content in bread by about 4 percentiles (table C11, column 1). It is a meaningful increase compared to the 41% pre-WIC level of the median household that later received the post-reform WIC vouchers. For both non-WIC grain products[34] (column 2) and all non-WIC product categories (column 3), the estimates of the short-term effect are imprecise, but allow to rule out with 95% probability positive and negative externalities of the same magnitude as the main effect (column 1).

The change in the WIC vouchers in the short-term decreased household's rank of saturated fat content in milk by about 3 percentiles (column 4). This is a meaningful decrease compared to the 64% pre-WIC level of the median household that later received the post-reform WIC vouchers. For non-WIC dairy products[35] (column 5), the estimates of the short-term effect allow to rule out positive and negative externalities of the same magnitude as the main effect (column 4). For all non-WIC product categories (column 6), the point estimate of the short-term effect is small in magnitude and the estimates allow to rule out large externalities.

[34] Non-WIC grain products include other bread and baked goods, flour, baking mixes, pasta, crackers, rice, barley, tortillas, rice cakes, and pretzels.

[35] Non-WIC dairy products include butter and margarine, cheese, cottage cheese, sour cream, yogurt, dairy desserts, cream, buttermilk, flavored milk.

Table C11: Externalities: The impact of post-reform WIC on fiber consumption in bread, non-WIC grains, and all non-WIC product categories, and on saturated fat consumption in milk, non-WIC dairy, and all non-WIC product categories

	Fiber			Saturated fat		
	Bread	Non-WIC grains	All non-WIC categories	Milk	Non-WIC dairy	All non-WIC categories
	(1)	(2)	(3)	(4)	(5)	(6)
β_1: Reformed WIC	3.958**	0.600	-0.242	-3.330**	-0.366	-0.675
	(1.903)	(1.481)	(1.761)	(1.459)	(1.451)	(1.907)
	[0.229,7.688]	[-2.303,3.503]	[-3.693,3.209]	[-6.190,-0.469]	[-3.210,2.478]	[-4.412,3.063]
β_2: After Reformed WIC Years 1-2	2.153	0.924	2.932	-1.483	-0.580	-0.814
	(2.089)	(1.554)	(1.930)	(1.922)	(1.741)	(2.802)
	[-1.942,6.249]	[-2.123,3.971]	[-0.851,6.715]	[-5.251,2.284]	[-3.992,2.833]	[-6.306,4.679]
α_1: WIC	-0.163	-1.726	0.830	-0.168	0.000	-0.395
	(1.357)	(1.169)	(1.357)	(1.030)	(1.123)	(1.613)
	[-2.822,2.497]	[-4.017,0.565]	[-1.830,3.490]	[-2.187,1.851]	[-2.201,2.201]	[-3.557,2.767]
α_2: After WIC Years 1-2	0.937	-2.089	-1.126	-0.415	1.759	-0.452
	(1.687)	(1.277)	(1.531)	(1.691)	(1.481)	(2.571)
	[-2.370,4.244]	[-4.592,0.415]	[-4.127,1.875]	[-3.729,2.899]	[-1.144,4.663]	[-5.492,4.588]
σ: Post-Reform	-0.486*	0.078	1.083***	-0.197	0.855***	1.070***
	(0.284)	(0.305)	(0.295)	(0.204)	(0.287)	(0.345)
	[-1.042,0.071]	[-0.519,0.675]	[0.504,1.662]	[-0.597,0.203]	[0.292,1.418]	[0.394,1.746]
Year-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Household FE	Yes	Yes	Yes	Yes	Yes	Yes
Wald test, $\beta_1 = \beta_2$, p-value	0.394	0.851	0.114	0.345	0.912	0.959
WIC households	301	301	301	301	301	301
Households	85421	85404	85421	80721	80688	80721
Household-quarters	1244237	1342951	1362822	1192570	1245769	1270611

Note: Each column presents estimates from a separate panel data fixed effects regression (equation (1)). A unit of observation is a household-quarter pair. Fiber content is measured in grams per thousand calories, and saturated fat content is measured as a percentage of calories from saturated fat. The outcome variable is household's percentile rank of fiber in bread (column 1), non-WIC grains (column 2), and all non-WIC product categories; and households percentile rank of saturated fat in milk (column 4), non-WIC dairy (column 5), and all non-WIC product categories (column 6). To calculate the percentile ranks, in each time period, all the households in the Nielsen Consumer Panel are ranked by fiber or saturated fat content. Household's percentile rank is defined based on the position in the national distribution of fiber or saturated fat content purchases relative to all others in a given time period. The sample includes households in columns 1–3 in table 1. Standard errors (in parentheses) are clustered at the household level. 95% confidence intervals are in the square brackets. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

D Online Appendix: Additional analysis for section 4

D.1 Identification of short-term state dependence

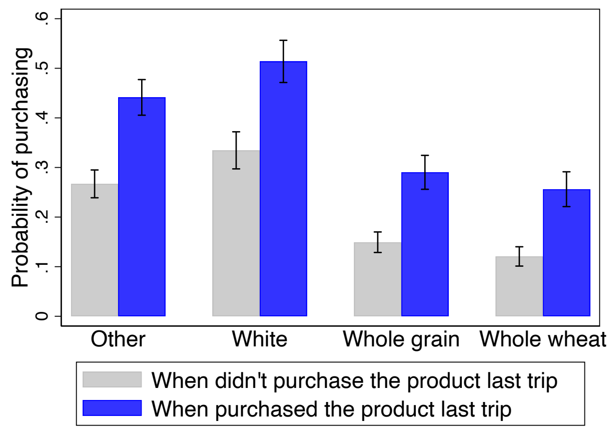


Figure D1: Within-household comparison of the conditional probability of purchasing the product when did or did not buy the same product in the previous trip; presented separately for the four bread types

Note: An observation is a household-product-state combination. Each bar is the average probability over households and the spiked interval presents the 95% confidence interval. When calculating the averages, only the households are included for whom for that product the probability in both states is calculated. When household never buys the product, then it is not possible to calculate the probability in a state of having bought the product in the previous trip. For comparison purposes, in that case, the household is excluded from calculating both conditional probabilities for that product. Therefore, probabilities are calculated for 248 households for Other bread, 228 for White bread, 224 for Whole grain, and 202 for Whole wheat.

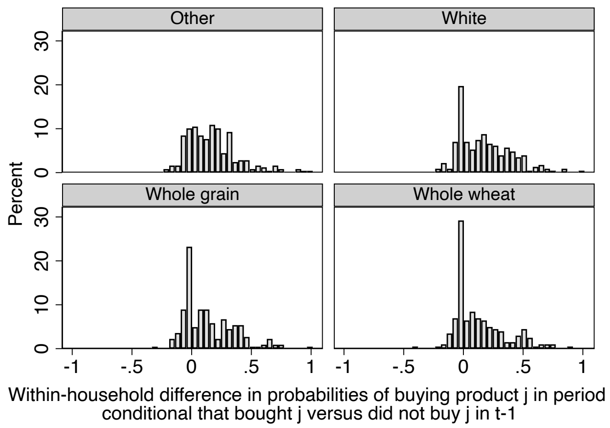


Figure D2: Histogram of within-household differences in the conditional probabilities in the two states: when did or did not buy the same product last trip; presented separately for the four bread types

Note: An observation is a household and product pair. The within-household difference in probabilities is only possible to calculate for households that bought the product in both states (when did or did not buy the same product in the previous trip). Therefore, the differences are calculated for 248 households for Other bread, 228 for White bread, 224 for Whole grain, and 202 for Whole wheat.

D.2 Demand model estimates: main specification

Table D1: Control function estimation. Dependent variable: price of bread

	Estimate	SE
η_p: Average price in other markets	0.885***	(0.044)
Whole grain bread	0.031	(0.022)
Whole wheat bread	0.009	(0.018)
White bread	-0.042	(0.025)
Market FE	Yes	
F-statistic of H0: $\eta_p = 0$	401.391	
R-squared	0.848	
Number of clusters	50	
Number of observations	80288	

Note: A unit of observation is a bread type (white, whole grain, whole wheat, or other), market, (aggregate) retailer, and week combination. Prices are measured in dollars per pound, deflated to 2015 dollars. Standard errors clustered on the market level are included in parenthesis. It should be noted that these are more conservative standard errors, while clustering at the market and bread type pair level would imply more precise estimates. *** Indicates significance at the 1 percent level, ** at 5 percent level, * at 10 percent level.

Table D2: Bread demand model estimates

	No control function		Control function	
	(1)		(2)	
	Estim.	SE	Estim.	SE
Mean				
$\bar{\alpha}$: Price	−1.029***	(0.088)	−1.316***	(0.287)
$\bar{\gamma}$: Short-term state dependence	0.693***	(0.025)	0.614***	(0.067)
β: Last year's share of whole wheat	3.872***	(0.121)	3.784***	(0.254)
$\bar{\theta}$: Whole wheat bread	−1.534***	(0.051)	−1.549***	(0.147)
$\bar{\delta}$: Whole grain bread	−0.271***	(0.059)	−0.606***	(0.195)
$\bar{\lambda}$: White bread	−0.403***	(0.077)	−0.809**	(0.362)
Residual			0.539*	(0.277)
Standard deviation				
σ_p: Price	1.903***	(0.132)	0.670	(1.248)
σ_s: Short-term state dependence	0.430***	(0.031)	0.487	(0.343)
σ_{WW}: Whole wheat bread	1.007***	(0.059)	0.875	(0.619)
σ_{WG}: Whole grain bread	1.511***	(0.093)	1.524**	(0.704)
σ_{White}: White bread	2.273***	(0.057)	1.931***	(0.670)
Covariance				
Price, short-term state dependence	0.458***	(0.079)	−0.169	(0.367)
Price, whole wheat bread	−0.088	(0.106)	−0.187	(0.469)
Price, whole grain bread	−0.612***	(0.177)	−0.676	(1.585)
Price, white bread	1.319***	(0.286)	−0.577	(0.783)
Short-term state dependence, whole wheat bread	−0.023	(0.034)	−0.100	(0.360)
Short-term state dependence, whole grain bread	0.105	(0.065)	0.164	(0.424)
Short-term state dependence, white bread	0.158***	(0.030)	−0.373	(0.689)
Whole wheat bread, whole grain bread	1.364***	(0.167)	0.957	(1.405)
Whole wheat bread, white bread	0.836***	(0.095)	0.686	(1.082)
Whole grain bread, white bread	1.106***	(0.190)	0.760	(1.386)
Log-likelihood	-22843		-22821	
Number of choices	112168		112168	
Number of households	252		252	

Note: The table presents the estimates from 2 random coefficient logit models. For each model, the first column presents parameter estimates and the second column standard errors. The base type of bread is *other* bread. The sample includes households from columns 1–2 in table 1 with a sufficient number of bread purchases; for each household the sample excludes the years when the household received WIC vouchers. Standard errors are clustered at the household level. In column 2, standard errors are obtained by bootstrap (over 2 steps of the estimation) with 100 bootstrap samples. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

D.3 Robustness analysis of demand model estimates

In this appendix, I explore the robustness of demand model estimates first with respect to the construction of the price variable, and then the initial conditions for state dependence.

First, to address the potential concerns about combining the Retail Scanner and the Consumer Panel data to construct prices, I re-estimate the model restricting the sample to the shopping trips to the stores in the Retail Scanner data and hence, using only the prices from the Retail Scanner data (tables D3–D4). Second, we might be concerned that prices from the Consumer Panel data were constructed based on too few households. Furthermore, for the area outside the major scantrack markets (which Nielsen aggregates into the remaining U.S.) prices were calculated by combining all the regions outside the major markets. To address the concerns about the potential measurement errors introduced in this way, I re-estimate the model while restricting the sample only to the Nielsen reportable markets and excluding the remaining US (tables D5–D6).

It is reassuring that the estimates from these models are similar to those from the main specification. Both the short- and long-term persistence parameter ($\bar{\gamma}$ and β) are similar across the specifications. Also the estimated price coefficients are rather similar, yielding the average own-price elasticity with control function estimates to equal 1.6 with only Retail Scanner prices (table D4) and 1.5 with reportable markets and without the remaining U.S. (table D6) compared to 1.6 in the main specification (table D2).

To address the concern about the initial conditions bias for state dependence, I use the method proposed by Simonov et al. (2020) to bound the true short-term state dependence value. The main specification (table 5) assumes that the initial state is exogeneous. Here I re-estimate the model assuming that the first observed purchase occasion has no state dependence (tables D7–D8). To do that I add to the sample the week that was used to calculate the first state for short-term state dependence and I assume that there is no short- or long-term state dependence during that week. Simonov et al. (2020) show that the estimates from these two alternative specifications bound the values of the state dependence. The estimates show that the possible bias is small as the bound for the short-term state dependence is tight: the point estimate of the short-term state dependence with alternative initial conditions is only slightly smaller (0.58 in column 2 table D8) compared to the estimate in the main specification (0.61 in column 2 of table 5).

Table D3: Control function estimation: prices only from the Retail Scanner dataset. Dependent variable: price of bread.

	Estimate	SE
η_p: Average price in other markets	0.884***	(0.024)
Whole grain bread	0.018	(0.022)
Whole wheat bread	0.016	(0.017)
White bread	-0.049**	(0.019)
Market FE	Yes	
F-statistic of H0: $\eta_p = 0$	1324.523	
R-squared	0.869	
Number of clusters	48	
Number of observations	37868	

Note: A unit of observation is a bread type (white, whole grain, whole wheat, or other), market, aggregate retailer, and week combination. Prices are measured as dollars per pound, deflated to 2015 dollars. Standard errors clustered by market are included in parenthesis. *** Indicates significance at the 1 percent level, ** at 5 percent level, * at 10 percent level.

Table D4: Robustness of the bread demand model estimates: prices only from the Retail Scanner dataset

	No control function (1)		Control function (2)	
	Estim.	SE	Estim.	SE
Mean				
$\bar{\alpha}$: Price	$-0.852{*}{*}{*}$	(0.081)	$-1.226{*}{*}{*}$	(0.306)
$\bar{\gamma}$: Short-term state dependence	$0.729{*}{*}{*}$	(0.033)	$0.645{*}{*}{*}$	(0.077)
β: Last year's share of whole wheat	$3.565{*}{*}{*}$	(0.197)	$4.015{*}{*}{*}$	(0.409)
$\bar{\theta}$: Whole wheat bread	$-1.802{*}{*}{*}$	(0.094)	$-1.515{*}{*}{*}$	(0.256)
$\bar{\delta}$: Whole grain bread	$-0.905{*}{*}{*}$	(0.100)	$-0.574{*}{*}$	(0.279)
$\bar{\lambda}$: White bread	$-0.158{*}{*}$	(0.066)	$-0.794{*}{*}$	(0.379)
Residual			0.463	(0.354)
Standard deviation				
σ_p: Price	$0.763{*}{*}{*}$	(0.144)	0.847	(1.231)
σ_s: Short-term state dependence	$0.445{*}{*}{*}$	(0.042)	$0.415{*}$	(0.250)
σ_{WW}: Whole wheat bread	$1.332{*}{*}{*}$	(0.081)	$1.178{*}{*}{*}$	(0.437)
σ_{WG}: Whole grain bread	$1.593{*}{*}{*}$	(0.085)	$1.824{*}{*}{*}$	(0.539)
σ_{White}: White bread	$2.028{*}{*}{*}$	(0.084)	$3.657{*}{*}{*}$	(0.550)
Covariance				
Price, short-term state dependence	$0.334{*}{*}{*}$	(0.064)	0.286	(0.556)
Price, whole wheat bread	$-0.635{*}{*}{*}$	(0.170)	-0.165	(0.669)
Price, whole grain bread	$-0.306{*}{*}$	(0.121)	0.225	(0.542)
Price, white bread	$-0.305{*}{*}{*}$	(0.053)	0.435	(1.079)
Short-term state dependence, whole wheat bread	$-0.404{*}{*}{*}$	(0.055)	-0.111	(0.293)
Short-term state dependence, whole grain bread	$-0.251{*}{*}{*}$	(0.053)	-0.218	(0.481)
Short-term state dependence, white bread	-0.076	(0.058)	0.539	(0.484)
Whole wheat bread, whole grain bread	$1.432{*}{*}{*}$	(0.164)	1.632	(1.280)
Whole wheat bread, white bread	0.048	(0.120)	2.265	(1.643)
Whole grain bread, white bread	$0.186{*}$	(0.109)	1.376	(2.175)
Log-likelihood	-9211		-9229	
Number of choices	43560		43560	
Number of households	218		218	

Note: The table presents the estimates from 2 random coefficient logit models. For each model, the first column presents parameter estimates and the second column standard errors. The base type of bread is *other* bread. The sample includes households from columns 1–2 in table 1 with a sufficient number of bread purchases; for each household the sample excludes the years when the household received WIC vouchers. The sample is restricted to the shopping trips to the stores in the Retail Scanner data. The prices are constructed only from the Retail Scanner data. Standard errors are clustered at the household level. In column 2, standard errors are obtained by bootstrap (over 2 steps of the estimation) with 100 bootstrap samples. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

Table D5: Control function estimation: sample includes only the major reportable markets and excludes the remaining U.S. Dependent variable: price of bread.

	Estimate	SE
η_p: Average price in other markets	0.865***	(0.050)
Whole grain bread	0.047**	(0.022)
Whole wheat bread	0.012	(0.020)
White bread	-0.050*	(0.029)
Market FE	Yes	
F-statistic of H0: $\eta_p = 0$	301.060	
R-squared	0.842	
Number of clusters	46	
Number of observations	67592	

Note: A unit of observation is a bread type (white, whole grain, whole wheat, or other), market, (aggregate) retailer, and week combination. Prices are measured as dollars per pound, deflated to 2015 dollars. Standard errors clustered by market are included in parenthesis. *** Indicates significance at the 1 percent level, ** at 5 percent level, * at 10 percent level.

Table D6: Robustness of the bread demand model estimates: sample includes only the major reportable markets and excludes the remaining U.S.

	No control function (1)		Control function (2)	
	Estim.	SE	Estim.	SE
Mean				
$\bar{\alpha}$: Price	-1.122***	(0.080)	-1.057***	(0.337)
$\bar{\gamma}$: Short-term state dependence	0.614***	(0.025)	0.556***	(0.063)
β: Last year's share of whole wheat	4.132***	(0.119)	3.977***	(0.266)
$\bar{\theta}$: Whole wheat bread	-1.296***	(0.057)	-1.376***	(0.183)
$\bar{\delta}$: Whole grain bread	-0.256***	(0.076)	-0.378	(0.240)
$\bar{\lambda}$: White bread	-0.287***	(0.063)	-0.123	(0.383)
Residual			0.621**	(0.299)
Standard deviation				
σ_p: Price	0.463***	(0.140)	1.440	(1.555)
σ_s: Short-term state dependence	0.471***	(0.027)	0.523*	(0.296)
σ_{WW}: Whole wheat bread	0.848***	(0.041)	0.941*	(0.571)
σ_{WG}: Whole grain bread	1.460***	(0.053)	1.550**	(0.740)
σ_{White}: White bread	1.988***	(0.056)	3.370***	(0.628)
Covariance				
Price, short-term state dependence	-0.178***	(0.055)	0.099	(0.350)
Price, whole wheat bread	0.074***	(0.026)	-0.387	(0.860)
Price, whole grain bread	-0.081*	(0.046)	-1.416	(2.444)
Price, white bread	-0.058***	(0.020)	2.811	(4.019)
Short-term state dependence, whole wheat bread	-0.115***	(0.025)	-0.293	(0.368)
Short-term state dependence, whole grain bread	0.232***	(0.030)	-0.081	(0.403)
Short-term state dependence, white bread	-0.239***	(0.047)	-0.937	(0.837)
Whole wheat bread, whole grain bread	0.827***	(0.085)	1.150	(1.364)
Whole wheat bread, white bread	0.958***	(0.084)	1.110	(1.794)
Whole grain bread, white bread	0.263***	(0.098)	-0.911	(2.819)
Log-likelihood	-17826		-17748	
Number of choices	83272		83272	
Number of households	198		198	

Note: The table presents the estimates from 2 random coefficient logit models. For each model, the first column presents parameter estimates and the second column standard errors. The base type of bread is *other* bread. The sample includes households from columns 1–2 in table 1 with a sufficient number of bread purchases; for each household the sample excludes the years when the household received WIC vouchers. The sample is further restricted to include households only in the major reportable markets and exclude the remaining U.S. Also the instruments (prices in other markets) are constructed only based on those markets in the sample. Standard errors are clustered at the household level. In column 2, standard errors are obtained by bootstrap (over 2 steps of the estimation) with 100 bootstrap samples. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

Table D7: Control function estimation: alternative initial conditions for state dependence. Dependent variable: price of bread.

	Estimate	SE
η_p: Average price in other markets	0.885***	(0.044)
Whole grain bread	0.032	(0.022)
Whole wheat bread	0.010	(0.017)
White bread	-0.042	(0.025)
Market FE	Yes	
F-statistic of H0: $\eta_p = 0$	402.1	
R-squared	0.848	
Number of clusters	50	
Number of observations	80928	

Note: A unit of observation is a bread type (white, whole grain, whole wheat, or other), market, aggregate retailer, and week combination. Prices are measured as dollars per pound, deflated to 2015 dollars. Standard errors clustered by market are included in parenthesis. *** Indicates significance at the 1 percent level, ** at 5 percent level, * at 10 percent level.

Table D8: Robustness of the bread demand model estimates: alternative initial conditions for state dependence

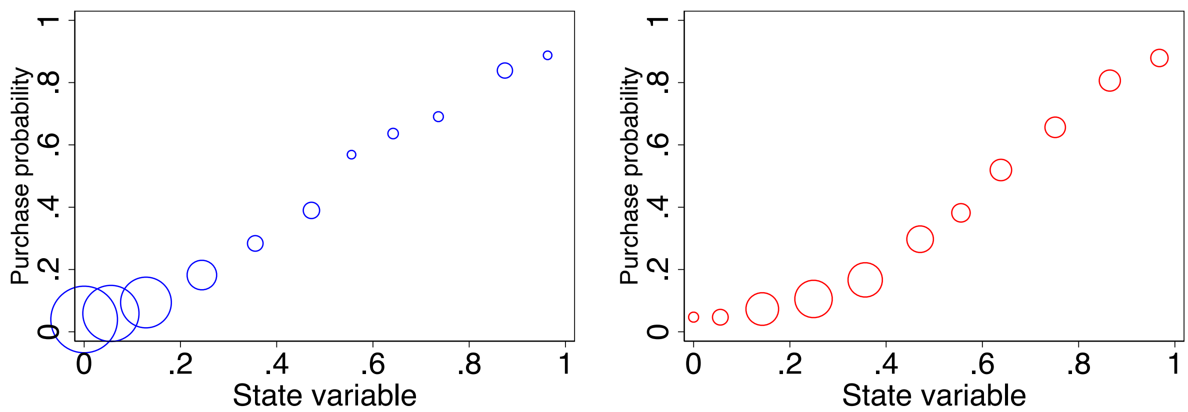
	No control function (1)		Control function (2)	
	Estim.	SE	Estim.	SE
Mean				
$\bar{\alpha}$: Price	−1.208***	(0.073)	−1.338***	(0.279)
$\bar{\gamma}$: Short-term state dependence	0.651***	(0.024)	0.577***	(0.060)
β: Last year's share of whole wheat	3.900***	(0.109)	4.049***	(0.263)
$\bar{\theta}$: Whole wheat bread	−1.683***	(0.048)	−1.387***	(0.139)
$\bar{\delta}$: Whole grain bread	−0.583***	(0.047)	−0.403**	(0.192)
$\bar{\lambda}$: White bread	−0.841***	(0.048)	−0.407	(0.350)
Residual			0.603**	(0.259)
Standard deviation				
σ_p: Price	1.260***	(0.081)	0.098	(1.327)
σ_s: Short-term state dependence	0.467***	(0.035)	0.521**	(0.247)
σ_{WW}: Whole wheat bread	0.830***	(0.037)	0.968*	(0.554)
σ_{WG}: Whole grain bread	1.283***	(0.040)	1.340*	(0.795)
σ_{White}: White bread	2.100***	(0.053)	2.184***	(0.675)
Covariance				
Price, short-term state dependence	0.135**	(0.060)	−0.051	(0.690)
Price, whole wheat bread	−0.221***	(0.060)	0.062	(0.828)
Price, whole grain bread	−0.605***	(0.075)	0.030	(0.358)
Price, white bread	1.141***	(0.137)	0.126	(1.760)
Short-term state dependence, whole wheat bread	−0.221***	(0.029)	−0.345	(0.432)
Short-term state dependence, whole grain bread	0.042	(0.031)	−0.189	(0.616)
Short-term state dependence, white bread	−0.085*	(0.045)	−0.680	(0.554)
Whole wheat bread, whole grain bread	0.774***	(0.065)	1.048	(1.408)
Whole wheat bread, white bread	−0.120**	(0.058)	1.320	(1.125)
Whole grain bread, white bread	−0.719***	(0.061)	0.998	(1.446)
Log-likelihood	-22974		-23066	
Number of choices	113208		113208	
Number of households	252		252	

Note: The table presents the estimates from 2 random coefficient logit models. For each model, the first column presents parameter estimates and the second column standard errors. The base type of bread is *other* bread. The sample includes households from columns 1–2 in table 1 with a sufficient number of bread purchases; for each household the sample excludes the years when the household received WIC vouchers. Standard errors are clustered at the household level. In column 2, standard errors are obtained by bootstrap (over 2 steps of the estimation) with 100 bootstrap samples. *** Indicates significance at the 1 percent level, ** 5 percent level, * 10 percent level.

D.4 Additional analysis for counterfactuals

To illustrate the role of the state variable κ_{WW} that is the share of whole wheat bread in past year's bread purchases, figure D3 presents the joint counterfactual distribution of the state variable and the probability of purchasing whole wheat bread. The figure on the left presents the distribution in the last period before receiving the vouchers, and on the right, in the first period after receiving the vouchers. On each figure, the 1000 counterfactual households are grouped into 11 intervals by the value of the state variable: 0, (0,0.1],(0.1,0.2],...,(0.9,1]. The size of the circle measures the number of households in the interval. Before receiving the WIC vouchers (figure D3a), the state variable equalled zero for a large share of households, while after receiving the vouchers (figure D3b) the number of people in that interval is small. Overall, after receiving the vouchers the distribution shifts upwards, increasing both in the state variable and purchase probability. The figures also show that as expected, the purchase probability of whole wheat bread is an increasing function of the state variable κ_{WW}. Note that the relationship between the state variable κ_{WW} and the purchase probability slightly differs before and after receiving the vouchers. This is because the probability also depends on which product was bought last period and that changes with the vouchers.

Figure D4 presents an example of a counterfactual household's purchase probability as a function of the state variable κ_{WW} and describes the evolution of the state variable over time. These results are presented for the time period (20 years) after receiving the vouchers. The black markers indicate the baseline counterfactual and the red markers indicate the counterfactual with the 10-cent whole wheat bread price discount. Figure D4a presents for the example counterfactual household the joint distribution of the purchase probability of whole wheat bread and the state variable κ_{WW}. Note that the state variable κ_{WW} takes 19 possible values from zero to one. This is because in the counterfactual, it is assumed that a year is 18 purchase occasions. Also, note that for a given value of state variable κ_{WW}, the purchase probability is the highest when whole wheat bread was purchased on the last trip



(a) Before WIC

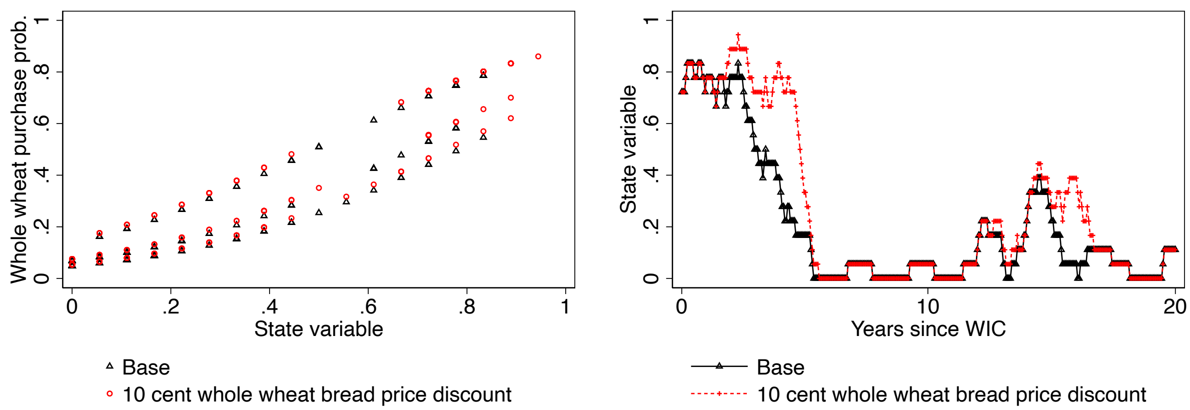
(b) After WIC

Figure D3: Counterfactual distribution of purchase probability of whole wheat bread as a function of state variable κ_{WW} before and after receiving the WIC vouchers

Note: The figures present the joint counterfactual distribution of state variable κ_{WW} that is the share of whole wheat in past year's bread purchases, and the probability of purchasing whole wheat bread. The figure on the left presents the distribution in the last period before receiving the vouchers, and on the right, in the first period after receiving the vouchers. Each figure is calculated using 1000 counterfactual households. The 1000 households are grouped into 11 intervals by the value of the state variable: 0, (0,0.1],(0.1,0.2],...,(0.9,1]. The size of the circle measures the number of households in the interval.

and takes lower values depending on which substitute bread product was bought. Figure D4b presents for the same example counterfactual household the evolution of the state variable over time.

Figure D5 presents various characteristics as a function of whole wheat bread purchase probability right before receiving WIC vouchers. One thousand counterfactual households are grouped into quartiles by the value of the purchase probability before receiving the vouchers. The x-axes displays the mean of the purchase probability before receiving the vouchers in each quartile. The y-axes shows for the households in each quartile either the average value of the price parameter α_i (figure D5a); the average relative preference for whole wheat bread, calculated as the difference between the preference for whole wheat parameter versus the mean of the preference for whole grain, white, and other (the base group) bread, $\theta_i - \frac{\delta_i+\lambda_i+0}{3}$ (figure D5b); the average purchase probability after receiving the vouchers (figure D5c); the

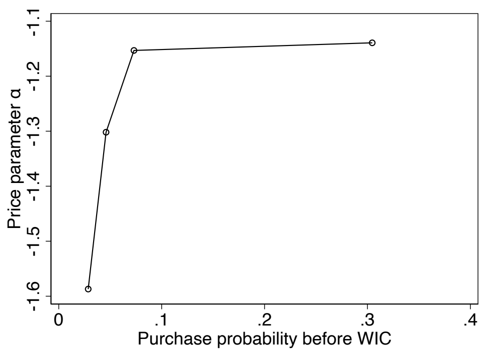


(a) State variable κ_{WW} vs purchase probability (b) State variable κ_{WW} over time

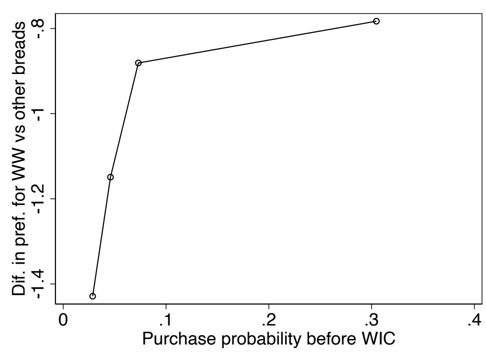
Figure D4: Counterfactual results for an example household

Note: The figure presents results for an example counterfactual household for the time periods (20 years) after receiving the vouchers. Figure D4a presents the joint distribution of the purchase probability of whole wheat bread and the state variable κ_{WW}, which is the share of whole wheat bread in past year's bread purchases. Figure D4b presents the evolution of the state variable over time. The black markers indicate the baseline counterfactual and the red markers indicate the counterfactual with the 10-cent whole wheat bread price discount.

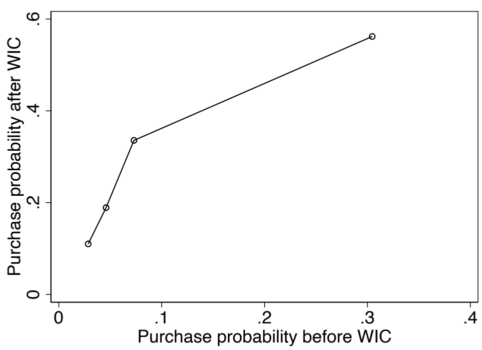
average change in the purchase probabilities from before to after receiving the vouchers (figure D5d); the average difference in the number of years it takes to return to pre-WIC purchase probability with a 10-cent discount on whole wheat bread versus the baseline counterfactual (figure D5e). The figure suggests that households that before WIC had lower whole wheat purchase probability, were more price elastic and had a lower preference for whole wheat relative to other breads, and for them the counterfactual of the whole wheat bread price decrease, has a larger effect.



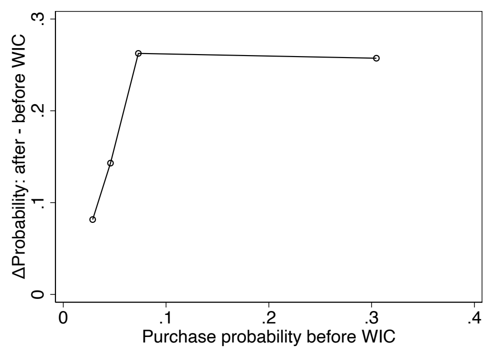
(a) Price parameter α



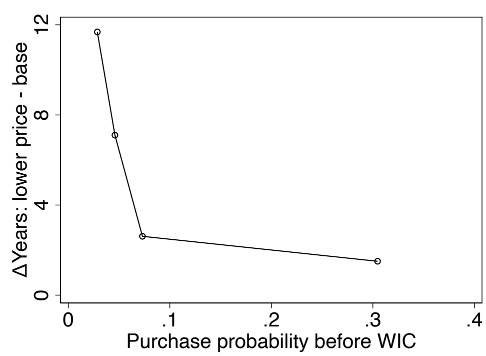
(b) Relative pref. for whole wheat



(c) Purchase probability after WIC



(d) Δ purchase prob. during WIC



(e) Effect of the price decrease

Figure D5: Counterfactual: characteristics as a function of pre-WIC whole wheat bread purchase probability

Note: One thousand counterfactual households are grouped into quartiles by the value of the whole wheat bread purchase probability before receiving the vouchers. The x-axes displays the mean of the purchase probability before receiving the vouchers in each quartile. The y-axes shows for the households in each quartile either the average value of the price parameter α (figure D5a); the average relative preference for whole wheat bread, calculated as the difference between the preference for whole wheat parameter versus the mean of the preference for whole grain, white, and other (the base group) bread, $\theta_i - \frac{\delta_i+\lambda_i+0}{3}$ (figure D5b); the average purchase probability after receiving the vouchers (figure D5c); the average change in the purchase probabilities from before to after receiving the vouchers (figure D5d); the average difference in the number of years it takes to return to pre-WIC purchase probability with a 10-cent discount on whole wheat bread versus the baseline counterfactual (figure D5e).

References for Online Appendix

de Chaisemartin, C. and X. D'Haultfœuille (2020a): "Difference-in-Differences Estimators of Intertemporal Treatment Effects," SSRN Scholarly Paper ID 3731856, Social Science Research Network, Rochester, NY.

——— (2020b): "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110, 2964–2996.

DellaVigna, S. and M. Gentzkow (2019): "Uniform Pricing in U.S. Retail Chains," *The Quarterly Journal of Economics*, 134, 2011–2084.

Dubois, P., R. Griffith, and A. Nevo (2014): "Do Prices and Attributes Explain International Differences in Food Purchases?" *American Economic Review*, 104, 832–867.

Goodman-Bacon, A. (2021): "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 225, 254–277.

Institute of Medicine (2006): *WIC Food Packages: Time for a Change*, Washington, D.C.: The National Academies Press.

Johnson, P., L. Giannarelli, E. Huber, and D. Betson (2014): "National and State-Level Estimates of Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) Eligibles and Program Reach, 2011," Report WIC-14-ELIG, U.S. Department of Agriculture, Food and Nutrition Service, Office of Research and Analysis.

Oster, E. (2018): "Diabetes and Diet: Purchasing Behavior Change in Response to Health Information," *American Economic Journal: Applied Economics*, 10, 308–348.

Seo, B. (2019): "Firm Scope and the Value of One-Stop Shopping in Washington State's Deregulated Liquor Market," SSRN Scholarly Paper ID 2863094.

USDA (2018a): *USDA Branded Food Products Database*, U.S. Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory.

——— (2018b): *USDA National Nutrient Database for Standard Reference*, U.S. Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory.