Retail Globalization, Households' Diets and the Effectiveness of Sin-Food Taxes

Emilio Gutierrez, Beata Javorick, Wolfgang Keller, Faqiang Li, Ricardo Miranda, Kensuke Teshima, James Tybout

Preliminary and Incomplete

this version: July 6, 2024
Please click here for the latest version

Abstract

We study the impact of retail globalization on calorie consumption under alternative policy regimes. Specifically, we first examine the effects of Walmart openings in Mexican cities on household consumption patterns using household-level surveys and home scanner data. In doing so, we document an eight percent permanent increase in households' purchased calories that coincides with the timing of Walmart openings, and we show that this increase traces to greater consumption of unhealthy foods. Next, we show that when Mexico introduced a tax on highly caloric foods in 2014, caloric intake fell among Walmart shoppers, who substituted for cheaper and healthier food options. Finally, building on Thomassen et al. (American Economic Review, 2017), we estimate a structural model of households' choices concerning the stores they visit and the products they consume. This model provides a basis for counterfactual analyses of calorie taxes (inter alia), and it allows us to link changes in caloric intake among different types of households to Walmart openings.

1 Introduction

Rising obesity rates and its health consequences are a significant concern for developed and developing nations. Since 1980, obesity rates have almost tripled worldwide (World Health Organization, 2002). One first step toward combating obesity is to recognize that it is caused by increases in the difference between calories consumed and calories burned (Hill, Wyatt, and Peters, 2012). Researchers have proposed two reasons for the increase in calorie consumption: price reductions of high caloric content foods and increases in their availability (Cawley, 2015; Currie et al., 2010; Cutler, Glaeser, and Shapiro, 2003). In response to the increase in obesity rates, sin-food taxes have emerged as a countermeasure. Using the case of Walmart openings and expansions in Mexican cities, we explore how retail globalization, which affects the prices and availability of foods with high caloric content, impacts the composition of household diets and the effectiveness of sin-food taxes.

Mexico offers an ideal setting to investigate the relationship between retail globalization, nutrition, and health policies. It is a country in which the fraction of the population that is obese or overweight has grown the most in recent decades. According to the Mexican Health and Nutrition Survey, known locally as ENSANUT, in 2012, 35% of the Mexican population was obese. In January 2014, the Mexican government enacted a tax on high-calorie foods and beverages to address this alarming trend. Moreover, in recent years, Mexico has experienced a rapid increase in the number of Walmart stores, a large foreign retail chain, which has transformed the composition of the retail industry (Atkin, Faber, and Gonzalez-Navarro, 2018; Iacovone et al.,2015). In addition, rich micro-level data are available for our research. We use two types of detailed household data: the National Survey of Household Income and Expenditure, or ENIGH, a nationally representative household consumption and expenditure survey, and scanner data collected by KANTAR Worldpanel.

In the first part of the analysis, using an event study approach, we document an 8% permanent increase in the calories consumed by households, which coincides with the timing of Walmart openings. We find that this increase is concentrated in unhealthy foods. Moreover, we find that after entry, households source the largest part of their weekly consumption of packaged products from Walmart stores more frequently. In the second part of the analysis, we study the effectiveness of the tax on high-caloric content foods imposed by the Mexican government in January 2014. We find that households who frequently source most of their packaged product consumption from Walmart reduced their total caloric intake as opposed to the rest of the households. They reduce their calorie consumption mainly from untaxed products. This phenomenon is explained by Walmart's provision of products in the middle range of calorie intensity is relatively more minor. This induced households to substitute for cheaper and healthier products of low calorie intensity. We rationalize these results with a grocery shopping model developed by Thomassen et al. (2017).

Our study is related to several strands of the literature. First, it relates to studies linking globalization and obesity. Cross-country studies find mixed results. For example, Miljkovic et al. (2015) and Vogli et al. (2014) find a positive and significant association between trade openness and obesity and body mass index, while no such relationship is found in other studies (Oberlander, Disdier, and Etilé, 2017; Costa-Font and Mas, 2016). Our study is unique in focusing on the role of retail globalization in both rising obesity and the effectiveness of policies against it.

Our study also contributes to the literature on food deserts and the nutritional implications of shopping costs. While we do not find that Walmart entries on their own improve the diet of households, we show how retail globalization increases the availability of alternatives to high caloric content foods in developing countries (Rose and Richards 2004; Morland et al., 2002; Cotterill and Franklin, **cotteriimpact**; Weinberg, 1995) and we document how this allowed consumers to substitute toward a wider variety of products and sizes after the tax enactment (Allcott et al., (2019)[1]; Chung and Myers Jr, 1999; Kaufman, 1998; Kaufman et al., 1997). This result, along with our finding that after entry, households start sourcing most of their packaged products consumption from Walmart stores, points in the direction of studies that have analyzed the implications of travel costs and one-stop shopping, such as Thomassen et al. (2017). The implications of these purchasing behaviors for the effectiveness of sin-food taxes remain to be studied.

Our study relates to the growing literature that has linked Walmart and other supermarket openings to

[1]While Allcott et al., (2019) do not find that supermarket entries affect nutritional outcomes, our setting differs from theirs in at least two aspects. First, travel costs are different in Mexico than in the USA. For instance, only 50% of our sample households have a car. Second, the alternatives to globalized retail that domestic retailers offer differ between developed and developing countries.

US households' body mass index and diet composition. (Volpe, Okrent, and Leibtag, 2013; Courtemanche and Carden, 2011)[2] Our contribution to this literature is to document whether the relationship between Walmart openings and the composition of the diets of households is different in Mexico than in developed countries.

Our paper also contributes to the understanding of the impact of retail globalization in low and middle-income countries (Iacovone et al. 2015; Javorcik and Li 2013; Javorcik, Keller, and Tybout, 2006). While some existing papers have focused on Mexico and Walmart openings, they differ from ours in terms of methodology and research questions. In particular, our study is the first to investigate the impact of Walmart on the caloric intake and diet composition of Mexican households. The closest to our research in terms of methodology and context is Atkin, Faber, and Gonzalez-Navarro (2018). They use an event-study method in the Mexican context. However, their analysis is broader, as it measures the impact of supermarket stores' entry on overall prices and welfare distribution.

Finally, our study is linked to research that has explored the relationship between food prices, taxes on foods and beverages and the diet of households, or obesity rates in Mexico (Aguilar, Gutierrez, and Seira, 2019; Colchero et al., 2016; Gracner, 2015) and other contexts (Harding and Lovenheim, 2017; Dubois, Griffith, and Nevo, 2014; Grossman, Tekin, and Wada, 2014; Fletcher, Frisvold, and Tefft, 2010; Powell and Chaloupka, 2009; Beydoun, Powell, and Wang, 2008). Our contribution to this literature is to identify retail globalization as one of the forces behind these price changes and analyze its implications on households' calorie intake and the effectiveness of food tax policy.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 introduces the empirical strategy. Section 4 presents the preliminary results. Section 5 presents the quantitative model. Section 6 introduces the estimation of the model. Section 7 shows the structural estimation results. Section 8 concludes.

2 Data

2.1 ENIGH: Composition of Households' Diets

To explore the composition of households' diets following Walmart's entry into Mexican municipalities, we use the ENIGH surveys. They are administered every two years. Although the municipalities and households covered change across surveys, they are designed to be representative at the national level. For this reason, all surveys include households from most of the large municipalities of Mexico. We analyze the ENIGH surveys of 2008, 2010, 2012, 2014, and 2016.

In the survey, households are asked to report all their food purchases during the three months before the survey. Their expenditures are then classified into multiple categories, one of which corresponds to foods and

[2] Holmes (2011), Jia (2008), Hausman and Leibtag (2007), and Basker (2005) also explore the relationship between Walmart store openings and other outcomes for the US context.

beverages. This category is partitioned into 247 different subcategories, 211 corresponding to foods, 24 to beverages, and 12 to tobacco, food for animals, meals eaten outside the household, and in-kind transfers.

While ENIGH includes detailed information on expenditures and quantities purchased for each subcategory listed, it does not contain nutritional information. To assign caloric contents to each food and beverage subcategory, we use the National Nutrient Database for Standard Reference (NNDSR) published by the United States Department of Agriculture. It is a data set on the nutritional content of most products consumed in the United States. It comprises 1,137 categories, and, in total, 85% of the food subcategories of the ENIGHs are covered. We also collected caloric content information on the ENIGH subcategories not included in the NNDSR, mainly traditional Mexican foods. When the NNDSR lists products in more detail than ENIGH, we assign the average calories per kg/liter of NNDSR categories to the corresponding unique subcategory in ENIGH.

Apart from expenditures, the ENIGH surveys also include detailed information about household socio-economic characteristics. They include household size, living-place characteristics, and all household members' age, employment, and health information. The exact location of households is not provided, but the surveys report the population of the localities in which households are located within municipalities.

2.2 KANTAR: Purchases of Packaged and Taxed Products

In addition to the ENIGH surveys, we exploit scanner panel data on households' consumption of packaged products collected by KANTAR Worldpanel. For each household, this high-frequency dataset registers household purchases at the store and barcode level for an extended period. In addition to stores and bar codes, it registers the quantities and prices of each purchase and the exact date at which each transaction occurred. Our data range from 2011 to 2015.

As for the ENIGH surveys, the KANTAR data do not include information on the nutritional content of each bar code. Thus, we obtain it from a database on the nutritional content of packaged products in Mexico that was specifically collected to be merged with the KANTAR Worldpanel data by Aguilar, Gutierrez, and Seira (2019). Nutritional content was directly collected for 71% of the bar codes in the data set. This represents 68% of the observed expenditures. For the remaining 29% of bar codes, caloric content was imputed at the bar code level from those barcodes for which caloric content was directly collected.

2.3 Walmart Store Entries

Our main data source on the dates Walmart stores entered Mexican markets is Walmart's monthly financial reports from the Walmart website of Mexico. They show the exact month and city of store openings. From the financial reports, we collected the months of 55 Walmart supercenters' opening in the cities included in the KANTAR data and the opening months of an additional 30 stores in the cities included in the ENIGH surveys for our analysis. These entries occurred from 2012 to 2015.

Because the ENIGH surveys are collected every two years, most of the entries we identify from the reports occurred between ENIGH 2012 and ENIGH 2014 waves. This limits the variation in the Walmart entries' timing required for an event study analysis. To address this issue, we complement our dataset on Walmart entries using the registries of Walmart supercenters in the Mexican National Directory of Economic Units (DENUE). This registry database is updated yearly for economic units with more than 100 employees. The frequency of these updates is sufficient for us to discern which Walmart entries occurred between ENIGH 2010 and ENIGH 2012. This, therefore, increases the time variation of the entries in our sample. Through this exercise, we recovered 44 additional entries outside the time covered by the financial reports of Walmart of Mexico.

3 Empirical Strategy

In this section, we discuss our method for estimating the effect of Walmart entries on households' calorie consumption and purchasing patterns. We also explain how we test whether consumers respond differently to a tax on high-calorie content (HCC) foods based on which stores they buy at, Walmart or other stores.

3.1 Effect of Walmart Store Entries on Calorie Consumption

We use an event study approach based on the ENIGH surveys to estimate the effect of Walmart entries to Mexican cities. These surveys cover all household food purchases, enabling us to observe the whole composition of households' diets and how they changed in response to the Walmart entries.

We include all entries from Walmart of Mexico websites between 2012 and 2015 in our regressions. As noted above, to expand the period covered by our analysis and achieve sufficient time variation to perform an event study, we use DENUE as an additional source to identify entries. DENUE is updated yearly for stores with over 100 employees, which is the case for Walmart supercenters. ENIGH surveys are collected every two years. Hence, for an event study, the registry dates to DENUE are sufficient to define whether a purchase is pre- or post-entry of Walmart stores.

In the ENIGH surveys, municipalities are further divided into smaller localities. From the DENUE data, 93% of Walmart supercenters are in localities with more than 15,000 inhabitants. We restrict the sample for our main analysis to those localities.[3] In our analysis, we retain only those municipalities appearing in at least two surveys before and after Walmart entries. We repeat those observations when more than one entry occurs in the same city but on different dates relative to the ENIGH surveys. This gives us 93 different entry time–city combinations, representing 129 Walmart store openings distributed across 80 municipalities over four years.

We are interested in the caloric intake of households and figuring out the product types in which the changes in diets are concentrated. Following Hut and Oster (2015), we classify all the products that appear in the ENIGH surveys as healthy if they are "obviously healthy." This refers to products that unambiguously

[3]Results for localities with less than 1,500 inhabitants are excluded from this version of the paper. We find no effect of Walmart store entries on households' diets in those localities.

are not harmful to human health. The healthy category includes fruits, vegetables, low-fat and fresh protein sources, such as fish and chicken. It excludes packaged products, such as cereals, candies, snacks, sodas, processed juices, and prepared meals. It also excludes foods with high fat, such as pork and beef. All the excluded products are labeled as unhealthy. To measure caloric consumption, we exploit information from the NNDSR and compute the monthly caloric intake of households from healthy products, unhealthy products, and the sum of calories from both categories.

Our event study specification is as follows:

$$log\ Y_{tmh} = \sum_{k=-2}^{3} \delta_k \mathbf{I}_{\{t-E_{jm}=k\}} + \tau_t + \eta_m + \alpha_t a_h \mathbf{I}_{\{t\}} + \beta_m s_h \mathbf{I}_{\{m\}} + \varepsilon_{tmh}, \tag{1}$$

where Y_{tmh} is the calorie from the food category of interest consumed by household h from municipality m at time t. The term $\mathbf{I}_{\{t-E_{jm}=k\}}$ is an indicator of whether the year of the observation, t, is equal to the $k-th$ year of the ENIGH wave (i.e., two years) after the entry, E_{jm}, of Walmart store j in municipality m. The terms τ_t and η_m are time and municipality fixed effects, respectively. The term a_h denotes household head age, and s_h denotes household size. The functions $\mathbf{I}_{\{t\}}$ and $\mathbf{I}_{\{m\}}$ are time and municipality indicators. Finally, ε_{tmh} is an error term. As noted above, the households in the ENIGH survey change from one survey to another. Therefore, the panel is balanced at the municipality level but not at the household level. For this reason, we control for the observables of households in addition to time (survey wave) and municipality fixed effects. Since purchased calories depend on households' size and that households' tastes might vary across cities, we include interactions between city indicators and household size. Moreover, consuming calories varies at one's different life stages, and such effects might also vary over time. For this reason, we control for household head age interacted with time fixed effects in our specification.

3.2 Timing, Purchasing Patterns, and Calories from Packaged Products

Using the ENIGH surveys, we can assess the long-term effects of Walmart entries on households' diets. However, we do not observe the timing of the consumption or the store choice changes. To answer these problems, we repeat our event study analysis using the KANTAR scanner panel data. Here, we can observe monthly consumption and the stores where purchases occur.

To ensure that our household panel is balanced, we restrict our sample to households for which we can observe consumption for at least six months before and after a Walmart entry. As in our analysis of the ENIGH surveys, we repeat observations when multiple entries to the same city occur at different times. This leaves us with 2,576 households in 28 cities, for which we observe 55 Walmart entries. Because we want to observe the consumption of households at the highest possible frequency (monthly), we only use entries from the Walmart of Mexico website.

Having household panel data allows us to include household fixed effects in our analysis and control for unobserved household tastes and characteristics that affect calorie consumption and purchasing patterns.

Thus, our event study is performed at the household–month level. Our specification is as follows:

$$logY_{tmh} = \delta_{<-6}\mathbf{I}_{\{t-E_j=<-6\}} + \sum_{k=-6}^{6} \delta_k \mathbf{I}_{\{t-E_j=k\}} + \delta_{>6}\mathbf{I}_{\{t-E_j=>6\}} + \tau_t + \eta_h + \beta_m t \mathbf{I}_{\{m\}} + \varepsilon_{tmh}, \tag{2}$$

where Y_{tmh} is the outcome of interest for household h's at time t in city m. $\mathbf{I}_{\{t-E_j=<-6\}}$ indicates whether j's entry is six months earlier to time t. $\mathbf{I}_{\{t-E_j=k\}}$ indicates whether it is k-th month till time t since j's entry. τ_t and η_h are time and household fixed effects, respectively. $\mathbf{I}_{\{m\}}$ is a city indicator; $\mathbf{I}_{\{m\}}$ allows for city-spoecific trends. ε_{tmh} is an error term

The outcome of focus is the total calories each household purchases. We are also interested in how purchasing patterns change after Walmart's entry. In particular, we explore if there is an increase in purchases at Walmart and if Walmart becomes the main store from which households source their consumption of packaged products. The answers to these questions help us better understand the degree of relevance of Walmart as a retailer to households' shopping decisions. Hence, for every month and every household, we compute the proportion of weeks in which a Walmart store had the largest share of the weekly observed expenditures. This indicates whether Walmart stores are the household's primary source of packaged products. Finally, we measure the share of total expenditures and calories corresponding to Walmart purchases during weeks in which a Walmart store was the households' main source of packaged products.

3.3 Effect of Taxing High Caloric Content Food on Calorie Consumption and Substitution

The prices and availability of products might differ across Walmart and other retailers. This may induce differences in consumers' responses to price changes depending on where they buy groceries. In particular, the response to changes in the price of HCC products due to the introduction of the tax on them might differ between Walmart shoppers and shoppers of other stores. We explain the test of this hypothesis below.

To prepare for the tests, we classify all products that appear in the KANTAR data set according to their calorie content. We begin by identifying the products that are subject to the sin-food tax. The tax applies to all beverages with added sugar and all foods with more than 275 kilocalories per 100 grams and are not considered an essential component of the diet of Mexicans (such as oil or tortillas). We define all taxed products as HCC products. Then, we divide untaxed bar codes into two broad categories: foods and beverages. Within each broad category, products that fall in the first quartile of *total* calories are defined as low caloric content (LCC), and all remaining products are defined as middle caloric content (MCC). The definition of LCC products does not imply that these products are healthier because both caloric density and size must be considered. For instance, sufficiently small presentations of unhealthy products could be labeled LCC products.

To test which products are relatively cheaper at Walmart, we exploit the KANTAR data set at its most disaggregated level. We regress the log of observed purchase prices on dummies, which indicates if a purchase occurred at Walmart and its interactions with indicators of the type of product that was bought. We are interested in the sign of the Walmart indicator and its interactions. Our specification is as follows:

$$logPrice_{bstp} = \gamma_b + \mu_t + \omega \mathbf{I}_{\{s=Walmart\}} + \sum_{i \in h,m,l} \beta_{iw} \mathbf{I}_{\{s=Walmart\}} \mathbf{I}_{\{b=i\}} + \varepsilon_{bstp}, \tag{3}$$

where $Price_{bstp}$ is the price of bar code b at store s in month t for purchase p. γ_b is the barcode fixed effect and μ_t is the month fixed effect. $\mathbf{I}_{\{s=Walmart\}}$ is an indicator of the purchase occurring at Walmart, and $\mathbf{I}_{\{b=i\}}$ is an indicator of product type (HCC, MCC or LCC) of b. ε_{bstp} is an error term.

Furthermore, we examine how the composition of households' diets changes due to the tax and if this effect is the same for Walmart shoppers as for the rest. To do this, we classify households as Walmart shoppers according to their pre-tax purchasing behavior. We proceed as follows: first, we identify for each household the weeks in which the largest share of their expenditures in packaged products was spent at Walmart stores. Then, we define Walmart shoppers as all the households that regarded Walmart as their main store for at least one week of the month for at least nine out of the twelve months in 2013. While these variables are defined concerning weekly expenditures, we aggregate them by month so that the level of measurement coincides with the time disaggregation level at which we observe Walmart entries in our sample.

Finally, to fulfill the key purpose of revealing the relationship between calorie intake, sin-food tax, and Walmart entries, we design the following tests. We first compute the total amount of purchased calories from all product types (HCC, MCC, and LCC products) and regress them on the indicator of the tax enactment and its interaction with an indicator of whether a household is a Walmart shopper. We repeat this analysis by separately inspecting calories from product types of the taxed, i.e., HCC, and the untaxed, i.e., LCC and MCC. The specification for this test is:

$$logY_{ht} = \eta_h + \tau_t + \beta_T \mathbf{I}_{\{t \in T\}} + \beta_{TW} \mathbf{I}_{\{t \in T\}} \mathbf{I}_{\{h \in Walmart\}} + \varepsilon_{ht}, \tag{4}$$

where Y_{ht} represents total purchased calories by household h from the category of interest (HCC, MCC, LCC, and their relevant combinations) at time t. Time and household fixed effects are represented by η_h and τ_t. $\mathbf{I}_{\{t \in T\}}$ indicates whether t is after tax-enactment. $\mathbf{I}_{\{h \in W\}}$ indicates that household h is a Walmart shoppers according to its pre-tax purchasing behavior. The error term is represented by ε_{ht}.

Note that if there is a positive association between Walmart entries and caloric intake, households who start attending Walmart after the tax would experience an increase in their caloric intake. This could be a confounding factor for our analysis. Therefore, we exclude from our sample the households who had no purchases at Walmart during 2013 but started attending it in 2014 or 2015.

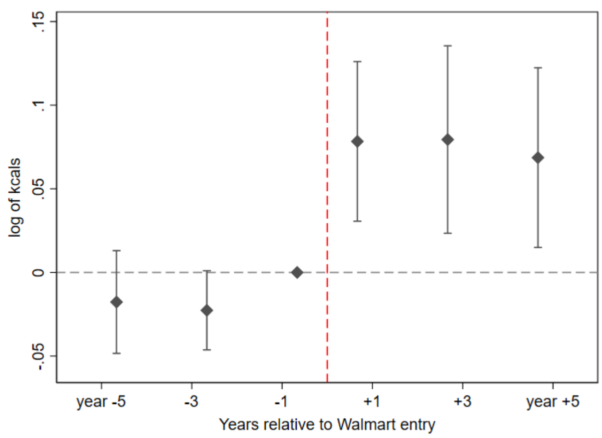
4 Preliminary Results

In this section, we present the results of our empirical implementation. We find that households' calorie consumption increased after the Walmart entries. In addition, we provide evidence that, after the enactment of the tax, households' substitution of taxed products with untaxed products differed depending on whether they shopped at Walmart or other stores. We conclude that the tax was more effective in reducing calorie consumption for households that could buy from Walmart.

4.1 Effect of Walmart Store Entries on Calorie Consumption

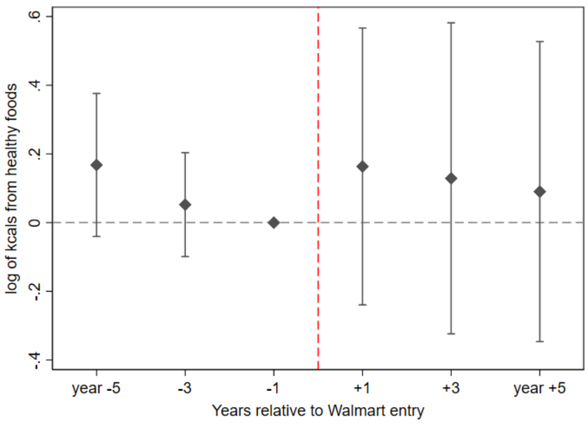
The event study using the ENIGH surveys shows that Walmart's entry into Mexican municipalities transformed households' diets. Figure 1 shows the event study coefficients from which we find that there was a permanent 8% increase in households' total calorie consumption after Walmart entered the local Mexican markets. There is no evidence of an increasing trend before its entry. Figures 2 and 3 show the decomposition of such increase into healthy and unhealthy products. We find that there is a 6% increase in calories from unhealthy products. This indicates that most of the increase comes from households consuming more unhealthy products. These findings are summarized in columns 1 to 3 of Table 1, where we report the average effect for all post-entry periods on total caloric intake and calories from healthy and unhealthy products.

Figure 1: Effect of Walmart Entries on Calorie Consumption of Households



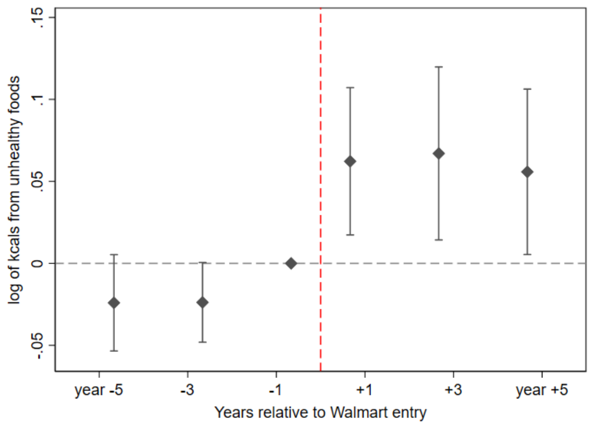
Notes: This figure depicts the event study coefficients for the log of households' calorie consumption from all products. The coefficients correspond to each ENIGH survey immediately before and after Walmart entries. The sample is restricted to municipalities that appear in at least two surveys before and after Walmart entries.

Figure 2: Effect of Walmart Entries on Households' Calorie Consumption from Healthy Products



Notes: This figure depicts the event study coefficients for a log of households' calorie consumption from healthy products. The coefficients correspond to each ENIGH survey immediately before and after Walmart store entries. The sample is restricted to municipalities that appear at least in two surveys before and after Walmart entries.

Figure 3: Effect of Walmart Entries on Households' Calorie Consumption from Unhealthy Products



Notes: This figure depicts the event study coefficients for the log of households' calorie consumption from unhealthy products. The coefficients correspond to each ENIGH survey immediately before and after Walmart store entries. The sample is restricted to municipalities that appear at least in two surveys before and after Walmart entries.

Table 1: Event Study Estimates: the Effect of Walmart Entries on Households' Calorie Consumption

	ENIGHs			KANTAR			
Variable	(1) log of calories All products	(3) log of calories Healthy products	(3) log of calories Unhealthy products	(4) log of purchased calories	(5) Prop. of weeks during which WM was the main store	(6) Share of exp. From WM as main store	(7) Share of cals. from WM as main store
Before entry	-0.025	0.0311	-0.025	0.00	0.001	0.00261	0.002
	(0.015)	(0.088)	(0.015)	(0.007)	(0.001)	(0.001)	(0.001)
After entry	0.079***	0.164	0.063**	0.014*	0.006***	0.007***	0.005***
	(0.029)	(0.246)	(0.027)	(0.007)	(0.001)	(0.002)	(0.001)
Municipality FE	Yes	Yes	Yes				
Time/survey FE	Yes	Yes	Yes				
H. Head education	Yes	Yes	Yes				
Household size	Yes	Yes	Yes				
Household and Date FE				Yes	Yes	Yes	Yes
City-specific trend				Yes	Yes	Yes	Yes
	0.226	0.104	0.262	0.576	0.542	0.547	0.533
	232,795	232,795	232,795	381,177	381,177	381,177	381,177

Notes: Estimates for the effect of Walmart entries on the calorie consumption of households. Results are shown aggregating across all pre- and post-entry periods for four years in the case of ENIGH surveys and six months in the case of the KANTAR data set. For ENIGH, the sample is restricted to the households in municipalities that appear at least in the two consecutive periods before and after entries. Results are reported in columns 1 to 3. The first column shows results for all calories in products purchased by households. Columns 2 and 3 show results for calories from healthy and unhealthy products. For results from the KANTAR data set, regressions are performed using a household monthly scanner panel that is perfectly balanced for the six months before and after entry. Results are shown in columns 4 to 7. Column 4 reports the results for using the log of purchased calories as the outcome under inspection. For columns 5 to 7, we use different outcome variables. We identified the weeks in which a single Walmart store had the largest share of total observed expenditures and, therefore, was the main store from which packaged products were purchased. Column 5 reports the effect of entries on the proportion of weeks in a month during which a Walmart store was the households' main source of packaged products. Columns 6 and 7 report the effect of entries on the share of total expenditures and purchased calories corresponding to the products bought from Walmart stores when they were the households' main source of packaged products.

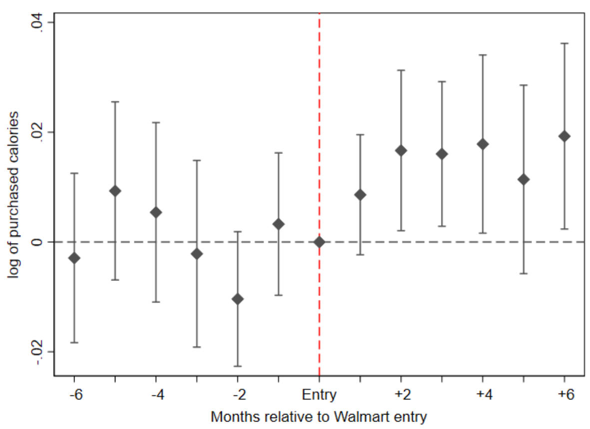
4.2 Timing, Purchasing patterns, and Calories from Bar-coded Products

Because of the low frequency of the ENIGH surveys, our estimates in the previous section could raise the question of whether the increase in calorie consumption coincides with the Walmart store entries. We inspect this using the KANTAR data and find that calorie increases from bar-coded products coincide with the months in which Walmart stores enter the Mexican municipalities. We show this in figure 4 and column 4 of Table 1. We find a 1.4% increase in household calorie intake from purchased packaged products, consistent with the ENIGH surveys' findings. Hence, we conclude that Walmart stores' entries are behind Mexicans' transition toward more calorie-intense diets.

Moreover, from Figures 5, 6, and 7, we find that recently opened Walmart stores have become one of the main sources of packaged products for households. Figure 5 shows a significant increase of 0.6% in the number of weeks per month during which a Walmart store had the largest share of observed household expenditures. Figures 6 shows a similar increase, by 0.05%, for the share of calories from the purchases for which Walmart stores the main stores. Figure 7 shows the share of total expenditures corresponding to purchases at Walmart stores during weeks where the households' main source of packaged products significantly rises by 0.7%. From these figures, we find no evidence of pre-entry trends in households' expenditure, purchased calories, or the frequency with which Walmart stores are the main source of packaged products for households in our sample for these tests. Columns 5 to 7 of Table 1 provide the results in values.

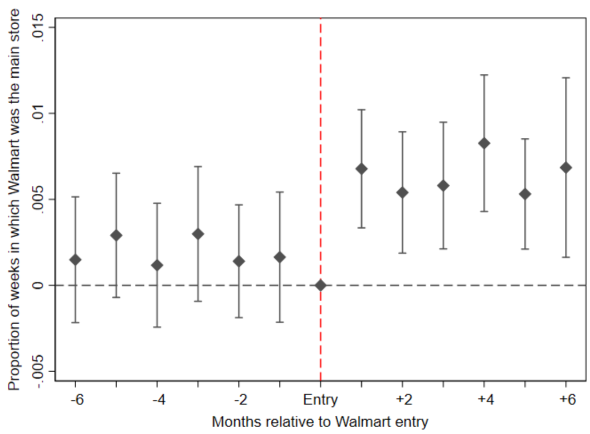
Note that the identified effects are driven by the average change among households in the city of entry. While, on average, there is a significant increase in the count of weeks in which Walmart stores are the main stores from which packaged products are bought, this effect is unlikely to be homogeneous across households. Only 200 out of the selected 2576 households in our sample used Walmart as their main store for at least one week every month after Walmart entry, and only 18 households used Walmart as their main store every week after Walmart entry. After its entry, Walmart became a relevant retailer, but other retailers remain the primary source of packaged products for many households.

Figure 4: Walmart Entries and Purchased Calories from Packaged Products



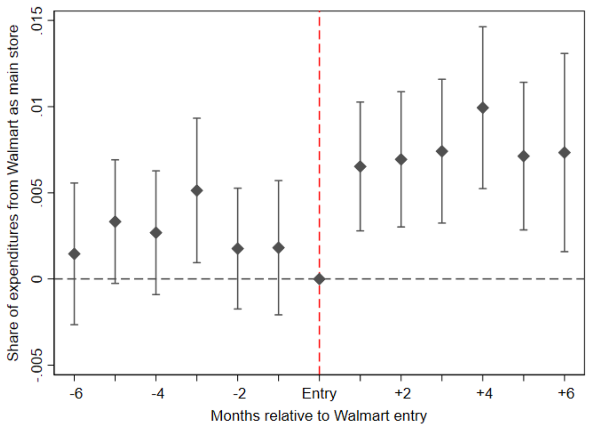
Notes: This figure depicts the event study coefficients for the log of households' observed purchased calories in the KANTAR data set. The coefficients correspond to each of the six months before and after entry. The sample is a perfectly balanced panel of households. The sample is restricted to cities that experience at least one entry during the sampling period.

Figure 5: Walmart Entries and Walmart as Main Source of Packaged Products using Expenditure



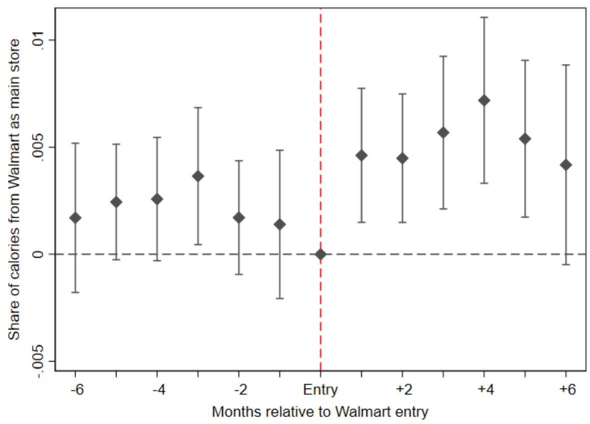
Notes: This figure depicts the event study coefficients for the number of weeks per month in which a single Walmart store had the largest share of total observed expenditures and was the household's main source of packaged products. The coefficients correspond to each of the six months before and after entry. The sample is a perfectly balanced panel of households restricted to cities that experience at least one entry during the sampling period.

Figure 6: Walmart Entries and the Share of Total Expenditures from Walmart as the Main Source of Packaged Products



Notes: For this figure, we identified the weeks in which a single Walmart store had the largest share of household observed weekly expenditures. Then, the dependent variable is the share of monthly expenditures in the identified store-week combinations, namely, when a Walmart store was the households' main source of packaged products. The figure depicts the event study coefficients for each six months before and after entry. The sample is a perfectly balanced panel of households restricted to cities that experience at least one entry during the sampling period.

Figure 7: Entries and the Share of Total Calories from Walmart as the Main Source of Packaged Products



Notes: For this figure, we identified the weeks in which a single Walmart store had the largest share of total household observed weekly expenditures. Then, the dependent variable is the share of monthly purchased calories from products purchased at the identified store-week combinations, namely, when a Walmart store was the households' main source of packaged products. The figure depicts the event study coefficients for each six months before and after entry. The sample is a perfectly balanced panel of households restricted to cities that experience at least one entry during the sampling period.

4.3 Effect of Tax on Calorie Consumption and Substitution

The key goal of this paper is to explore if the potential change in the composition of the households' diets induced by the tax on HCC foods was heterogeneous across those households who bought from Walmart and other households. Moreover, we seek to understand the underlying forces that could drive these heterogeneous responses.

Our analysis of prices shows that, on average, all products are 6% cheaper at Walmart than at other stores. Moreover, we find that both taxed and LCC packaged products are relatively cheaper at Walmart than at other stores. These results are summarized in Table 2. These differences in prices and relative prices provide suggestive evidence that households' response to the tax might vary between consumers depending on where they shop.

Table 2: Prices in Walmart by Types of Product Relative to the Rest of Stores During 2013 and 2014

	(1)	(2)	(3)	(4)	(5)	(6)
	Pre tax		Post tax		All periods	
Variables	log Price	log Price	log Price	log Price	log Price	log Price
Walmart	-0.0631***	-0.0396***	-0.0665***	-0.0488***	-0.0660***	-0.0452***
	(0.00426)	(0.00286)	(0.00274)	(0.00282)	(0.00404)	(0.00306)
Walmart#Taxed (HCC)		-0.0488***		-0.0380***		-0.0442***
		(0.00414)		(0.00293)		(0.00377)
Walmart#Untaxed Low caloric cont. (LCC)		-0.0173***		-0.0135***		-0.0134***
		(0.00440)		(0.00392)		(0.00322)
Barcode FE	Yes	Yes	Yes	Yes	Yes	Yes
Month of year FE	Yes	Yes	Yes	Yes	Yes	Yes
Period	2013	2013	2014	2014	2013-2014	2013-2014
R-squared	0.953	0.953	0.948	0.949	0.944	0.944
Observations	6367353	6367353	6227351	6227351	12594704	12594704

Robust standard errors clustered at the store level in parentheses
*** p<0.01 ** p<0.05 * p<0.1

Notes: The table shows estimates for prices in Walmart compared to other stores. The dependent variable is the log of the observed purchase price. The observations correspond to all the registered transactions at the date–store–bar-code level made by households in the KANTAR data set from 2013 to 2014. Each observation is weighted by the quantities purchased. The variable "Walmart" is a dummy that indicates the purchase occurring at Walmart. The variable "Taxed (HCC)" indicates all purchases of beverages and foods subject to the tax. The variable "Untaxed low caloric cont. (LCC)" indicates beverages and foods in the lowest quartile of caloric content among products not subject to the tax. The middle caloric content (MCC) product, which is also untaxed, is defined by those with caloric content above the first quartile of caloric content among products not subject to the tax. They are the excluded category. The first two columns were estimated using only pre-tax observations, whereas columns 3 and 4 were estimated using only post-tax observations. The last two columns were estimated using all observations.

Our estimates for the responses of households to the tax enactment show that substitution patterns were different depending on whether they shopped at Walmart or elsewhere. Column 1 of Table 3 shows that the tax had no significant impact on the total calorie consumption from bar-coded products of the households in our sample. However, we find that the tax caused an effective reduction in the total calories purchased by Walmart shoppers. This can be explained by the fact that the increase in calories from purchases of untaxed products induced by the tax was significantly lower for Walmart shoppers than for other shoppers. Disaggregating this effect, we find that Walmart shoppers substituted significantly more for LCC products, and their consumption of MCC products increased less than that of the rest of the households. Hence, we conclude that the tax was more effective in reshaping the composition of household diets for Walmart shoppers than it was for other consumers.

Table 3: Effect of the Tax and the Option to Attend Walmart on the Types of Calorie Consumption

Variables	(1)	(2)	(3)	(4)	(5)
	Purchased calories				
	All products	Taxed products	Untaxed products		
	All	HCC	MCC and LCC	MCC	LCC
Tax	227.5	-1,503***	1,730***	1,698***	31.80
	(340.8)	(187.8)	(215.8)	(208.6)	(55.71)
Tax#Walmart customer	-1,599**	-536.5	-1,063**	-1,278***	215.2**
	(702.3)	(353.0)	(469.5)	(450.5)	(101.7)
Household FE	Yes	Yes	Yes	Yes	Yes
Month of year FE	Yes	Yes	Yes	Yes	Yes
Dep. var mean	78,872	34,479	44,393	39,549	4,843
Tax+Tax#Walmart F-statistic	4.989	46.56	2.562	1.109	8.427
p-value	0.0256	0	0.110	0.292	0.00372
R-squared	0.540	0.559	0.477	0.482	0.484
Observations	131,796	131,796	131,796	131,796	131,796

Notes: The table shows the effect of the tax enactment on calorie consumption by types of its source of Walmart shoppers and the rest of households in the KANTAR data set. Walmart shoppers are defined according to their pre-tax purchasing behavior using the following procedure: For each household, we identify if there were weeks in which a Walmart store had the largest share of observed expenditures among all the stores from which the household bought at least one item. Then, a household is defined as a Walmart shopper if this was the case for at least one week of the month for at least nine out of the twelve months of the year. To avoid the bias induced by households who started attending Walmart in 2013 and therefore increased their calorie consumption, all households who made no purchases at Walmart during 2013 but began shopping at Walmart in 2014 or 2015 were excluded from the sample. The first column shows the effect on purchased calories from all products. Columns 2 and 3 show the effect on purchased calories from taxed (i.e., HCC) and untaxed (i.e., MCC and LCC) products, respectively. The last two columns disaggregate the effect shown in column 4 into MCC and LCC products. See Table 2 or in the paper for the definition of HCC, MCC, and LCC products.

5 Model

To put structure behind the patterns documented thus far and to thereby provide a basis for counterfactual analysis, we set up a demand model adapting Thomassen et al.'s (2017). In our model, heterogeneous consumers choose products, the stores where they purchase them, and the number of trips they will make to fulfill the purchase. Different stores have different products available and at various prices. Furthermore, stores differ in the commute time required to reach them, varying across households. So, taking stock of their tastes and budgets, households must weigh commute times, grocery offerings, and prices when choosing which store or stores to visit.

The mechanism underlying the model can explain the data patterns observed. Due to the shopping cost, Walmart's existence might affect households' caloric intake because households might prefer to fulfill their weekly shopping at fewer stores. Therefore, they like to shop at stores that offer a wide selection of products and competitive prices or when these stores are conveniently located. Generally, Walmart stores provide a wide range of competitively priced and high-calorie products. Households are more likely to purchase these products than they would have been if Walmart stores had not appeared in their neighborhood. Further, when sin-food taxes are introduced, Walmart shoppers may also have ready access to healthier substitutes that are not subject to the tax.

We now describe the model, following Thomassen et al.'s (2017) exposition. First, we describe the household utility function, suppressing household subscripts. Next, we characterize optimal shopping strategies. Finally, we introduce heterogeneity in households' tastes to bring the model to the data.

5.1 Preferences

Households derive utility from consuming K-dimensional bundles of groceries and derive disutility from trips to any stores in their available set, $\mathcal{J}$. Their weekly optimization problem is choosing how much of each grocery category and where to buy it. Prices are store-specific, and not all stores carry all food categories so that households may visit multiple stores in the same week. To contain the dimensionality of the problem, we follow Thomassen et al. (2017) by assuming that they visit no more than two stores per week and that they purchase a category of grocery from only one store. We provide evidence that supports these assumptions in the empirical strategy section.

To proceed, we introduce additional notations. First, let the set of stores chosen by a household from the set $\mathcal{J}$ be c, and let $\mathcal{C}$ be the set of all possible shopping choices. This set contains all the possible pairs of stores from $\mathcal{J}$ and all the singletons in $\mathcal{J}$. Next, let $\Gamma(c)$ denote the time and travel cost associated with visits to the store set c. Finally, let q_k be the quantity of grocery category k purchased, and let $d_k \in c$ indicate the store it is purchased from. Collected into the $K \times 1$ vectors, $\mathbf{q} = \{q_k\}_{k=1,K}$ and $\mathbf{d} = \{d_k\}_{k=1,K}$ jointly and fully characterize households' purchasing decisions.

We now lay out the household optimization problem. Households gain utility of a particular consumption bundle given the quantity and their sourcing choice of each grocery category:

$$u(\mathbf{d}, \mathbf{q}) = \mu_{\mathbf{d}}' \mathbf{q} - 0.5\mathbf{q}'\mathbf{\Lambda q}. \tag{5}$$

Here, the $K \times 1$ vector $\mu_{\mathbf{d}}$ captures category- and store-specific differences in appeal. Since products and prices are differentiated across stores, therefore $\mu_{\mathbf{d}}$ depends upon the sourcing choice, $\mathbf{d}$. The symmetric $K \times K$ matrix Λ indicates convexity in quantity demanded for one grocery category while allowing grocery categories to be complements or substitutes. Subtracting the utility losses from food expenditures and shopping trips from (5), we obtain the net value to households of any particular choice of quantities and sources, given price $\mathbf{p_d}$:

$$\begin{aligned} U(\mathbf{d}, \mathbf{q}) &= u(\mathbf{q}, \mathbf{d}) - \alpha \mathbf{p}_{\mathbf{d}}' \mathbf{q} - \Gamma(c_{\mathbf{d}}) + \varepsilon_c \\ &= (\boldsymbol{\mu}_{\mathbf{d}} - \alpha \mathbf{p}_{\mathbf{d}})' \mathbf{q} - 0.5\mathbf{q}'\Lambda\mathbf{q} - \Gamma(c_{\mathbf{d}}) + \varepsilon_{c_{\mathbf{d}}}. \end{aligned} \tag{6}$$

Here α measures the marginal disutility of expenditures, $\Gamma(c_{\mathbf{d}})$ measures the disutility of shopping trips, and

ε_{c_d} is a shopping preference shock.

As emphasized by Thomassen et al. (2017), this model admits two sources of product differentiation: category composition (captured by μ) and store accessibility c (captured by $\Gamma(c)$). The first source of differentiation captures the fact that different stores offer different varieties of groceries and different store amenities (e.g., store size). In particular, Walmart provides a relatively broad range of groceries and affords its consumers relatively more scope for substitution toward alternative products in response to taxes. The second source of product differentiation is the spatial variation in store locations relative to the consumer.

5.2 Solution Algorithm

To solve the household decision problem, we proceed in several stages. First, for any given set of stores, c, households choose an optimal source and quantity for each grocery category. Next, after determined the maximum welfare associated with each possible store set, the household decides the set that yields the highest net utility of shopping.

Specifically, for a store set c, the first stage maximization yields utility:

$$w(c, \mathbf{p_d}) = \max_{\mathbf{d} \in \mathcal{D}_c} \max_{\mathbf{q} \geq 0} \left[(\boldsymbol{\mu}_\mathbf{d} - \alpha \mathbf{p_d})' \mathbf{q} - 0.5 \mathbf{q}' \boldsymbol{\Lambda} \mathbf{q} \right]. \tag{7}$$

where $\mathcal{D}_c$ is the set of sourcing vectors $\mathbf{d}$ consistent with store set c. Note that the quadratic form admits simple closed-form solutions. The second stage optimization over store sets determines the maximized net value of shopping, which is the difference between utility in the first stage and the shopping cost:

$$\max_{c \in \mathcal{C}} \left[w(c) - \Gamma(c) + \varepsilon_c \right], \tag{8}$$

. We solve the discrete choice problem in (8) given the distributional assumption on ε_c. The choice of c can be expressed as a list of choice probabilities for all possible shopping options in $\mathcal{C}$.

5.3 Heterogeneity

We now introduce heterogeneity in $\boldsymbol{\mu}, \alpha$, and $\boldsymbol{\Gamma}$ across households and time. Consider first the grocery appeal parameter, $\boldsymbol{\mu}$. Earlier, for a given sourcing vector $\mathbf{d}$, we collected the grocery appeal indices for all grocery categories in the vector $\boldsymbol{\mu_d}$. So implicitly, if the k^{th} good was sourced from store j, the k^{th} element of this vector measured the appeal of store j's version of grocery category k, hereafter μ_{jk}. To let these μ_{jk} values depend upon household i's taste at time t, we write:

$$\mu_{itjk} = \xi_{jk} + \beta_{0k} \left(\beta_1 hz_i + \beta_2 sz_j + \mathbf{T}_t \boldsymbol{\beta}_T + \sigma_1 \nu_i^\mu + \sigma_2 \nu_{it}^\mu + \sigma_3 \nu_{ik}^\mu + \sigma_4 \nu_{ijk}^\mu \right), \tag{9}$$

where ξ_{jk} captures store-category effects common to all consumers, and the remaining terms capture household- and time-specific deviations from that effect. Specifically, hz_i stands for household size, sz_j stands for store size (measured by number of employees), and T_t is a vector of dummies allowing seasonal and year effects. Finally, we include four unobserved random shocks, each i.i.d. standard normal: a household effect ν_i^{μ} , a household-time effect ν_{it}^{μ} , a household-category effect ν_{ik}^{μ}, and a household-store-category effect ν_{ijk}^{μ}. Weighted by their standard deviations, $(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$, these ν-terms introduce horizontal product differentiation at the store-category level, enabling each household to perceive stores differently given a store-category.

Next, to allow for price sensitivity to depend upon household characteristics, we set the price coefficient to vary across households with their income y_i, household size hz_i, and unobserved characteristics ν_i^{α}:

$$\alpha_i = (\alpha_1 + \alpha_2 / (y_i/hz_i))\, \nu_i^{\alpha} \tag{10}$$

The term ν_i^{α} is a Rayleigh(1) random shock which introduces heterogeneity in a parsimonious way while ensuring positive price sensitivity $\alpha_i > 0$ for all i, as long as α_1 and α_2 are positive.

Finally, we let shopping costs depend upon the number of stores visited per trip, $n(c)$, total travel distance $dist_{ic} = 2\sum_{j \in c} dist_{ij}$, and unobserved standard normal shocks, ν_{i1}^{Γ} and ν_{i2}^{Γ}:

$$\Gamma_i(c) = \left(\gamma_{11} + \gamma_{12}\nu_{i1}^{\Gamma}\right) 1[n(c) = 2] + \left(\gamma_{21} + \gamma_{22}\nu_{i2}^{\Gamma}\right) dist_{ic}. \tag{11}$$

In our context, shopping costs vary across households and within households due to store entry. These sources of variation and associated variations in household choice sets are critical to identifying Walmart's impact on caloric intake and the interaction between Walmart accessibility and sin-food tax effects.

6 Structural Estimation

In this section, we explain our approach to structural estimation. We start by describing the key variables and primitives in the model. We then provide evidence that the data are consistent with the fundamental model assumptions and discuss sample selection. Finally, we present our estimator.

6.1 Measurement

We make various measurement choices to apply the model to the data and investigate our research questions. We group stores into store types, hereafter the "chain." We construct the relevant catchment area of stores for shoppers. We group barcode-level products into grocery categories. We describe our strategy for each below.

Chain Groups (f) To estimate the chain group-category specific parameters, ξ_{jk}, it was necessary to group stores according to their characteristics. For those stores belonging to major chains, like Walmart Supercenters, we create a store category for the chain. For smaller chains and independent stores, we create groups using cluster analysis. The clustering variables included the mean and standard deviation of their category-specific price indices, median market shares, the number of products they offered (measured at the city level) within each category, their median store size (as measured by the number of employees), and the total number of cities in which their stores appeared in the data. The resulting grouping gives us eight "chain" types: traditional corner stores, other traditional stores, small chains, medium chains, and big chains except for Walmart-owned stores, Walmart Supercenter, Aurrera (smaller Walmart stores targeted at low-income households), and Aurrera Express (also owned by Walmart). We hereafter index these chains by f, and replace ξ_{jk} with $\xi_{f_j k}$, where f_j maps each store j into a particular chain f.

Store choice sets ($\mathcal{J}$) Following Thomassen et al. (2017), we limit households' choice set to the nearest 30 stores, provided they fell within a 20-mile radius.[4] This makes $|J| = 30$ and the number of all the possible pairs of stores and singleton store, i.e., $|\mathcal{D}_c|$, to be $\frac{30\times 29}{2} = 465$. Below, we elaborate on our practices in measuring each key model component.

In KANTAR data, we only know the name of the affiliated chain of a store to which a household goes, while we do not observe the exact location and characteristics of a store. To identify the stores in each household's choice set, we match all the stores that appear in DENUE with all the chains appearing in households' purchase records from KANTAR data according to stores' chain affiliation, available in both datasets. With this procedure, we can assign the location and store characteristics for about 76% of all chains appearing in KANTAR data. They account for about 90% of the observed expenditures in the KANTAR data. With this procedure, we can match 80% of the total expenditures of the average household in our data to stores within less than 5 kilometers of the household's neighborhood. Having distance from every household to every store, we define a household's choice set as the set of all stores within a 20-kilometer distance.[5] More details and results of this procedure are provided in appendix C.

Categories of Groceries (k) We first group all barcodes appearing in households' purchases into five broad categories: self-care and toiletries, household goods, beverages, dry foods, and dairy-based products.[6] Then, to incorporate a healthiness dimension, we further split each of the last three categories into healthy and unhealthy. This leaves us with eight final categories: self-care and toiletries, household goods, healthy beverages, unhealthy beverages, healthy dry foods, unhealthy dry foods, healthy dairy-based products, and unhealthy dairy-based products.

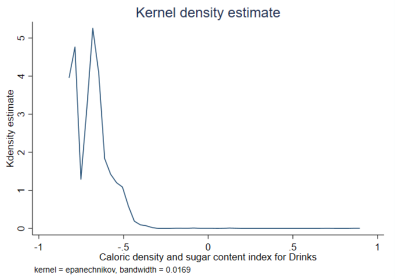
We calculate the first principal component of each product's caloric density and sugar content to distinguish healthy from unhealthy groceries. With the index in hand, we define all products with an index above the 75th within-category percentile as unhealthy and the remainder as healthy. This focus on caloric density and sugar content is consonant with the sin-food tax, based on the same information.

[4] A small number of households have less than 30 stores within 20 miles. For these, we assigned "virtual stores" with prohibitively high prices in all grocery categories, bringing their choice to 30.

[5] Note that KANTAR does not record expenditures at the store level but only at the chain level. When two or more stores that belong to the same chain appear in the choice set of a household, we only keep the closest one.

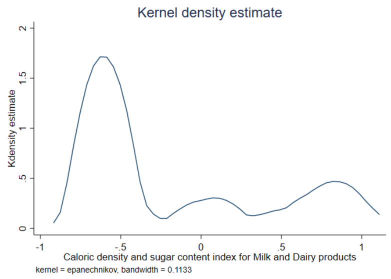
[6] In general, our categories follow Thomassen et al. (2017), except that we do not observe purchases of meat, fruit, vegetables, and non-packaged bakery products; thus, we merge the bakery and dry grocery to one category.

Figure 8



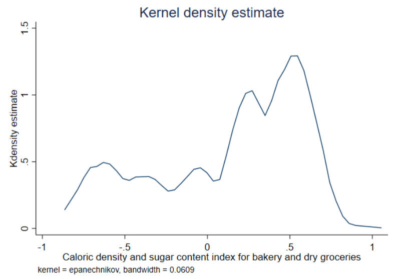
Note: Kernel density estimate for the distribution of the calorie intensity and sugar content index among beverages.

Figure 9



Note: Kernel density estimate for the calorie intensity and sugar content index distribution among milk and dairy products.

Figure 10



Note: Kernel density estimate for the calorie intensity and sugar content index distribution among bakery and dry groceries.

Price indices (p) Since chains generally price the products uniformly at all locations, we construct our price indices for each grocery category at the chain level. Specifically, we calculate our price variable following the two-level procedure in Thomassen et al. (2017). In the first step, the goal is to capture intra-store preferences for specific types of products within the broader categories. This ensures that the aggregate price index accurately reflects the price of products frequently purchased within the store. In the second step, the aggregation is done based on total revenue per grocery group (without considering store level variation) so that product type-store combinations not frequently purchased are not over-represented in the final index. See the appendix B for specific details on the imputation algorithms.

Calorie indices To link households' grocery choices predicted by the model to calorie consumption, we need to calculate the calories underlying the purchased groceries. Moreover, since calorie information is available at the product level, which is finer than the model's prediction, we aggregate the calorie information to the grocery category level. To this end, we construct calorie indices for each store-category combination that measure the total calories per spent peso and consider within-store preferences. This is done in three steps. First, we aggregate calories at the category-store-product level. Second, for every store-category-product combination, we calculate total expenditures, total calories, and calories per peso spent (*i.e., the price of a unit calorie from each product at each store*). Finally, for every store-category combination, we calculate the weighted average of the calories per peso; in this step, we use weights proportional to each product's observed expenditures. The resulting weighted average is the calorie index for each observed store-category pair.

Household size (hz) and employment size (sz) The information on the demographics of households can be found in the KANTAR. However, the dataset does not provide household income. Instead, each household is grouped into six socioeconomic categories according to the household's income and wealth. With this information, we impute household income levels for each socioeconomic group according to a public report that maps each socioeconomic group's average Mexican household income during the sample period.

Employment size information is available in the DENUE dataset. However, it only allows the number of employees to fall into six exclusive intervals of employee size (e.g., "having less than five employees", "having

10 to 100 employees). We use the logarithm of the lower bound of such intervals as the employment size measure for stores.

Quantities of groceries (q) To get the quantities of grocery demand at the category level as the model requires, we first aggregate expenditure to the store-category-week level for each household and divide by price index as constructed.

6.2 Sample Selection and Supporting Evidence for Key Model Assumptions

In selecting households for structural estimation, we first inspect the fitness of households' characteristics and behaviors to the key model assumptions. We eliminate unfit households from the full sample and do not consider them as the candidates to enter structural estimation. We then construct a subsample from the remaining households by randomly selecting them and their weekly purchase data. We elaborate on the implementation details below.

In the first step of elimination, we drop households that do not have long enough panels of weekly purchase data. Primarily, the reason for dropping them is that the information provided by these households may be noisy and low-quality.[7] Further, including households with too short panels stops us from using the timing of purchases, which brings temporal variation in prices and store availability. From the practical perspective, we also need the households selected for the structural estimation to have sufficiently distant weeks to facilitate further randomization of households.

In the second step of elimination, besides these households with short panels, we eliminate some households based on their behaviors. In this step, we concurrently inspect whether households' purchase patterns largely satisfy our assumptions for tractability in estimation. As shown in appendix A, although most households meet the behavioral patterns that support the key model assumptions, some deviate from the patterns. Therefore, we drop them from the eligible households that could enter the structural estimation. We also document in the appendix A the data loss in each elimination step and the summary statistics of the remaining samples surviving each elimination. Since the remaining households have similar summary statistics as the full sample, we believe these remaining households are reasonable representatives of the full sample. Finally, all the elimination steps leave us with around 7920 households. We randomly pick households further for the structural estimation because the model is computationally costly if we use all the eligible households. We select a subsample, a panel of 2,000 consumers, and three weeks per household, following the randomization procedure applied in Thomassen et al. (2017).

6.3 Estimation Strategy

As in Thomassen et al. (2017), we use the household purchase panel to estimate the model with a simulated method of moments. For each sampled household i in a week (or period) t, the model predicts the stores and

[7] For example, these households have scarce records because, even though they are surveyed, they might not fully understand how to use the expenditure recorder device properly. This raises doubt on the credibility of their data.

the quantities of all grocery categories the household chooses. We seek to match the predicted outcomes to the corresponding outcomes observed in the data. The parameters of interest are denoted as $\boldsymbol{\theta}$ and $\boldsymbol{\theta}$ is the vector $(\boldsymbol{\beta}, \boldsymbol{\xi}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\Lambda}, \boldsymbol{\gamma})$ from the model. Let $\mathbf{x}_{itc}^{w} = \left(\mathbf{x}_{itj}^{w}\right)_{j \in c}$ denotes the explanatory variables, including prices. Let $\boldsymbol{\nu}_{itc}^{w} = \left(\nu_{i}^{\mu}, \nu_{it}^{\mu}, \nu_{ik}^{\mu}, \nu_{ijk}^{\mu}, \nu_{i}^{\alpha}\right)_{j \in c, k=1,\ldots,K}$ denote the unobserved household heterogeneity. For an endogenous variable of interest, Y, let Y^{*} represent the observed and the corresponding population variable. Let the variable's symbol without a star and with $\boldsymbol{\theta}$ and $\boldsymbol{x}$ as arguments, i.e., $Y(\boldsymbol{x}, \boldsymbol{\theta})$, represent the variable predicted by the structural model. As in Thomassen et al. (2017), we further assume the observed value of the endogenous variable Y^{*} is the function (namely, $Y(.)$) of $\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{\nu}, \boldsymbol{\varepsilon}$ implied by the structural model evaluated at the true parameters $\boldsymbol{\theta}_0$:

$$Y^{*} = Y\left(\boldsymbol{\theta}_0, \mathbf{x}, \boldsymbol{\nu}, \boldsymbol{\varepsilon}\right)$$
$$\nu, \varepsilon \mid \mathbf{x} \sim \text{ with known density } f(\nu, \varepsilon \mid \mathbf{x})$$

We assume independence between $(\boldsymbol{\nu}, \boldsymbol{\varepsilon})$ and x so that $f(\boldsymbol{\nu}, \boldsymbol{\varepsilon} \mid \mathbf{x}) = f(\boldsymbol{\nu}, \boldsymbol{\varepsilon})$. Then, we can write the population conditional expectation in terms of the model's primitives as

$$Y(\boldsymbol{\theta}, \mathbf{x}) = E[Y(\boldsymbol{\theta}, \mathbf{x}, \boldsymbol{\nu}, \varepsilon) \mid \mathbf{x}] = \iint Y(\boldsymbol{\theta}, \mathbf{x}, \boldsymbol{\nu}, \varepsilon) f(\boldsymbol{\nu}, \varepsilon \mid \mathbf{x}) d\varepsilon d\boldsymbol{\nu}.$$

This gives us the following conditions to form moment conditions.

$$E\left[Y^{*} - Y\left(\boldsymbol{\theta}_0, \mathbf{x}\right) \mid \mathbf{x}\right] = 0. \tag{12}$$

We define the endogenous variables used in (12) to form moment conditions. For each household-week it, Q_{itcjk}^{*} is the quantity purchased of category k in store j from the shopping choice c; D_{itcjk}^{*} is the corresponding visit by i in time t in store j for grocery category k, which equals one if Q_{itcjk}^{*} is positive and zero otherwise; I_{itcjk}^{*}, the indicator of whether an c is chosen, which equals one if and only if the set of stores i in t visits is exactly c. From the model, the relationship between these endogenous variables and other explanatory variables x, unobserved individual heterogeneity ν, and parameter θ are:

$$\begin{aligned}
Q_{cjk}\left(\boldsymbol{\theta}, \mathbf{x}_{it}\right) &= \int q_{cjk}\left(\boldsymbol{\theta}^{w}, \mathbf{x}_{itc}^{w}, \boldsymbol{\nu}_{itc}^{w}\right) P_{c}\left(\boldsymbol{\theta}, \mathbf{x}_{it}, \boldsymbol{\nu}_{it}\right) f\left(\boldsymbol{\nu}_{it}\right) d\boldsymbol{\nu}_{it} \\
D_{cjk}\left(\boldsymbol{\theta}, \mathbf{x}_{it}\right) &= \int 1\left[q_{cjk}\left(\boldsymbol{\theta}^{w}, \mathbf{x}_{itc}^{w}, \boldsymbol{\nu}_{itc}^{w}\right) > 0\right] P_{c}\left(\boldsymbol{\theta}, \mathbf{x}_{it}, \boldsymbol{\nu}_{it}\right) f\left(\boldsymbol{\nu}_{it}\right) d\boldsymbol{\nu}_{it} \\
I_{c}\left(\boldsymbol{\theta}, \mathbf{x}_{it}\right) &= \int P_{c}\left(\boldsymbol{\theta}, \mathbf{x}_{it}, \boldsymbol{\nu}_{it}\right) f\left(\boldsymbol{\nu}_{it}\right) d\boldsymbol{\nu}_{it},
\end{aligned}$$

where

$$\begin{aligned}
P_{c}\left(\boldsymbol{\theta}, \mathbf{x}_{it}, \boldsymbol{\nu}_{it}\right) &= \int I_{c}\left(\boldsymbol{\theta}, \mathbf{x}_{it}, \boldsymbol{\nu}_{it}, \varepsilon_{it}\right) f\left(\varepsilon_{it}\right) d\varepsilon_{it} \\
&= \frac{\exp\left[w_{it}\left(c, \mathbf{p}_{t}\right) - \Gamma_{i}(c)\right]}{\sum_{c' \in \mathcal{C}_{it}} \exp\left[w_{it}\left(c', \mathbf{p}_{t}\right) - \Gamma_{i}\left(c'\right)\right]}
\end{aligned}$$

by the assumption that ε_{itc} follows i.i.d. type-1 extreme value distribution for each c and independent of ν_{it}.

As in (12), we assume the expectation of prediction error of quantities, visits, and the indicator of store choice has a zero mean conditional on the explanatory variables. These moment conditions are shown in the equations below:

$$E\left[Q_{itcjk}^{*}-Q_{cjk}\left(\boldsymbol{\theta}_{0},\mathbf{x}_{it}\right)\mid\mathbf{x}_{it}\right]=0,\ \text{for } k=1,\ldots,K$$

$$E\left[D_{itcjk}^{*}-D_{cjk}\left(\boldsymbol{\theta}_{0},\mathbf{x}_{it}\right)\mid\mathbf{x}_{it}\right]=0,\ \text{for } k=1,\ldots,K$$

$$E\left[I_{itc}^{*}-I_{c}\left(\boldsymbol{\theta}_{0},\mathbf{x}_{it}\right)\mid\mathbf{x}_{it}\right]=0.$$

Further, these moment conditions imply orthogonality conditions between the prediction errors and functions of the explanatory variable. The sample analogs of this relationship can be used to form the following empirical moments with a selected set of instruments. We denote this set of moment conditions as $\mathbf{g}_i^{(1)}$ in equation (13).

$$\mathbf{g}_i^{(1)}(\boldsymbol{\theta})=\begin{bmatrix}\sum_{t=1}^{T}\sum_{c\in\mathcal{C}_{it}}\sum_{j\in c}\mathbf{Z}_{itcj1}^{Q}\left(Q_{itcj1}^{*}-Q_{cj1}\left(\boldsymbol{\theta},\mathbf{x}_{it}\right)\right)\\ \vdots\\ \sum_{t=1}^{T}\sum_{c\in\mathcal{C}_{it}}\sum_{j\in c}\mathbf{Z}_{itcjK}^{Q}\left(Q_{itcjK}^{*}-Q_{cjK}\left(\boldsymbol{\theta},\mathbf{x}_{it}\right)\right)\\ \sum_{t=1}^{T}\sum_{c\in\mathcal{C}_{it}}\sum_{j\in c}\mathbf{Z}_{itcj1}^{D}\left(D_{itcj1}^{*}-D_{cj1}\left(\boldsymbol{\theta},\mathbf{x}_{it}\right)\right)\\ \vdots\\ \sum_{t=1}^{T}\sum_{c\in\mathcal{C}_{it}}\sum_{j\in c}\mathbf{Z}_{itcjK}^{D}\left(D_{itcjK}^{*}-D_{cjK}\left(\boldsymbol{\theta},\mathbf{x}_{it}\right)\right)\\ \sum_{t=1}^{T}\sum_{c\in\mathcal{C}_{it}}\mathbf{Z}_{itc}^{I}\left(I_{itc}^{*}-I_{c}\left(\boldsymbol{\theta},\mathbf{x}_{it}\right)\right)\end{bmatrix}\tag{13}$$

We also include cross-period moments $\mathbf{g}_i^{(2)}$ and cross-category moments $\mathbf{g}_i^{(3)}$ suggested by Thomassen et al. 2017. These moments are particularly useful in identifying the variance of household-invariant heterogeneity. In equation (14), we construct $\mathbf{g}_i^{(2)}$ by matching the empirical moments with the predicted moments from the model. The picked moments of this kind are the overall expenditure (R), category-specific quantities, usage incidence of a given store j for category k, one-stop shopping (OS), and total distance traveled ($DIST$). In equation (13), C_{ijt} is the set of shopping choices that include store j for consumer i in period t, and p_{ijk} is the price of category k at store j and time t.

$$\mathbf{g}_i^{(2)}(\boldsymbol{\theta})=\begin{bmatrix}\sum_{t=2}^{T}\left(R_{it}^{*}-R\left(\boldsymbol{\theta},\mathbf{x}_{it},\mathbf{x}_{i(t-1)}\right)\right)\\ \sum_{t=2}^{T}\sum_{k=1}^{K}\left(Q_{itk}^{*}-Q_{k}\left(\boldsymbol{\theta},\mathbf{x}_{it},\mathbf{x}_{i(t-1)}\right)\right)\\ \sum_{t=2}^{T}\sum_{j\in\mathcal{J}_{it(t-1)}}\sum_{k=1}^{K}\left(D_{itjk}^{*}-D_{jk}\left(\boldsymbol{\theta},\mathbf{x}_{it},\mathbf{x}_{i(t-1)}\right)\right)\\ \sum_{t=2}^{T}\left(OS_{it}^{*}-OS\left(\boldsymbol{\theta},\mathbf{x}_{it},\mathbf{x}_{i(t-1)}\right)\right)\\ \sum_{t=2}^{T}\left(DIST_{it}^{*}-\mathrm{DIST}\left(\boldsymbol{\theta},\mathbf{x}_{it},\mathbf{x}_{i(t-1)}\right)\right)\end{bmatrix}\tag{14}$$

In equation (15), we construct $\mathbf{g}_i^{(3)}$ by matching the average product between spending on k and k' within (R^{in}) and across (R^{cr}) households' adjacent time periods.

$$\mathbf{g}_i^{(3)}(\boldsymbol{\theta})=\begin{bmatrix}\sum_{t=2}^{T}\sum_{k=1}^{K}\sum_{k'=k+1}^{K}\left(R_{itkk'}^{in*}-R_{kk'}^{in}\left(\boldsymbol{\theta},\mathbf{x}_{it},\mathbf{x}_{i(t-1)}\right)\right)\\ \sum_{t=2}^{T}\sum_{k=1}^{K}\sum_{k'=k+1}^{K}\left(R_{itkk'}^{cr*}-R_{kk'}^{cr}\left(\boldsymbol{\theta},\mathbf{x}_{it},\mathbf{x}_{i(t-1)}\right)\right)\end{bmatrix}\tag{15}$$

Finally, we write $\mathbf{g}(\boldsymbol{\theta}) = N^{-1}\sum_{i=1}^{N}\mathbf{g}_i(\boldsymbol{\theta})$, where $\mathbf{g}_i(\boldsymbol{\theta})$ vertically stacks the three sets of moments $(\mathbf{g}_i^{(1)}(\boldsymbol{\theta}), \mathbf{g}_i^{(2)}(\boldsymbol{\theta})), \mathbf{g}_i^{(3)}(\boldsymbol{\theta}))$, and the estimator is

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \mathbf{g}(\boldsymbol{\theta})'\mathbf{W}^{-1}\mathbf{g}(\boldsymbol{\theta}),$$

where the weighting matrix is the inverse of a consistently estimated variance-covariance matrix $\mathbf{W} = N^{-1}\sum_{i=1}^{N}\mathbf{g}_i(\tilde{\boldsymbol{\theta}})\mathbf{g}_i(\tilde{\boldsymbol{\theta}})'$.

7 Structural Estimation Results

This section shows the structural estimates and tests the goodness-of-fit using various metrics. Following these results, we propose our plan for the counterfactual simulations.

7.1 Structural Estimates

Table 4 shows a selected set of estimates of the model in which the cross-category elasticities, the off-diagonal elements of Λ, are set to be zero for all grocery category pairs. The category-specific scaling parameters, β_{0k}, are all positive and precisely estimated. Together with their corresponding quadratic parameters, the diagonal elements of Λ, they show a decreasing and positive marginal utility for consuming more quantities of any grocery category. Regarding households' taste in shopping in larger stores (measured by employment size), our estimate is lower than that in Thomassen et al. 2017. The positive coefficient on household size indicates that households with more members also extract higher values in consuming groceries. Regarding the scaling parameters of the standard deviation of household heterogeneity, the time-varying and store- or category-specific parameters are not precisely estimated. This reflects that conditional on other covariates, the rest of the unobserved heterogeneity that influences household choices is explained mainly by household-category level variation. The rest of the critical parameters, price sensitivity, the disutility of long-distance shopping, and the shopping cost of traveling to two stores in one purchase, all have intuitive signs and are estimated precisely. They reflect that households prefer lower prices and shopping with lower travel frequencies and distances. Different from Thomassen et al. 2017, the parameter that reflects the variability of price sensitivity to per-head income is not estimated precisely. This might be partly due to the lack of variability in the discrete-valued household income variable.

Table 4: Structural Estimates of the Demand

	Estimate	SE
Panel A. Store-category taste effects (β) and scaling terms (σ)		
H-bakery & dry grocery	1.697	0.075
H-dairy & milk	1.366	0.058
H-drinks	0.790	0.044
L-bakery & dry grocery	1.532	0.082
L-dairy & milk	1.310	0.155
L-drinks	0.926	0.063
Household goods	0.573	0.029
Employment size X non-traditional store	0.146	0.079
Household size	3.274	0.197
Fixed across category/store	2.436	0.128
Time-varying	1.326	1.161
Category specific	4.814	0.578
Store/category specific	0.761	0.826
Panel B. Second-order quadratic parameters (Λ)		
H-bakery & dry grocery	20.724	1.136
H-dairy & milk	12.154	0.646
H-drinks	12.376	1.004
L-bakery & dry grocery	12.620	0.771
L-dairy & milk	13.296	1.867
L-drinks	3.312	0.162
Household goods	7.365	0.524
Personal care	7.906	0.469
Panel C. Price parameters (α)		
Constant	1.831	0.001
1/[weekly income per head]	32.881	35.946
Panel D. Shopping costs (γ)		
Two store dummy	16.794	1.450
Standard deviation	0.891	0.173
Distance	17.586	2.378
Standard deviation	2.879	0.424

Notes: The table shows structural estimates for the unknown parameters of the model. Parameters are estimated using 6,000 consumer-week observations. Standard errors are corrected for simulation noise. Chain-category effects not reported.

7.2 Goodness of Fit

We inspect the goodness of fit using various metrics to test how accurately the estimated model matches the observed household choices. The choices under inspection are the chain group-category (f, k) level quantities Q_{fk}, visits D_{fk}, and revenue R_{fk}. They are defined as follows in equations (16), (17), and (18), respectively. To get them, for example, Q_{fk}, we take the summation of Q_{cjk}, the expected quantities from all purchase records used in estimation from a store j in store pair c for category k.

$$Q_{fk}(\mathbf{p}) = \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{j \in \mathcal{J}_f} \sum_{c \in \mathcal{C}_{ijj}} Q_{cjk}\left(\hat{\boldsymbol{\theta}}, \mathbf{x}_{it}\right) \tag{16}$$

$$D_{fk}(\mathbf{p}) = \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{j \in \mathcal{J}_f} \sum_{c \in \mathcal{C}_{ijj}} D_{cjk}\left(\hat{\boldsymbol{\theta}}, \mathbf{x}_{it}\right) \tag{17}$$

$$R_{fk}(\mathbf{p}) = \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_f}\sum_{c\in\mathcal{C}_{ijj}} R_{cjk}\left(\hat{\boldsymbol{\theta}}, \mathbf{x}_{it}\right) \tag{18}$$

In Table 5 panels A and B, we first investigate the closeness between the observed and the predicted household choices aggregated to all 64 pairs of chain groups and categories (eight chain groups and eight categories). We apply two metrics: the correlation coefficients (of quantity and visit) and the mean absolute prediction errors (of the chain group shares of category demand). The estimated model fits the data better if the correlation coefficients are closer to one or the mean absolute prediction errors are closer to zero. We find that the observed and predicted quantities positively correlate, with a high correlation coefficient of 0.993. This value is also high for visits, although the coefficient is smaller: 0.951. The mean absolute prediction error for chain group shares of category demand across 64 pairs of chain groups and categories is 0.023. It indicates that, on average, the predicted chain group market shares deviate from the observed ones by about 2.3%. We also investigate the same metric using observations of the one-stop shoppers and get 3.1% mean prediction errors.

Table 5: Goodness of Fit

	In-sample	
	Quantity	Visits
Panel A: Correlation between predicted and observed demands		
$\rho\left(Q_{fk}, Q^*_{fk}\right), \rho\left(D_{fk}, D^*_{fk}\right)$	0.993	0.951
Panel B. Mean absolute prediction errors		
B1. Firm share of category demand (all firms and categories)		
$\lvert s_{fk} - s^*_{fk}\rvert$	0.023	
B2. 1-stop shopper share of category demand (all categories)		
$\left\lvert s_{1ss,k} - s^*_{1ss,k}\right\rvert$	0.031	

Panel C. Chain market shares

	In-sample			
	Revenue		Visit	
	Pred	Obs	Pred	Obs
Corner store	0.524	0.588	0.576	0.498
Other traditional store	0.126	0.071	0.117	0.100
Big Chain (non-Walmart)	0.127	0.103	0.110	0.134
Medium Chain	0.006	0.022	0.004	0.019
Small Chain	0.026	0.014	0.030	0.024
Aurrera	0.149	0.170	0.127	0.188
Aurrera Express	0.011	0.004	0.014	0.008
Walmart Supercenter	0.032	0.028	0.023	0.029

Notes: The table shows the evaluation of the fitness of structural estimates for the unknown parameters of the model using correlation coefficient, mean absolute prediction errors, and the closeness of chain market shares.

In figure 11, we plot the observed market share of each chain and category measured in visits against the predicted one. The title of each subplot indicates the category of focus, and the dots in each subplot, therefore, map to all chains' market share of the category. The fitness is largely satisfactory, except that the market share predictions regarding high-caloric bakeries are distant from the observed share for several chain groups.

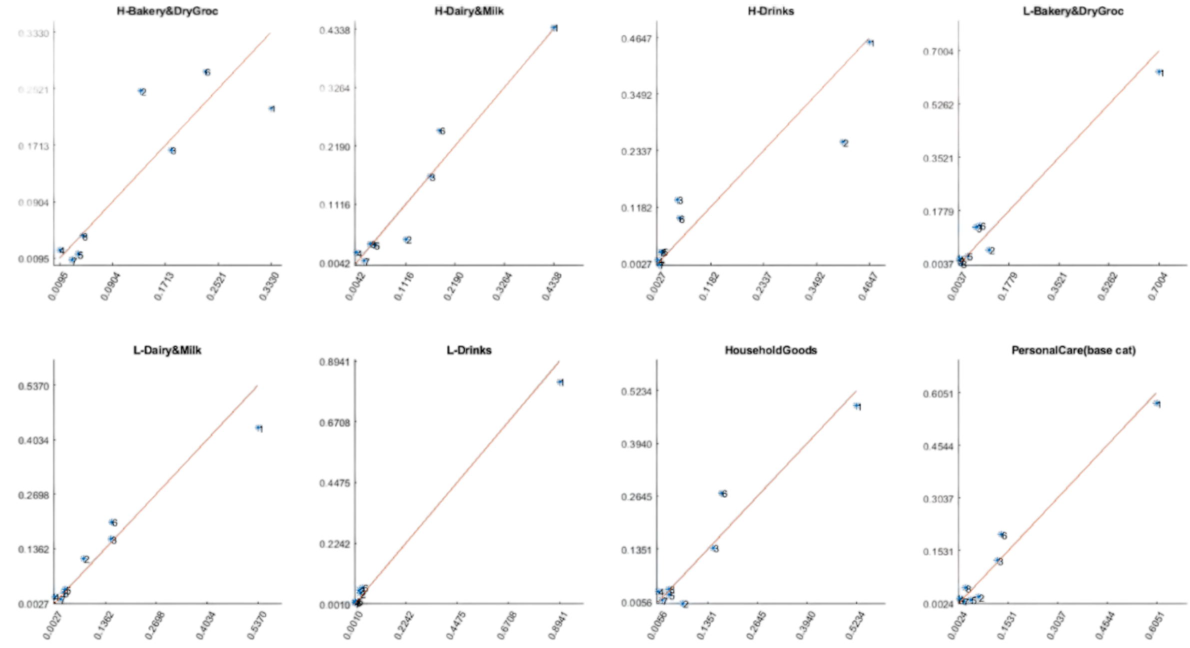


Figure 11: Predicted and Observed Shares of each Chain Group-Category in Visit

Notes: This figure shows observed (x) and predicted (y) market shares by category in terms of shoppers using the 6,000 consumer-weeks in the estimation sample (and their taste draws).

In Table 5, panel C, we compare the observed overall revenue and visit share by chains to their predicted counterparts to investigate the model's fitness at the chain level. For a certain grocery category k, We define the market share of a chain f for revenue naturally as the ratio between the total revenue purchased by all households in all periods in a chain and that of all chains, namely $R_{fk}(\mathbf{p})/\sum_{f'} R_{f'k}(\mathbf{p})$. The market share for visit is defined similarly as $D_{fk}(\mathbf{p})/\sum_{f'} D_{f'k}(\mathbf{p})$. The predictions are reasonably close to the observed revenue share of chains. However, one must be careful when using the current estimates to infer consumer choices on related corner stores and other non-corner traditional stores. For other non-corner traditional stores, we over-estimate their revenue share, and for the corner stores, we over-estimate their visit share, both in concerning magnitudes. Notably, we lack store-level characteristics, such as location and employment size, for corner stores and other traditional stores, and therefore impute them by assuming they are located at exactly households' residential location and have one employee. This imputation might be associated with the inaccurate prediction.

Overall, the estimated model provides a relatively satisfactory fit to the data. We will use the estimated model to investigate the interactive effect of Walmart's entry and the introduction of the sin-food tax. We will conduct counterfactual experiments in which three scenarios are considered: Walmart-owned stores had not entered; the sin-food tax was not introduced; or both happened concurrently. The main goal of these experiments is to evaluate the aggregate effects on households' intake of calories under the introduction of the tax and the heterogeneity of the impact across different Mexican households of distinct demographic backgrounds and retail environments.

8 Conclusion

We show that retail globalization significantly impacts households' calorie consumption. We reveal an increase in the consumption of unhealthy food coincided with the timing of Walmart openings. We also show that introducing a tax on highly caloric foods in Mexico in 2014 decreased caloric intake among Walmart shoppers, who opted for cheaper and healthier food alternatives. We construct and estimate a quantitative model adapted from Thomassen et al. 2017 to consider counterfactual experiments to investigate the synergic effect of retail globalization as represented by Walmart's entry and sin-food tax on households' grocery choices.

References

[1] Arturo Aguilar, Emilio Gutierrez, and Enrique Seira. "The effectiveness of sin food taxes: evidence from Mexico." In: *Instituto Tecnológico Autónomo de México (ITAM) Working Paper* (2019).

[2] Hunt Allcott et al. "Food deserts and the causes of nutritional inequality." In: *The Quarterly Journal of Economics* 134.4 (2019), pp. 1793–1844.

[3] David Atkin, Benjamin Faber, and Marco Gonzalez-Navarro. "Retail globalization and household welfare: Evidence from mexico." In: *Journal of Political Economy* 126.1 (2018), pp. 1–73.

[4] Emek Basker. "Job creation or destruction? Labor market effects of Wal-Mart expansion." In: *Review of Economics and Statistics* 87.1 (2005), pp. 174–183.

[5] May A Beydoun, Lisa M Powell, and Youfa Wang. "The association of fast food, fruit and vegetable prices with dietary intakes among US adults: is there modification by family income?" In: *Social science & medicine* 66.11 (2008), pp. 2218–2229.

[6] John Cawley. "An economy of scales: A selective review of obesity's economic causes, consequences, and solutions." In: *Journal of health economics* 43 (2015), pp. 244–268.

[7] Chanjin Chung and Samuel L Myers Jr. "Do the poor pay more for food? An analysis of grocery store availability and food price disparities." In: *Journal of consumer affairs* 33.2 (1999), pp. 276–296.

[8] M Arantxa Colchero et al. "Beverage purchases from stores in Mexico under the excise tax on sugar sweetened beverages: observational study." In: *bmj* 352 (2016), h6704.

[9] Joan Costa-Font and Núria Mas. "'Globesity'? The effects of globalization on obesity and caloric intake." In: *Food Policy* 64 (2016), pp. 121–132.

[10] Ronald Cotterill and Andrew Franklin. *The urban grocery store gap.* Tech. rep. 1995.

[11] Charles Courtemanche and Art Carden. "Supersizing supercenters? The impact of Walmart Supercenters on body mass index and obesity." In: *Journal of Urban Economics* 69.2 (2011), pp. 165–181.

[12] Janet Currie et al. "The effect of fast food restaurants on obesity and weight gain." In: *American Economic Journal: Economic Policy* 2.3 (2010), pp. 32–63.

[13] David M Cutler, Edward L Glaeser, and Jesse M Shapiro. "Why have Americans become more obese?" In: *Journal of Economic perspectives* 17.3 (2003), pp. 93–118.

[14] Pierre Dubois, Rachel Griffith, and Aviv Nevo. "Do prices and attributes explain international differences in food purchases?" In: *American Economic Review* 104.3 (2014), pp. 832–67.

[15] Jason M Fletcher, David E Frisvold, and Nathan Tefft. "The effects of soft drink taxes on child and adolescent consumption and weight outcomes." In: *Journal of Public Economics* 94.11-12 (2010), pp. 967–974.

[16] Tadeja Gracner. "Bittersweet: how prices of sugar-rich foods contribute to the diet-related disease epidemic in Mexico." In: *Job Market Paper* (2015).

[17] Michael Grossman, Erdal Tekin, and Roy Wada. "Food prices and body fatness among youths." In: *Economics & Human Biology* 12 (2014), pp. 4–19.

[18] Matthew Harding and Michael Lovenheim. "The effect of prices on nutrition: comparing the impact of product-and nutrient-specific taxes." In: *Journal of health economics* 53 (2017), pp. 53–71.

[19] Jerry Hausman and Ephraim Leibtag. "Consumer benefits from increased competition in shopping outlets: Measuring the effect of Wal-Mart." In: *Journal of Applied Econometrics* 22.7 (2007), pp. 1157–1177.

[20] James O Hill, Holly R Wyatt, and John C Peters. "Energy balance and obesity." In: *Circulation* 126.1 (2012), pp. 126–132.

[21] Thomas J Holmes. "The diffusion of Wal-Mart and economies of density." In: *Econometrica* 79.1 (2011), pp. 253–302.

[22] Stefan Hut and Emily Oster. "Changes in Household Diet: Determinants and Predictability." In: (2015).

[23] Leonardo Iacovone et al. "Supplier responses to Walmart's invasion in Mexico." In: *Journal of International Economics* 95.1 (2015), pp. 1–15.

[24] Beata S Javorcik and Yue Li. "Do the biggest aisles serve a brighter future? Global retail chains and their implications for Romania." In: *Journal of International Economics* 90.2 (2013), pp. 348–363.

[25] Beata Javorcik, Wolfgang Keller, and James Tybout. *Openness and industrial response in a Wal-Mart world: a case study of Mexican soaps, detergents, and surfactant producers.* The World Bank, 2006.

[26] Panle Jia. "What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry." In: *Econometrica* 76.6 (2008), pp. 1263–1316.

[27] Phil R Kaufman. "Rural poor have less access to supermarkets, large grocery stores." In: *Rural America/Rural Development Perspectives* 13.2221-2019-2662 (1998), pp. 19–26.

[28] Phillip R Kaufman et al. *Do the poor pay more for food? Item selection and price differences affect low-income household food costs*. Tech. rep. 1997.

[29] Dragan Miljkovic et al. "Globalisation and obesity." In: *The World Economy* 38.8 (2015), pp. 1278–1294.

[30] Kimberly Morland et al. "Neighborhood characteristics associated with the location of food stores and food service places." In: *American journal of preventive medicine* 22.1 (2002), pp. 23–29.

[31] Lisa Oberlander, Anne-Célia Disdier, and Fabrice Etilé. "Globalisation and national trends in nutrition and health: A grouped fixed-effects approach to intercountry heterogeneity." In: *Health economics* 26.9 (2017), pp. 1146–1161.

[32] World Health Organization. *The world health report 2002: reducing risks, promoting healthy life*. World Health Organization, 2002.

[33] Lisa M Powell and Frank J Chaloupka. "Food prices and obesity: evidence and policy implications for taxes and subsidies." In: *The Milbank Quarterly* 87.1 (2009), pp. 229–257.

[34] Donald Rose and Rickelle Richards. "Food store access and household fruit and vegetable use among participants in the US Food Stamp Program." In: *Public health nutrition* 7.8 (2004), pp. 1081–1088.

[35] Øyvind Thomassen et al. "Multi-category competition and market power: a model of supermarket pricing." In: *American Economic Review* 107.8 (2017), pp. 2308–51.

[36] Roberto De Vogli et al. "Economic globalization, inequality and body mass index: a cross-national analysis of 127 countries." In: *Critical Public Health* 24.1 (2014), pp. 7–21.

[37] Richard Volpe, Abigail Okrent, and Ephraim Leibtag. "The effect of supercenter-format stores on the healthfulness of consumers' grocery purchases." In: *American Journal of Agricultural Economics* 95.3 (2013), pp. 568–589.

[38] Zy Weinberg. *No Place to Shop: The Lack of Supermarkets in Low-income Neighborhoods: Analysis of a University of Connecticut Study on Low-income Communitites and Households Receiving Public Assistance in 21 Metropolitan Areas*. Public Voice for Food and Health Policy, 1995.

A Sample Selection for Structural Estimation

We conducted three elimination steps on the full sample from the KANTAR dataset, which included 14116 unique households from 2012 to 2015. The first elimination applies to the households with lower-quality observations, and the second and third eliminations apply to a relatively small proportion of households that severely violate the model assumptions relative to other households. To illustrate, denote household expenditure as E henceforth. We summarize household expenditure to the household-week-store-category level, denoted by E_{itjk}. To further assist exposition, we use E_i, E_{it}, E_{itj}, and E_{itk} to represent the expenditure summed at the corresponding levels intuitively. For example, $E_{it} = \sum_j \sum_k E_{itjk}$.

To recall, three assumptions were made in the model for traceability in estimation. The second and third steps of elimination center on making the remaining households meet these assumptions relatively more. These assumptions are: (1) households' choice sets are the nearest 30 stores to their residential locations ; (2) for each i in each week t, at most two stores are visited; (3) for each i in each week t, if the purchase behaviors take place in two stores, then for each k purchased, only one store is visited.

A.1 First Elimination

In the first step, we eliminate households with short weekly expenditure records or poor data quality. Truncating these households helps us in two ways. First, we are concerned that households with short purchase records bring too many measurement errors. For example, these households may not fully understand how to use the expenditure recorder device properly. Second, as in Thomassen et al. (2017), we need to randomly select households' purchase records in different weeks for the structural estimation sample. Therefore, the available weeks we can choose from need to be sufficiently distant from each other within a household. This truncation is particularly important to us since we expect to analyze Walmart's entry on purchase, which is less likely to be observed for households with short panels.

In practice, we drop households if they survive in the KANTAR weekly expenditure dataset for less than two quarters or have less than 24 weeks of purchase records. In this step, 21% (2950/14116) of the households are dropped.[8]

Additionally, there were 1596 households from 2012 to 2015 whose identifiers appear in the purchase records but not in the household characteristics dataset. Most such households were newly entered into the KANTAR expenditure dataset in 2015. Therefore, we drop the union of these households and the above households with a shorter sample length.

In this first step, we are left with 10378 households. These households have relatively long purchase records and no missing variables. In the following elimination steps, all the descriptive statistics we provide are restricted to the set of survivors of the first elimination.

[8]Thomassen et al. (2017) also drops households with too short records to conduct further analysis, resulting in 23% of the households dropping from their sample. This is the only considerable elimination Thomassen et al. (2017) does because the three assumptions are primarily satisfied in their case.

A.2 Second Elimination

In the second step, we inspect assumption (1). We delete households whose total expenditure shares are large outside the 30 nearest stores of their choice set. We find that more than 90% of households spend more than 83.9 % within their 30 nearest stores. Therefore, assumption one is plausible given the majority of households spend most of their expenditures within the 30 closest stores.

Therefore, in practice, we drop a household i if

$$\frac{\sum_{j:30nearest} E_{ij}}{E_i} < 0.9.$$

This results in 1562 households (15%) dropping from 10378 households, leaving us 8816 households.

A.3 Third Elimination

In the third step, we inspect assumptions (2) and (3). In KANTAR, among expenditures of all pairs of household-week, i.e., $i-t$, roughly 46% of them are spent on less than three stores and 47% of them on three or four stores. Hence, the frequency of the appearance of weekly expenditures spent in more than two stores is high, which makes assumption two seem implausible. However, assumption two can still be reasonable given the following descriptive statistics. We show the expenditure share from the two stores where households spend the most relative to the overall expenditure of that weekly purchase for all households. Formally, t is defined as the expenditure share of the highest and the second highest E_{itj} across all js of the i, t cell. We find the summary statistics of the distribution of such measurement over more than 1.4 million pairs of $i-t$. It shows that more than 90% of the $i-t$ pairs involve only two significant stores, in which they account for more than 73.3% of the total $i-t$ expenditure.

Continuing to inspect the plausibility of assumption 3, we provide the distribution of the higher expenditure share spent among the two stores for each $i-t$ and category k. For example, suppose we see that in a weekly expenditure profile indexed by i and t, the expenditure E on category k is distributed as 64% on one store j_1 and 36% on the other store j_2. In that case, we record the number 0.64 for the $i-t-k$ cell. We show the percentile of the distribution of these statistics across all trios of $i-t-k$. The 10th percentile is 0.63, indicating that 90% of the expenditures E_{itk} involve one store taking more than 63% of the expenditure on the category k. Moreover, the 50th percentile is 1, meaning 50% of the E_{itk} involves expenditure on one store.

Therefore, in practice, to implement the third elimination, we first tag each E_{it} as "healthy" if it satisfies

$$\frac{E_{itj_1} + E_{itj_2}}{E_{it}} > 80\%,$$

where j_1 and j_2 indicates the first and the second highest stores in which i spends money on in t. Similarly, we tag each E_{itk} as "healthy" if

$$\frac{E_{it\tilde{j}k}}{E_{itk}} > 90\%,$$

where $\tilde{j}$ means the store in which the expenditure k household i in t spends on is the highest. Therefore, a "healthy" expenditure record can be understood as being closer to satisfying the assumptions 2 or 3 compared to an unhealthy expenditure record. Further, each household gets a measure of the degree of "healthiness" according to the two criteria. We then drop the union of the households with the lowest 5% healthiness in both measures, resulting in around 500 households loss.

Therefore, in the end, all the elimination leaves us with around 7920 households. These households are the candidates for constructing the structural estimation sample. We continue to trim all candidates' weekly expenditure records to satisfy the three assumptions exactly and use them in the structural estimation. To do this, firstly, all expenditures outside of the 30 nearest stores are coded as zero. Secondly, their expenditures of all purchases outside the two stores where they spend the most are coded as zero. Thirdly, in each weekly expenditure record, the expenditures spent on each store category with a lower share are coded as zero. From the elimination steps, we know these trimmings do not severely distort the information stored in these households' expenditures.

A.4 Sample Comparison

To ensure the selected set of households for structural estimation are representative of the full sample, we compare samples remaining under each elimination with the full sample in the mean and variance of household characteristics. Across samples of households surviving each elimination, they are balanced in terms of the selected list of household characteristics. This can be seen in Table A1.

Table A1: Comparison of Selected Household Characteristics Across Samples after Eliminations

	Full sample **(1)**	**First Surv.** **(2)**	**Second Surv.** **(3)**	**Third Surv.** **(4)**
Age of wife	41.21	41.67	41.49	41.34
	(14.19)	(14.03)	(14.03)	(14.10)
Employment of wife	0.26	0.25	0.25	0.25
	(0.437)	(0.432)	(0.433)	(0.433)
Education of wife	5.18	5.13	5.04	5.05
	(2.674)	(2.647)	(2.612)	(2.617)
Age of husband	43.77	44.11	43.94	43.76
	(14.37)	(13.91)	(13.90)	(13.99)
Employment of husband	4.01	4.03	4.01	4.01
	(1.328)	(1.336)	(1.313)	(1.320)
Education of husband	5.59	5.55	5.46	5.45
	(3.064)	(3.074)	(3.083)	(3.083)
Location socioeconomic level	3.74	3.74	3.78	3.79
	(1.373)	(1.381)	(1.368)	(1.370)
Count of all members	4.22	4.28	4.27	4.24
	(1.764)	(1.773)	(1.779)	(1.767)
Count of adults	2.54	2.57	2.56	2.54
	(1.107)	(1.106)	(1.104)	(1.092)
Count of adolescent	1.55	1.59	1.58	1.58
	(1.120)	(1.122)	(1.125)	(1.125)
Count of female members	2.25	2.29	2.28	2.27
	(1.195)	(1.203)	(1.202)	(1.202)
Count of male members	1.95	1.98	1.98	1.96
	(1.216)	(1.218)	(1.225)	(1.218)
Has a car	0.50	0.51	0.50	0.49
	(0.644)	(0.647)	(0.642)	(0.643)
Has a computer	0.44	0.46	0.44	0.44
	(0.707)	(0.720)	(0.714)	(0.714)
Has a TV	1.58	1.58	1.57	1.56
	(0.973)	(0.981)	(0.963)	(0.952)
Body mass index 1	27.56	27.71	27.69	27.62
	(5.417)	(5.268)	(5.272)	(5.204)
Body mass index 2	27.36	27.46	27.42	27.35
	(4.961)	(4.848)	(4.832)	(4.724)
Diabete rate	0.04	0.04	0.04	0.04
	(0.199)	(0.194)	(0.192)	(0.189)
N	15733	10365	8788	7920

Note: This table compares the household characteristics of the various samples. The household characteristics of interest are age, employment status, education level of the housewife and husband of the household, socioeconomic status at the residential location, the number of family members of different kinds, the property holdings (car, computer, TV), and health status(two body mass indices, diabetes rate). Column 1 shows summary statistics of the full sample, and the subsequent columns show summary statistics for the households surviving each elimination step.

B Price Index Construction

Purchases are observed in the raw data at the store-household-date-bar code level. To estimate the model, we need a price for every category at every store, date, and city. To aggregate from barcode level prices to grocery category level prices, we use an intermediate level called a grocery group. The definition of grocery groups is inherited from the KANTAR dataset, which defines comparable products. A barcode belongs to one and only one grocery group and a category. For example, 200 ml Head and Shoulders belongs to Shampoo, which belongs to the Self Care grocery category.

The price variable is constructed in two steps: We first aggregate individual bar code prices into grocery group-store prices. And then, we aggregate grocery group-store prices into category-store prices.

The described procedure requires a price for all relevant bar code-store combinations every week. However, not all bar codes are purchased every week in all stores. Thus, we impute unobserved prices using the following procedure:

1. Using the purchase data, we created dummies for all possible barcodes at the store-region-time combinations. We allow for region to be at either the city, region, or nation level and for time to be at either week, quarter, or year level. These combinations are possible imputation levels.

2. We regress observed prices on the obtained dummies. This fully saturated model can be seen as a non-parametric price estimator.

3. We sort imputation levels by predictive power (R^2).

4. We calculate each imputation level's average and median bar code price.

5. We impute the median price to each bar code at the corresponding price level of interest, e.g., a store-city-week unit observed in the data from the available level with the highest explanatory power.

B.1 From Bar Codes to Grocery Groups

The first step is obtaining a price for each grocery category in every store and weekly. Let p_{swybgc} be the price in chain s of bar code b that belongs to grocery group g and category c during week w of year y. In this step, the prices are obtained from the imputation algorithm and are aggregated using weights that aim to capture intra-category and intra-store preferences.

If all bar codes were sold in every chain, the weight ω_{bg} assigned to bar code b within the grocery group g would be defined by its share of total observed sales (measured by volume) among barcodes that comprise g. However, not all chains sell all barcodes. Thus, the weights are computed only among chains that sell barcode b. This means that if $V_{gs(b)}$ is the total volume of grocery group g observed sales in chains $s(b)$ that sold bar code b at least once and V_b is the total volume of bar code b sold. ω_{bg} is computed as $\frac{V_b}{V_{bs(g)}}$.

The previous procedure leads to weights that do not add up to one within chain-grocery group combinations. Thus, weights are normalized according to the following rule:

$$\hat{\omega_{bg}} = \frac{\omega_{bg}}{\sum\limits_{b \in B_{gs}} \omega_{bg}}$$

Where B_{gs} is the set of barcodes in grocery group g sold by chain S.

The price for grocery group g of category c at chain s in week w is computed as:

$$p_{wysgc} = \sum_{b \in g} \hat{\omega_{bg}} p_{swybgc}.$$

Finally, p_{wysgc} is normalized by the value of the price index during the first week we observe in our sample.

B.2 From Grocery Groups to Categories

In the second step, we calculate weights according to total observed expenditures per grocery group. Note that at this level, within-store preferences are not considered. Let $z_g c$ be the total observed expenditures in grocery group g of category c and let z_c be the total observed expenditures in category c. Then, the price for category c in chain s during week w of year y is obtained as:

$$P_{cswy} = \sum_{g \in c} \frac{z_g c}{z_c} P_{wysgc}.$$

In our data, we do not have a price for every grocery group and every store. In this case, we assume that the price was equal to the highest observed price for that product.

C Matching Stores to Households

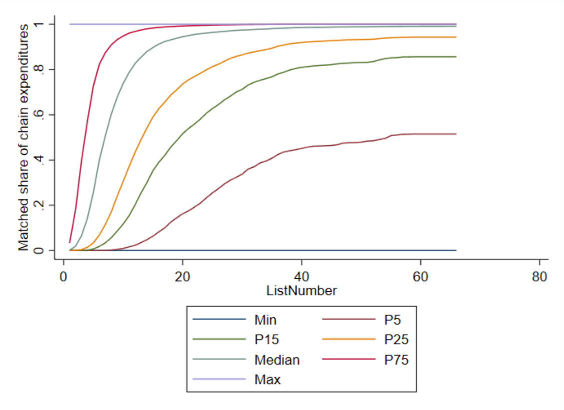
We use string search to find the location of stores in DENUE for all the chains that appear in the KANTAR dataset. The KANTAR dataset only records the chain when purchases occur at relatively large chains. This implies that a large proportion of households' expenditures can only be associated with a type of store but not a chain.

Our first step is to link chains in the KANTAR dataset to stores registered in DENUE. This first step is done using string searches.

Then, for every household in the KANTAR dataset, we created a list comprising all the stores within a 20-kilometer radius around the household's location. Then, for every household, we calculated the proportion of total observed chain expenditures captured by its associated list of stores. We found that 25 % of households had a share of matched chain expenditures below .3. This implied that for 25 % of households we were matching less than 65 % of their total observed expenditures to either traditional (unmatcheable) retailers or chains successfully linked to at least one store in DENUE.

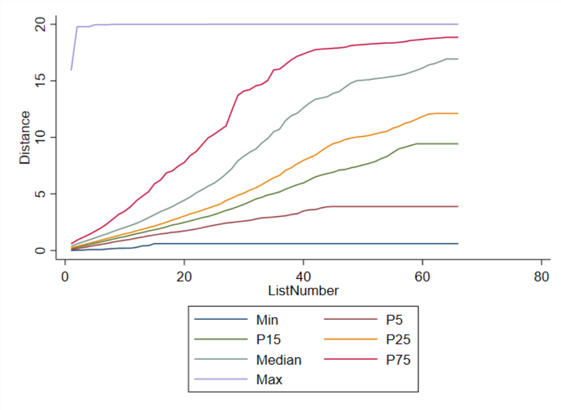
The main cause for these "pathological" households was that most appear only a few times in the purchase data (In less than three different weeks). Ignoring these households leads to a clear improvement in the quality of the match.

Figure A1: Quality of Match by Share of Chain Store Expenditures



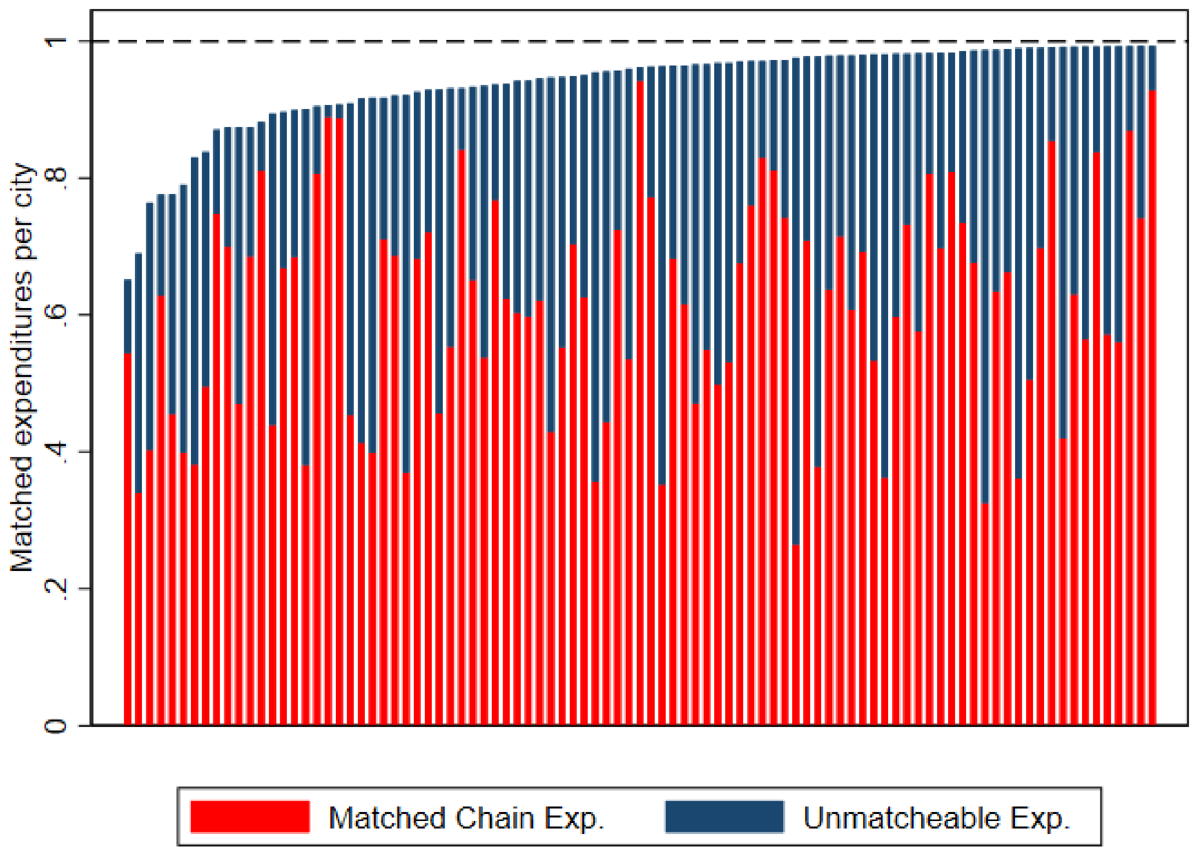
Note: The x-axis represents stores sorted increasingly by distance relative to each household. The y-axis represents the accumulated share of matched chain store expenditures. The sample is restricted to households whose purchases were observed for at least three consecutive weeks.

Figure A2: Quality of Match by Distance to Households



Note: The x-axis represents stores sorted increasingly by distance relative to each household. The y-axis represents the actual distance between stores and households. The sample is restricted to households whose purchases were observed for at least three consecutive weeks.

Figure A3



Note: The figure reports matched chain expenditures within 20 kilometers and traditional expenditures as a proportion of total observed expenditures at the city level. Each bar corresponds to a city.