Single-Threshold Food Labeling Policies

Nano Barahona Cristóbal Otero Sebastián Otero Joshua Kim[a]

Abstract: We examine the aggregate and heterogeneous effects of front-of-package single-threshold food labeling policies. Following the introduction of a regulation in Chile mandating such food labels, consumers reduced their overall sugar and caloric intake by 9% and 6%, respectively. These changes are explained by both consumers purchasing healthier products and firms reformulating their offerings. On the demand side, labels prompt consumers to substitute within categories rather than between categories. Such within-category responses are more pronounced when labels provide new information that contrasts with prior beliefs, highlighting the role of misinformation. On the supply side, we observe bunching at regulatory thresholds, although there is substantial heterogeneity across categories. These heterogeneous responses are consistent with differing costs of product reformulation. We conclude that taking into account heterogeneous supply- and demand-side responses to the policy is key for effective policy design.

JEL Codes: D12, D22, I12, I18, L11, L81

[a]This version: December 2023. An earlier version of this paper circulated under the title "On the Design of Food Labeling Policies." We want to thank Hunt Allcott, Liran Einav, Matthew Gentzkow and participants at Haas and the Virtual Quantitative Marketing Seminar for valuable comments and suggestions. We also thank Camila Corvalán and Marcela Reyes for beneficial conversations on institutional details; Christine Von Dessauer, Roberto Cases, and Romina Quagliotti for excellent research assistance; and Alejandro Guin-Po and Fernanda Mediano for their contributions to the data-collection process. Finally, we thank Walmart-Chile and Instituto de Nutrición y Tecnología de los Alimentos (INTA) and OK to Shop for sharing the data for the project, and the Stanford Center of Population Health Sciences (PHS) for providing a secure environment to store and analyze the data. Barahona: UC Berkeley, `nanobk@berkeley.edu`. C. Otero: Columbia Business School, `c.otero@columbia.edu`. S. Otero: Columbia University, `seb.otero@columbia.edu`. Kim: Facebook, `joshkim@fb.com`.

1. Introduction

The average American adult weighs nearly 23 pounds more today than in 1975 (NCHS, 2018). This dramatic rise in obesity is not specific to the United States: Over the same period, obesity in the world has tripled, and today, roughly 40% of the world's adult population is considered to be either obese or overweight (WHO, 2018). In response to this health pandemic, governments around the world are grappling with how to design policies that effectively improve diet quality.

An increasingly popular policy is to provide simplified information about products' healthiness to consumers through front-of-package (FOP) warning labels. These labels are simple symbols that clearly signal to consumers when a product is considered unhealthy based on whether targeted critical nutrients—such as sugar and calories—exceed certain concentration thresholds. These are single-threshold labels in the sense the regulatory thresholds are uniform across product categories and only make a distinction between solids and liquids. Chile was a pioneer country in implementing government-mandated FOP warning labels in the Chilean Food Act in 2016. Since then, more than 25 countries have approved or are considering similar regulations based on single-threshold binary labels.[1]

A policymaker faced with determining a single regulatory threshold for FOP warning labels encounters a complex task. To illustrate this challenge, assume two different product categories: cereal and pastries, with the latter having a much higher sugar content. The categories represent an equal share of consumer spending and are considered potential substitutes, as both serve as breakfast food items. The policymaker can establish the threshold for any level of sugar. Products to the right of the threshold receive a label, and products to the left remain unlabeled.

The choice of the threshold depends on several product features. First, it hinges on the extent to which labels are informative. If consumers are only uninformed about the sugar content of cereal but not of pastries, then the regulatory threshold should seek to increase the substitution from labeled to unlabeled cereal. Second, it also depends on the degree of substitution within and between categories. In the extreme case of perfect substitution between pastries and cereal, and assuming that labels are equally effective in both categories, the regulatory threshold should encourage substitution from pastries to cereal. Conversely, if there is minimal substitution across categories, the policy threshold should aim at curbing consumption of labeled products within pastries—the more sugary

[1] Countries that have passed similar regulations include Argentina, Canada, Chile, Colombia, Israel, Mexico, Peru, Uruguay, and Venezuela. See Table A.1 for the full list.

category. Third, the policymaker should also take into account the potential supply-side responses of the policy consisting of firms reformulating products to avoid labels. For instance, it may be easier for pastries than for cereals to reduce the amount of sugar.

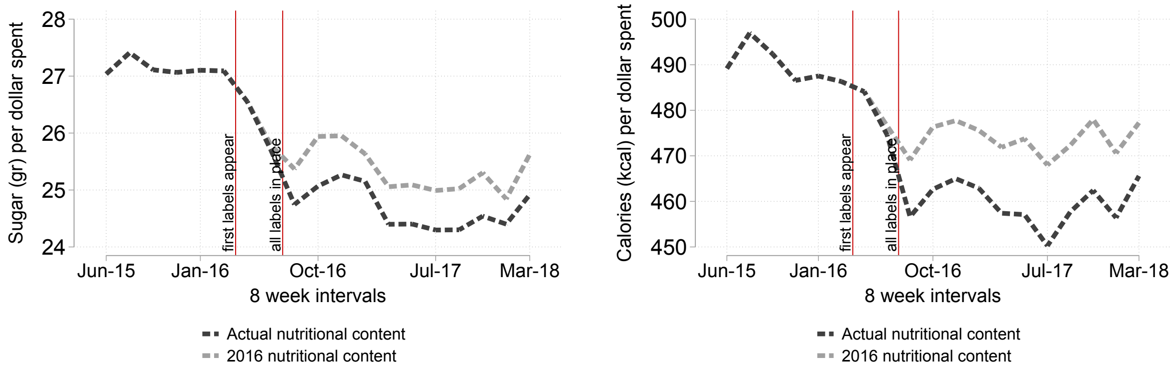
In this article, we review the Chilean experience and study the effects of food labeling policies across a wide range of product categories. In previous work, we constructed an equilibrium model to study the Chilean Food Act in a single category (Barahona et al., 2023). The study shows that the regulation was effective in reducing the consumption of unhealthy nutrients in the breakfast cereal category. This article builds on this study in two ways. First, we evaluate the policy's impact on overall nutritional intake. Second, we decompose the overall impact of the policy into supply- and demand-side responses and document substantial heterogeneity in substitution effects of labels and product reformulation across categories, which are key for effective policymaking.

To investigate the impact of the regulation across all product categories, we leverage access to Walmart's scanner data in Chile. The data contains the universe of food purchases made in Chile between 2015 and 2018. Walmart is the country's largest food retailer and accounts for more than 40% of supermarket sales. We combine these data with the product's self-collected nutrition facts tables before and after the policy. Finally, we use Walmart's loyalty program to follow consumers over time and produce individual-level measures of nutritional intake.

We start by documenting a sharp overall decrease in sugar and caloric intake of 9% and 6% per dollar spent, respectively, immediately after the policy was phased in. We present these findings in Figure 1. The reduction in sugar and caloric intake—which persists for the 2-year post-policy window in our data—is explained by a combination of demand- and supply-side responses. Consumers reacted to the regulation by making healthier choices, even when the nutritional content of products is kept constant over time (dashed curves). Firms responded by reducing the concentration of critical nutrients in their products, thus offering a healthier bundle of products (the difference between solid and dashed curves).[2]

We decompose the demand-side effects on *between*- and *within*-category substitution. First, we study whether the food labeling policy has the potential to shift consumer demand between categories. To test for the presence of cross-category substitution effects, we compare categories with different shares of labeled products and examine whether categories with a low share of labeled products increased their revenue relative to categories with a high share of labeled products. We find that the extent to which consumers substi-

[2] In Appendix A, Figure A.1, we show figures that divide nutritional intake by the volume of food purchased instead of by dollars spent. Overall, the findings and key takeaways are similar to the analysis above.



(a) Sugar intake

(b) Calorie intake

Figure 1: Nutritional intake per dollar spent before and after the policy

Notes: We produce this figure using a panel of Walmart consumers and computing their sugar and caloric intake during every 8-week period. After the labeling policy was introduced, total sugar intake decreased from 27.3 to 24.9 grams of sugar per dollar, and total caloric intake decreased from 488 to 457 kcal per dollar. The solid curve represents the total amount of sugar or calories purchased for every dollar spent in every 8-week period. The dashed curve is constructed in the same way as the solid curve, but fixing products' nutritional content at their 2016 values. The left vertical line corresponds to when the first labels appeared, and the right vertical line corresponds to when the Food Act became mandatory. We have two snapshots of nutritional information data: one from early 2016, before the policy was introduced, and one from 2018, after the policy was introduced. We assume that all changes in nutritional content occurred around the date of policy implementation (June 2016), and thus use these two snapshots for all pre-policy and post-policy nutritional values, respectively, in our calculations.

tuted between categories due to the presence of labels is negligible and cannot explain the patterns we document in Figure 1. The lack of between-category substitution is consistent with the limited demand effects observed in cross-category price promotions (Leeflang and Parreo-Selva, 2012). This phenomenon could be attributed to several factors: the spatial arrangement of cross-category products in stores (Kan et al., 2023), the pre-planned purchasing behavior targeting certain categories, often referred to as "destination categories" (Briesch et al., 2013), and the general stickiness and persistence of dietary habits (Hut and Oster, 2022; Kluser and Pons, 2023).

Next, we document important *within*-category substitution across labeled and unlabeled products in several product categories. We compare the quantities sold of labeled and unlabeled products before and after the introduction of the policy. We focus on categories for which there is enough variation in the share of labeled products and for which labeled and unlabeled products followed similar trends before the implementation of the policy. We document substantial heterogeneity in these substitution effects across different categories, spanning from 26% for cereals to 10% for soft drinks.

To understand the source of this heterogeneity, we implemented a survey in which we elicited consumers' beliefs about the nutritional content of soft drinks and cereal, the two categories with the smallest and largest substitution effects, respectively. We find that beliefs are very accurate for soft drinks but not for cereal. In line with Barahona et al. (2023), these results are consistent with the idea that food labels are more effective in shifting demand in categories in which labels are more informative.

On the supply side, firms might respond to labeling policies by reformulating their products and avoiding labels. To empirically assess these responses, we compare the distribution of nutritional content before and after the policy's implementation. We document a significant amount of bunching at the regulatory threshold in several product categories, with important heterogeneity across product categories. For instance, whereas virtually all products above the regulatory thresholds were reformulated in the yogurt category, only 5% of ex-ante labeled cookies changed their nutritional composition.

Product reformulation is more likely when the demand effects of labels are larger, when the threshold is close to the original nutritional content, and when reformulation costs are lower. For example, in categories such as yogurt or juice, firms can reformulate their products by substituting sugar with other low-cost sweeteners that mimic the products' taste. On the other hand, in cereal or cookies, sugar serves as a bulking agent, and replacing it with low-cost sweeteners may cause them to crumble. Consistent with this pattern, we also find that as the regulatory thresholds became tighter, the supply side responded by sequentially bunching at the stricter thresholds, but only in categories where reformulation was feasible at a low cost.

Our findings are important for policy design. First, the lack of between-category substitution implies that policymakers need to set regulatory thresholds to maximize the effects within specific categories. Second, the threshold for each nutrient should be set such that it targets categories that represent a large share of consumers' bundles in terms of the overall intake of that nutrient. In our Walmart dataset, categories representing 64% of spending had virtually all of their products labeled or none labeled. This highlights the crucial importance of the regulatory threshold selection, as it results in many categories lacking any significant indicators to assist consumers in making more informed purchasing decisions. Third, these categories need to have both healthy and unhealthy products that are close substitutes, and consumers must be misinformed about the healthiness status of those products. Fourth, the optimal threshold should also consider the extent to which reformulation is feasible at a low cost in the targeted categories.

This paper adds to the literature that investigates the effect of food labeling regulations on the demand for food. Most of this work has focused on specific food categories such as salad dressing (Mathios, 2000); microwave popcorn (Kiesel and Villas-Boas, 2013); sugar-sweetened beverages (Taillie et al., 2020); cheese and yogurt (Allais et al., 2015); ready-to-eat breakfast cereal (Zhu et al., 2015); and chain restaurants (Wisdom et al., 2010; Bollinger et al., 2011; Finkelstein et al., 2011). We add to this literature by providing evidence of and quantifying the effects of a national food labeling regulation on overall consumption of sugar and calories across all product categories.

Several other studies have also examined the Chilean Food Act. The main overarching finding is that labels induce demand effects within a category. Taillie et al. (2020) document a significant decline in purchases of labeled beverages following the policy's implementation. Araya et al. (2022) take advantage of the staggered introduction of labeled products to store inventories and find that—in the very short run—labels decrease demand in the breakfast cereal category, but not for chocolates or cookies. Alé-Chilet and Moshary (2022), Pachali et al. (2022), and Barahona et al. (2023) study the effects of the policy on breakfast cereal and find strong substitution from labeled to unlabeled products. Our paper contributes to these studies by focusing on substitution between categories and within multiple categories, and by rationalizing these heterogenous effects.

We also contribute to the literature that documents product reformulation responses to nutritional information policies (Unnevehr and Jagmanaite, 2008; Moorman et al., 2012; Griffith et al., 2017; Lim et al., 2020; Alé-Chilet and Moshary, 2022). In the context of the Chilean regulation, Reyes et al. (2020) and Quintiliano Scarpelli et al. (2020) show a reduction in critical nutrient concentration of multiple products after the policy's implementation. Alé-Chilet and Moshary (2022) and Barahona et al. (2023) provide evidence of bunching just below regulatory thresholds in the cereal market. Relative to these studies, we document important heterogeneity across categories in supply-side responses and discuss the drivers of these differences, and how they matter for policy design. We also add to these papers by looking at supply-side responses in the longer run after the most strict regulatory thresholds were in place.[3]

Our paper advances the debate on mandatory information disclosure and its potentially heterogeneous effect across categories and population groups (Cawley, 2015; Araya et al., 2022).[4] We show that food labeling policies can help to improve nutritional intake for consumers who do not respond to the labels via the product-reformulation channel. However, their effectiveness varies by product category and must be combined with com-

[3]As shown in Table A.2, the regulatory thresholds were gradually tightened as part of the policy implementation.

[4]It is worth noting that various FOP alternatives might exhibit different levels of effectiveness. See Dubois et al. (2021) and Ravaioli (2021) for a discussion.

plementary policies such as sugar taxes.

Finally, we contribute to a broader literature that studies how governments can help consumers make better nutritional choices. Allcott et al. (2019) examine whether improving access to healthy food in poor neighborhoods can decrease nutritional inequality; Dubois et al. (2017) analyze the effect of advertising on junk food consumption; and several other papers study the effects and design of taxes for sugar-sweetened beverages and calorie-dense food products (Falbe et al., 2015, 2016; Taylor et al., 2019; Silver et al., 2017; Dubois et al., 2020; Allcott et al., 2019a; Aguilar et al., 2021; Lee et al., 2019). Our paper focuses on an increasingly popular food label policy and shows that they can be an effective tool to improve diet quality and combat obesity when carefully designed.

The remainder of the paper is organized as follows. Section 2 describes the setting and the data. In Section 3, we present empirical evidence of how labels impact overall nutritional intake through demand- and supply-side responses. We discuss policy implications and conclude in Section 4.

2. Setting and data

In recent years, many countries have introduced FOP labels to help consumers make healthier food choices. Unlike nutrition facts tables, FOP labels simplify nutritional information, which makes it easier to use and interpret in a context in which shoppers make quick purchasing decisions (Temple, 2020). A recent development is warning labels that indicate whether a food product has a relatively high content of a critical nutrient, such as sugar, sodium, fat, or calories. Relative to other FOP labels, warning labels are simple binary symbols that clearly signal to consumers when a product has a high concentration of a given nutrient. Perhaps due to their simplicity, warning labels have become popular in the last few years. Following Chile's implementation in 2016, more than 25 countries, including Argentina, Brazil, Canada, Israel, and Mexico, have either implemented or are discussing the implementation of country-wide mandatory food labeling policies.[5]

Mandatory warning labels can be justified from both a demand- and supply-side perspective. In terms of consumer behavior, labels might help mitigate biases that drive consumers to over-purchase unhealthy products beyond their true preferences, such as lack of self-control, inattention to potentially harmful health effects, and poorly calibrated beliefs over products' nutritional content (Bernheim and Taubinsky, 2018; Allcott et al., 2019b; Barahona et al., 2023).

[5]Table A.1 presents the list of countries, including the critical nutrients targeted in the regulation.

Food labels can also incentivize firms to produce healthier products. In the absence of mandatory labels, firms do not have incentives to invest in reducing the concentration of critical nutrients in their products.[6] If consumers increase the demand for healthier products in response to labels, firms can benefit from reducing the concentration of regulated nutrients to avoid getting a label. In equilibrium, regulated markets with labels can induce consumers to make better nutritional decisions and firms to offer a healthier bundle of products (Barahona et al., 2023).

2.1. *The Chilean Food Act*

Chile was the first country to introduce a nationwide mandatory FOP warning-label policy.[7] In response to high rates of obesity, the most prevalent chronic disease in the country, in 2016 Congress passed Law 20.606 (hereafter, the Food Act), which introduced FOP warning labels to inform consumers about products' healthiness and help guide purchasing decisions.[8] The rationale was that nutritional information available in the form of a fact table on the back of the product was too complex and "did not allow [consumers] to make an informed decision" (Historia de la Ley 20.606, 2011, p. 170).

The Chilean Food Act mandated that products with calories, added sugar, saturated fats, and sodium higher than a given threshold must include a FOP warning label for each nutrient threshold surpassed. Figure 2 shows what Chilean FOP warning labels look like and how they are displayed on actual products. The thresholds were established uniformly for all food products, depending on whether the product is a solid or a liquid. To define the thresholds, the legislators chose the 90th percentile of the distribution of the concentration of critical nutrients from non-processed food products according to the USDA. The introduction of the thresholds was gradual and implemented in three stages. Stages 1, 2, and 3 took place in June 2016, 2018, and 2019 respectively. Threshold values

[6]Although suppliers of healthy products have incentives to solve these inattention and informational problems in a well-functioning market, several frictions make it unlikely that the market itself will self-regulate and thus provide ground for government intervention. Dranove and Jin (2010) discuss several conditions that are required for markets to unravel.

[7]Chile was the first country to approve mandatory national FOP warning labels for food high in calories, added sugars, saturated fats, and sodium and to implement the warning labels for all processed food products. Before 2016, several countries had implemented voluntary FOP labels. For example, Sweden, Denmark, Norway, Lithuania, and Iceland used the Keyhole logo; The Netherlands, Belgium, and Poland used the Choices logo; Korea and the United Kingdom used traffic-light labels; and Singapore used the Healthier Choice Symbol. Finland implemented a mandatory warning label in 1993 but only for some products high in salt. Thailand introduced a mandatory GDA label in 2007 but only for five categories of snacks. Also, Ecuador and Iran implemented mandatory traffic-light labeling for all processed products in 2014 and 2015, respectively.

[8]The Food Act also included regulations to ban selling labeled products in schools and a ban on advertising labeled products aimed at children younger than 14 years of age.

are presented in Appendix A, Table A.2.[9]

Figure 2: FOP warning labels on selected products

Notes: The figure presents both the FOP warning labels implemented in Chile and how these are displayed on various food packages. The labels say, from left to right, "High in sugar," "High in saturated fat," "High in sodium," and "High in calories." Products can have from zero to four labels. Table A.2 presents the threshold values that determine the assignment of each label.

2.2. *Data*

The main outcome measures used in the analysis come from scanner-level data from Walmart, which we expand using nutritional information and a survey that captures consumers' beliefs about product healthiness.

2.2.1. *Walmart data*: We use data from Walmart-Chile, the largest food retailer in Chile, responsible for over 40% of supermarket sales. The data covers all Walmart store transactions between May 2015 and March 2018, and identifies products by a Universal Product Code (UPC). For each transaction, we have access to information such as a product's price, revenue, product name, brand name, and discounts.[10]

We use Walmart's loyalty program to connect transactions with individual shoppers over time. We focus on regular Walmart customers who visited a store at least once every 8 weeks during the study period, leading to a total of roughly 524,000 individuals. We

[9]The Food Act has a relevant exception that limits its scope to regulating only processed and packaged foods. This means that even if a product exceeds a certain threshold, it will not require a label if it does not contain added sugars, sodium, saturated fats, honey, or syrup.

[10]The data comprise over 9 billion transactions by over 5 million consumers for over 20,000 different food products.

have access to information on these customers, including their gender, age, and household income. The average customer in our study is 48 years old, and 69% are women. Before the policy was introduced, the median customer shopped at Walmart 24 times, at three different Walmart locations, and traveled about 3 kilometers to get to the nearest store.[11]

2.2.2. *Nutritional Information*: The nutritional data for packaged products come from three sources: (a) pre-policy data collected by the Institute of Nutrition and Food Technology (INTA) at the University of Chile, (b) post-policy data we collected and digitized in 2018, and (c) longer-run post-policy data collected by "OK to Shop" that is up-to-date for 2021.

INTA collected nutritional information for a sample of products in January 2016 at the UPC level. This included the nutrition facts table, whether the product is a liquid or a solid, and the package size. We developed a phone app that linked images of nutrition facts tables to UPC codes to collect nutritional information for the post-policy period. We then digitized the nutritional content of all available products in the three largest Walmart stores in Chile. The product assortment in our dataset includes 90% of Walmart's revenue from packaged food products. We collected this information in March 2018, two years after the first stage of the labeling law was implemented in June 2016. For nutritional information data after the third stage, we partnered with "OK to Shop", a startup founded in 2018 that collects detailed information on food products. We accessed 2021 data for all their available products and matched it at the UPC level with nutritional information from previous years.

To include information on non-packaged products, such as fresh produce or meat—which don't have nutrition facts tables—we consulted publicly available data from the USDA's FoodData Central. We used these data to complete any missing information on critical nutrients across all food categories.

2.2.3. *Consumer beliefs*: We surveyed 1,500 consumers to elicit beliefs about the nutritional characteristics of packaged food products. We asked participants to provide an estimate of the sugar and caloric content of certain cereal and soft drink products. We conducted the survey in Argentina in August 2019, when there was no food labeling policy in place.[12]

[11]We count as a visit anytime a customer spends at least $20 on food products.

[12]Although we would have preferred to conduct the survey in Chile before the policy's implementation, we used Argentina as a proxy due to its similar population and food market, but lack of exposure to any labeling policy.

3. Empirical evidence

This section provides evidence of demand- and supply-side responses to food labels. We investigate the degree to which labels prompt consumers to substitute products within and between categories, the importance of the accuracy of consumers' beliefs for the effectiveness of food labels, and the extent to which products have been reformulated to avoid being labeled.

3.1. *Demand-side responses*

3.1.1. *Between-category substitution*: We start by examining whether consumers shift consumption across food categories as a response to food labels. To do so, we define broader groups of products that contain multiple categories in which we could expect substitution to occur. For instance, we check whether there was substitution between categories for product categories that are likely to be eaten at breakfast: eggs, yogurt, bread, fruits, jams, and breakfast cereal. Then, within each broader group, we compare revenues before and after the policy for categories with a high and a low share of labeled products.

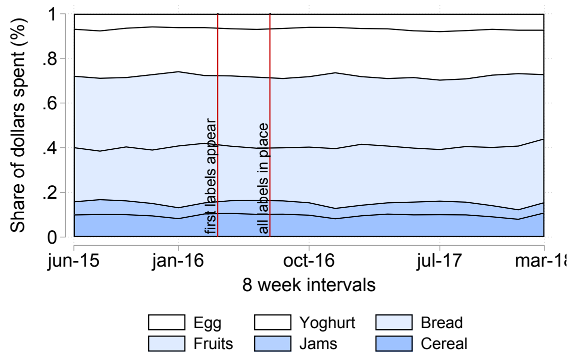
In Figure 3, we plot changes in the share of revenue over time of the food categories that fall into the breakfast and drinks food groups. In each group, categories are ordered from top to bottom according to the share of labeled products they contain (weighted by pre-policy revenue). The darker the area's color, the larger the share of labeled products. For instance, in breakfast products, 0% of egg products are labeled, while in cereals—the category with the highest share of labeled products in this group—62% of the products are labeled.

Figure 3 suggests there is little to no evidence that consumers are shifting consumption from highly labeled categories, such as breakfast cereals or soft drinks, to low-labeled categories, such as eggs or juices. For instance, in Panel (a), the share of breakfast spending on cereals averages 9.9% in both the pre- and post-policy period. In Appendix A Figure A.2, we show that this finding extends to several other food groups, such as carbs, meats, desserts, and snacks.

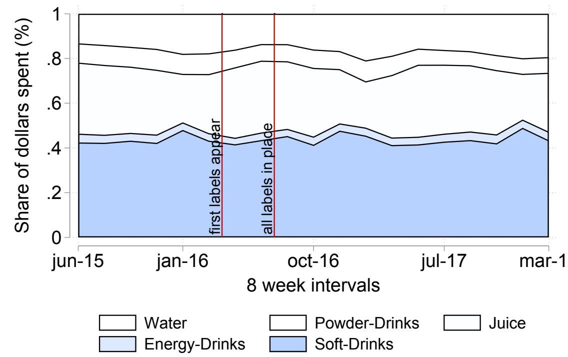
To formalize these results, we pool all food categories together (not only breakfast and drinks) and run the following regression:

$$\log(r_{cst}) = \beta_t \cdot L_c + d_{cs} + \delta_t + \varepsilon_{cst}, \tag{1}$$

where r_{cst} denotes the total revenue from products in category c sold in store s in pe-



(a) Breakfast



(b) Drinks

Figure 3: Share of dollars spent across categories

Notes: The figure shows the evolution of the share of dollars spent in each category within broader groups of products. Colors represent the share of labeled products within each category. White areas are categories in which no product received a label, and dark-blue areas are categories with a high share of labeled products. We show no differential changes in dollars spent between low-in-labels and high-in-labels categories.

riod t, and L_c is the (weighted) share of products in category c that have at least one label. Finally, δ_t denotes period fixed effects and d_{cs} refers to category-store fixed effects. We normalize the β_t coefficient corresponding to the first period post-adoption to zero. Observations are weighted by category-store pre-policy revenues, and standard errors are clustered at the category level.

Figure 4 displays the results from estimating Equation (1). We find that β_t estimates are small in magnitude, not statistically significant, and revolve around zero after the policy is in place. Reassuringly, we do not observe that categories with different shares of labeled products are differentially trending before the policy after accounting for store and year fixed effects. Regression results are consistent with the results for breakfast products and drinks presented in Figure 3. Overall, these results suggest that the extent to which consumers substituted toward other categories due to the presence of labels is negligible and cannot explain the patterns we document in Figure 1.

A caveat of our analysis is that the no interference assumption–standard in these types of research designs–does not hold. Naturally, a decrease in demand for unlabeled products will likely increase purchases of unlabeled ones. Since all markets are treated with the policy, no category or product can be used as a clean control group. This complicates the interpretation of the estimated coefficients. Nevertheless, the lack of significant coefficients still allows us to reject that consumers are shifting consumption from high- to low-labeled categories.

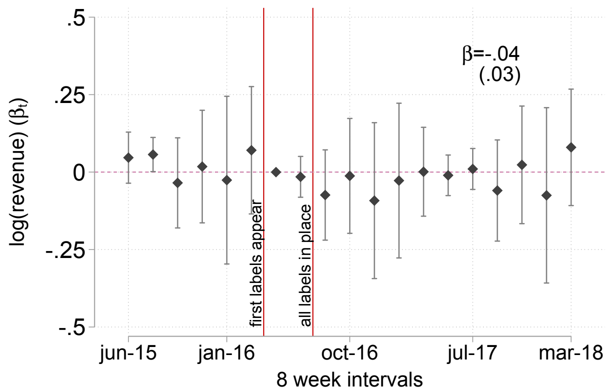


Figure 4: Changes in total spending per capita across categories

Notes: The figure presents the β_t coefficients from Equation (1). These regressions are run on a sample of 69 categories. The average share of labeled products in each category is 0.3, with a minimum of 0 and a maximum of 1. Vertical lines delimit the 95% confidence intervals of a test against the normalized coefficient (May 2016). Since different categories are affected by seasonality differently, coefficients closer to the winter season in Chile (May-July) are more precisely estimated when we compare them to the normalized coefficient.

3.1.2. *Within-category substitution*: Next, we examine the effects of food labels within each food category, where products are more likely to be close substitutes. For our analysis, we limit our attention to eight categories and subcategories that meet the following criteria: (a) products are sufficiently similar such that consumers would consider substituting from one to another as a result of the regulation, (b) there is sufficient variation in terms of the share of products that received a label, and (c) unlabeled and labeled products within the category follow similar pre-trends in the absence of the policy. We select eight categories that represent 5.7% of the pre-policy revenue of all food products and 11.9% of the pre-policy revenue of all labeled products. In Appendix B, we explain in more detail the selection process for these categories and sample coverage.

We define a product as the union of UPCs that share the same product name and brand. For example, we assign all *Diet Coke* the same product ID regardless of their can or bottle size. We assign labels to a product based on its 2018 nutritional content. We collapse our original data into product-store-period data bins (in which a period is defined as 8 consecutive calendar weeks) and estimate the following regression for each category:

$$\log(q_{jst}) = \beta \cdot L_{jt} + \gamma \cdot \log(p_{jst}) + \delta_{js} + \delta_t + \varepsilon_{jst}, \tag{2}$$

where q_{jst} denotes the grams (ml) of product j sold in store s in period t, p_{jst} refers to the product's price per 100 grams (ml), and L_{jt} is an indicator variable that takes the value of one if the product has one or more labels during the post-policy periods. Finally, δ_{js} refers to product-store fixed effects and δ_t to period fixed effects. Observations are weighted by product-store pre-policy revenues. Products that do not appear in the pre-period have zero weight and are thus excluded from the estimation sample. Standard errors are clustered at the product level.

In Figure 5, we plot the estimated changes in demand for each of the categories. We find a significant impact of the policy for all eight categories. These results confirm that consumers substituted from labeled to unlabeled products within all of these product categories and that the effect holds in the medium run. Overall these findings suggest that the policy is effective in shifting consumption in many product categories and thus complements the policy impact for breakfast cereal documented by Barahona et al. (2023).[13] Interestingly, the effects are highly heterogeneous by product category, ranging from 26% in the case of cereals to 10% in the case of soft drinks.

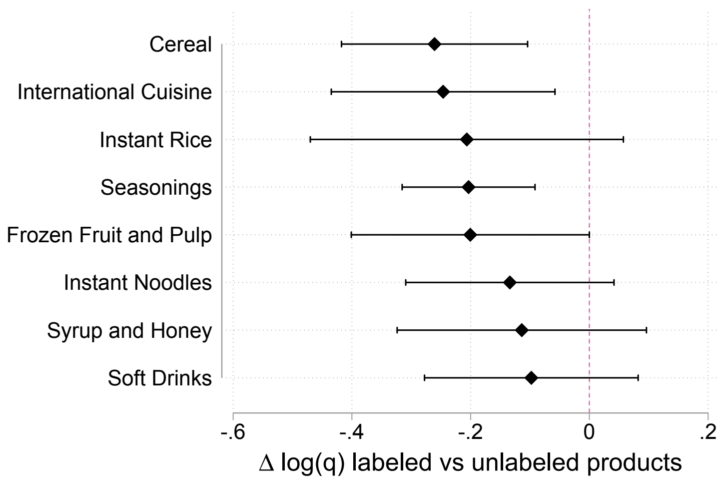


Figure 5: Changes in demand for selected categories

Notes: This shows the changes in equilibrium quantities between labeled and unlabeled products from estimating Equation (2) for different product categories. The set of products in this analysis represents 11.9% of the pre-policy revenue of labeled products in the sample. We provide more details on sample selection in Appendix B.

As in the previous analysis, consumers substitute from one product to another, and the no-interference assumption does not hold. In the extreme case of one-to-one substitution, a β of 10% would reflect a 5% decrease in labeled products and a 5% increase in unlabeled

[13] In Appendix A Figure A.3, we present nonparametric results. Pre-period coefficients are small and not significantly different from zero in all categories. After the introduction of the labels, we observe a noticeable drop in demand for labeled products relative to that for unlabeled products. This effect persists throughout our observational time window.

products. As a result, our coefficients should be interpreted as the before-after impact of the policy on the relative differences in equilibrium quantities between labeled and unlabeled products and not as the direct effect of labels on purchases of labeled products.

A second caveat of this analysis is that prices are endogenous, and including them as controls can complicate the interpretation of the coefficients even further. In the current specification of Equation (2), β accounts for the change in the relative differences in equilibrium quantities between labeled and unlabeled products before and after policy *conditional* on the observed prices. Omitting prices from Equation (2), β would capture changes in the relative differences in equilibrium quantities induced both by the labels and the endogenous changes in prices. In Appendix A, Figure A.4, we show that results omitting prices from Equation (2) look qualitatively similar.

Next, we explore a mechanism that can help make sense of the dispersion in observed effects.

3.1.3. *The importance of prior beliefs*: In related work, Barahona et al. (2023) show that the effect on cereals is mostly explained by substitution away from products that consumers believed to be healthy but which ended up with a label. In other words, at the product level, labels are more effective if they provide new information to consumers. To investigate how information and beliefs affect the extent of *within*-category substitution, we use the beliefs survey described in Subsection 2.2.3 and compare the effects of the policy on cereal and soft drinks.

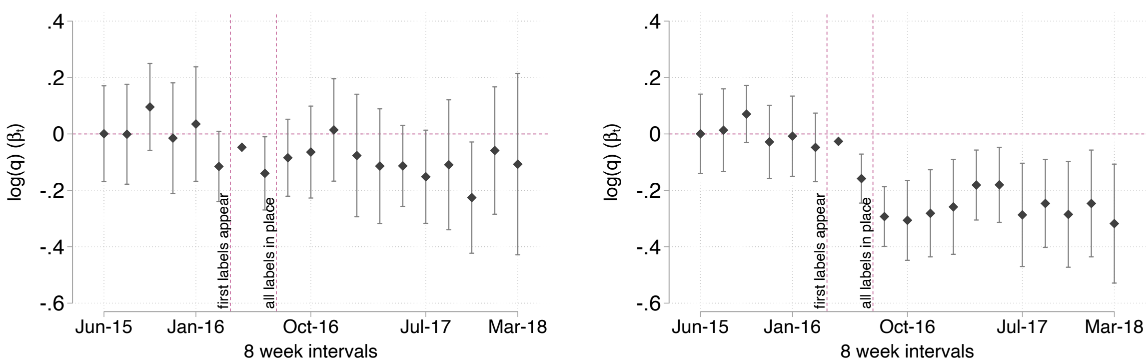
The pre-policy concentration of sugar in soft drinks follows a bimodal distribution driven by diet and non-diet drinks, which highly correlates with consumers' beliefs about sugar concentration in these products. The correlation between the average value of respondents' beliefs about the sugar concentration of each product and the product's observed pre-policy sugar concentration in the soft drinks category is 0.94. In cereal, however, consumers have mistaken beliefs about the caloric content of products. The correlation between the average value of respondents' beliefs about the caloric concentration of each product and the product's observed pre-policy caloric concentration is 0.23. We present the relationship between consumer beliefs and pre-policy nutritional content in Appendix A, Figure A.5.

The accuracy of beliefs about sugar content implies that labels came as no surprise in this category, which means that the effect of the policy should be smaller in soft drinks than in cereal. To test for this, we estimate the following regression for each category:

$$\log(q_{jst}) \;=\; \sum_k \beta_k \cdot L_j \cdot \mathbb{1}\{k=t\} + \gamma \cdot \log(p_{jst}) + \delta_{js} + \delta_t + \varepsilon_{jst}, \tag{3}$$

where all variables and specification details are defined as in Equation (2).

Figure 6 displays the results of estimating Equation (3) for soft drinks and cereal. Pre-period coefficients are small and not significantly different from zero in both categories. After the introduction of the labels, we observe a noticeable drop in the demand for labeled products relative to that for unlabeled products in the cereal market but not in the soft drinks market.



(a) Soft Drinks (b) Cereals

Figure 6: Dynamic changes in demand for soft drinks and cereals

Notes: This figure presents regression coefficient estimates from Equation 3. Panel (a) presents the β_k coefficients of the regression on the sample of 37 soft drinks that show up in the pre-and post-policy periods. The sample consists of 27 unlabeled and 10 labeled products for a total of 94,391 observations. Coca-Cola Company reformulated two products—Regular Fanta and Regular Sprite—to remove their high-in-sugar label in September of 2017. We define those products as unlabeled in Equation 3 (i.e., $L_j = 0$). Panel (b) displays the β_k coefficients of the regression on the sample of 68 ready-to-eat cereals that show up in the pre-and post-policy periods. The sample consists of 27 unlabeled and 41 labeled products for a total of 194,510 observations.

The main takeaway is that labels have a larger *within*-category effect when they provide useful information to consumers. This is consistent with the evidence provided by Araya et al. (2022), who study the short-term effects of the same regulation for breakfast cereals, chocolates, and cookies. They exploit the staggered rollout of labels across different stores before the law went into effect, which allows them to study the effect of labels in categories in which all products would have received a label after full deployment of the law. They find a null effect of food labels in chocolates and cookies, in which labels are likely to offer no new information to consumers.

3.2. *Supply-side responses*

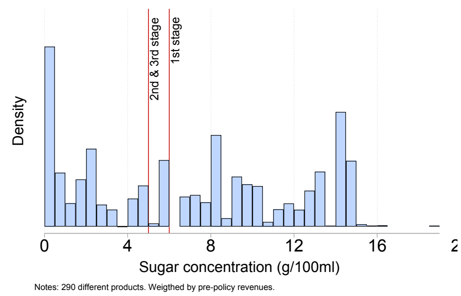
Finally, we examine supply-side responses to the regulation. We look at the extent to which firms reformulated their products in different categories due to the introduction of food labels. We focus on reformulation in sugar and restrict our attention to categories in which the distribution of sugar concentration is neither entirely to the left of the regulation threshold nor too far to the right, such that it would not be feasible to modify the nutritional content up to threshold levels. This gives us a total of 13 categories that represent 15.5% of the pre-policy revenue of all food products and 53.6% of the pre-policy revenue of products to the right of the threshold for which we collected nutritional content data. We discuss further details of the selection of categories and sample coverage in Appendix C.

To assess behavioral responses on the supply side, we compare the distribution of nutritional content before and after the policy was implemented and explore bunching at regulatory thresholds. Figure 7 plots the distribution of sugar concentration in 2016 (pre) and 2018 (post) for products in the juice and cereal categories. The size of the bars represents the pre-policy revenue of the products included in that bar. Each subfigure also includes three vertical lines that indicate the thresholds of the policy in each of the three stages.

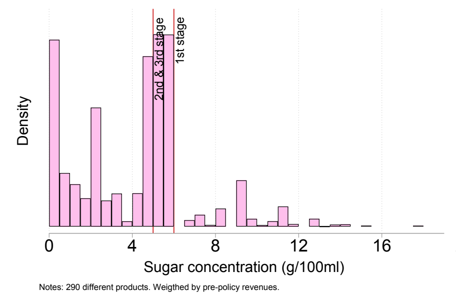
Before the introduction of food labels, we do not observe any noticeable pattern of bunching at any of the thresholds in either the juice or cereal category. However, in 2018 we find some evidence of bunching in both categories. In the juice category, we find that whereas in 2016, more than half of the products had a sugar concentration per 100 ml above the first-stage threshold, the distribution shifted to the left, and most products avoided first-stage labels. In cereal, we also find that some products were reformulated in order to be on the left side of the threshold. Nevertheless, reformulation occurred to a much lesser extent than in juice.

In Figure 8, we summarize the findings for all categories by plotting the (weighted by pre-policy revenue) share of the products in each category that surpassed the sugar threshold in the pre-policy period and were reformulated to be to the left of the threshold in the post-policy period. We show histograms for each category in Appendix A, Figures A.6 and A.7.

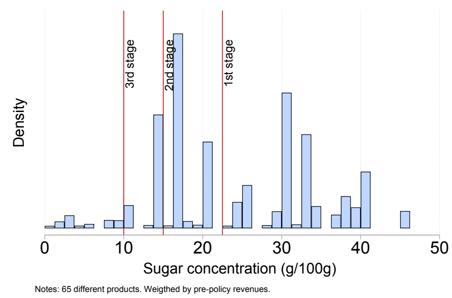
We find that while in some categories, 100% of the products were reformulated to cross the regulatory threshold, in other categories, less than 10% were reformulated. Three important features of a product category can affect the extent to which products are reformulated. First, firm responses depend on the expected impact of labels on product



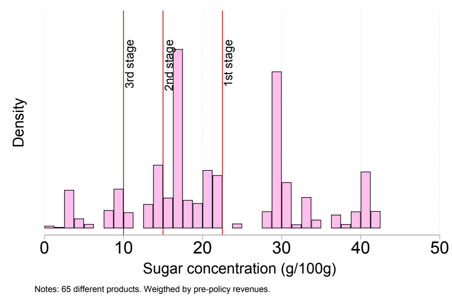
(a) Juice 2016



(a) Juice 2018



(b) Cereal 2016



(b) Cereal 2018

Figure 7: Distribution of sugar content pre- and post-policy in selected categories

Notes: This figure plots the distribution of sugar concentration for juice and cereal before and after policy implementation. Observations are weighted by pre-policy revenue.

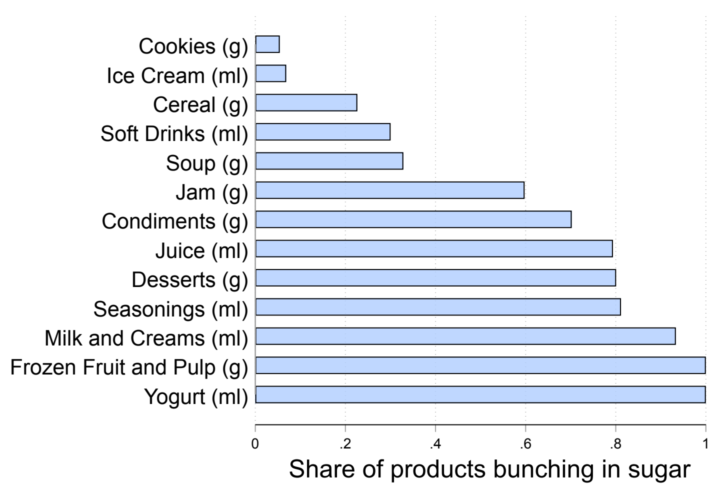


Figure 8: Share of products bunching in sugar

Notes: This figure summarizes the findings for bunching from Appendix A, Figures A.6 and A.7. It shows the pre-policy revenue-weighted share of products to the right of the threshold in sugar in the pre-policy period that reduced the concentration of sugar to be to the left of the regulatory threshold in the post-policy period. The products used represent 54% of the pre-policy revenue of all products to the right of the policy threshold in the pre-policy period. We provide more details on sample selection in Appendix C.

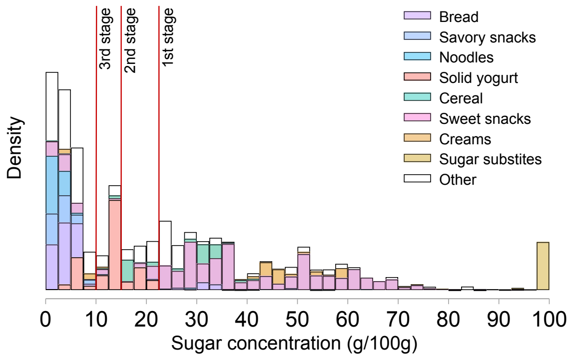
demand. In categories with close substitutes and in which labels can provide more information, the returns to reformulation are higher. Second, reformulation is a function of the distance between the products' current nutritional content and the regulatory threshold. Third, firms are more likely to reformulate products when they are able to do so without substantially affecting their quality (e.g., taste). For example, in categories such as yogurt or juice, firms can reformulate their products by substituting sugar with other low-cost sweeteners that mimic the products' taste. On the other hand, sugar serves as a bulking agent in cereal or cookies, and replacing it with low-cost sweeteners may cause them to crumble. Our results are consistent with this pattern, whereby categories with liquid products present a larger share of products that bunch.

3.2.1. *Long-run Supply-side Responses*: To study product reformulation following the implementation of the most stringent policy threshold, we analyze nutritional information data from 2021. This data is merged at the UPC level with data from 2016 (pre-policy) and 2018 (post-first stage). We focus on products (UPCs) for which nutritional information is available for all three periods, covering 44% of the pre-policy revenue for all solid products and 81% of the pre-policy revenue for all liquid products.

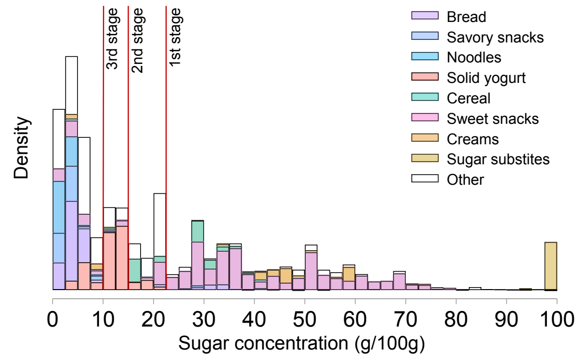
Figure 9 presents the results for solids. Panel (b) shows that, consistent with our findings across individual categories, some products that were close to the right of the threshold before the policy tend to bunch at the first-stage threshold afterward. Panel (c) reveals significant bunching at the third-stage threshold post-2021. Interestingly, much of this bunching is attributed to yogurts, which are occasionally categorized as solids rather than liquids; these are presumably easier to reformulate than other solid products.

In Figure 10, we examine the regulatory responses of liquid products. Note that prior to the policy implementation, multiple products in the milk category bunched at the second-stage regulatory threshold.[14] The reason is that the most stringent policy threshold was set at 5 grams of sugar per 100 ml, approximating the natural sugar content in whole milk. Once the first stage is enacted, the majority of the product distribution bunches to the left of either the first or second stage threshold. Panel (c) shows that after the second stage is implemented, most products bunch at that threshold. Notably, some soft drinks remain unchanged in formulation. This is because many of these products have sugar-free alternatives (e.g., *Coke* and *Diet Coke*). Overall, Figures 9 and 10 provide evidence of substantial reformulation in liquids and smaller responses in solids, particularly when considering that most of the reformulation in solids is driven by yogurts.

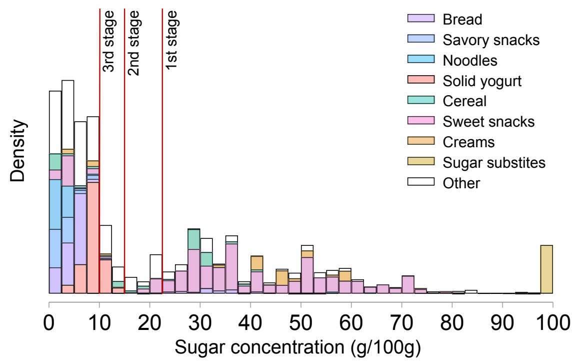
[14] For liquids, the third and second thresholds are the same, which is why we only label the second



(a) Pre-policy: 2016



(b) Post-policy (1st stage): 2018



(c) Post-policy (3rd stage): 2021

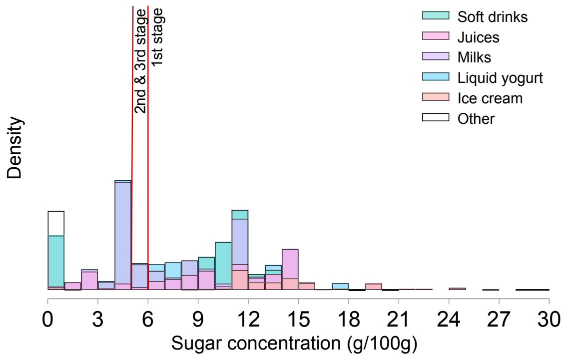
Figure 9: Distribution of sugar content pre- and post-policy for solid products

Notes: The figure shows the pre- and post-policy distribution of sugar concentration for all solid products (full bar) and selected categories (in colors). Panel (a) represents the pre-policy distribution, Panel (b) post-policy after the 1st stage of implementation, and Panel (c) after all stages were implemented. Product categories that are not covered by the law (e.g., fresh meat, fish, eggs, unpacked fruits and vegetables, beans, flour), that most do not include any added sugar (e.g. butter, rice, frozen meals) or that in some cases are sold packed and in others unpacked without labels (e.g., cheese, sausages) are excluded. Only products with nutritional information for the three periods are included - after matching by product UPC only. The products included are 44.27% of the pre-policy revenue of all solid products. Observations are weighted by pre-policy revenue. Stages 1, 2, and 3 in Chile took place in June 2016, June 2018, and June 2019.

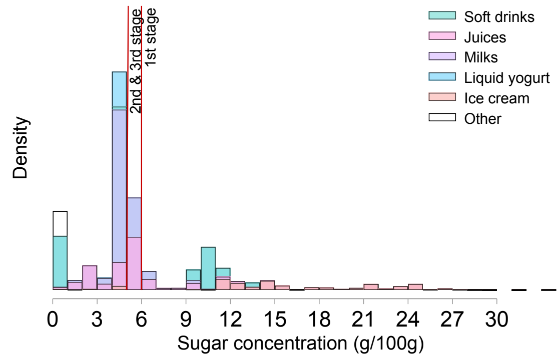
4. Discussion

The Chilean Food Act suggests that FOP warning labels have the potential to reduce the overall intake of calories and sugar. We use access to rich micro-data—the universe of Walmart transactions in Chile between 2015 and 2018—and perform several empirical exercises to unpack the mechanisms through which labels affect consumer and firm behavior to inform policy design. We find that labels are ineffective in shifting consumption across product categories. In other words, we do not find evidence of substitution from

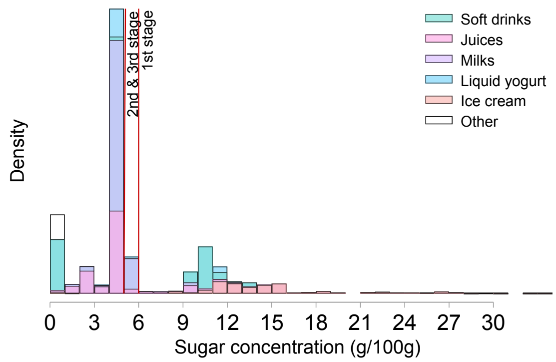
stage in the figure.



(a) Pre-policy: 2016



(b) Post-policy (1st stage): 2018



(c) Post-policy (3rd stage): 2021

Figure 10: Distribution of sugar content pre- and post-policy for solid products

Notes: The figure shows the pre- and post-policy distribution of sugar concentration for all liquid products (full bar) and selected categories (in colors). Panel (a) represents the pre-policy distribution, Panel (b) post-policy after the 1st stage of implementation, and Panel (c) after all stages were implemented. Product categories that are not covered by the law (e.g., alcohol, mineral water) or mostly do not include any added sugar (e.g. oils) are excluded. Only products with nutritional information for the three periods are included - after matching by product UPC only. The products included are 81% of the pre-policy revenue of all liquid products. Observations are weighted by pre-policy revenue. Stages 1, 2, and 3 in Chile took place in June 2016, June 2018, and June 2019.

unhealthy to healthy categories. Instead, most of the policy effects arise from substitution within a product category. We also find that labels are substantially more effective in product categories in which beliefs about product healthiness are poorly calibrated. Finally, we find that labels have the potential to promote product reformulation across several product categories, that these responses are highly heterogeneous, and that tighter regulatory thresholds implemented in later stages lead to increased reformulation.

These results are consistent with the predictions of the model of supply and demand for nutrients in Barahona et al. (2023). On the demand side, labels are most effective when they correct mistaken beliefs about product healthiness and when consumers are willing to substitute labeled and unlabeled products (i.e., whether labeled and unlabeled products are close substitutes). On the supply side, the model suggests that firms are more likely to reformulate their products and avoid being labeled when it is economically feasible to maintain taste consistency at a low cost.

Our findings provide insights to inform the design of effective food labeling policies. Policymakers seeking to implement labels need to decide where to set label thresholds to harness demand- and supply-side responses effectively. First, the lack of between-category substitution implies that food labels should be designed to focus on effects within specific categories. For instance, labels should not be designed to maximize substitution from cereal to fruit but instead from unhealthy to healthy cereal. Second, policymakers should target categories that (i) represent a large share of consumers' intake of critical nutrients, (ii) have both healthy and unhealthy products that are close substitutes, and (iii) in which consumers are misinformed about the health status of the products therein. Third, thresholds must be set to maximize substitutability within targeted categories and product reformulation, and should therefore consider the extent to which taste consistency can be preserved without significant increases in production costs.

It is important to bear in mind that the policy we examine in this study applied a uniform threshold for all solids and a uniform threshold for all liquids for each targeted nutrient. Originally, legislators in Chile considered introducing category-specific thresholds. A design with category-specific thresholds could work better as the regulator can take advantage of the heterogeneous demand- and supply-side responses across categories to maximize the policy effectiveness. Nevertheless, legislators ruled out such a design as it could be more challenging for consumers to understand, leading to undesirable perceptions about absolute healthiness levels across categories. It also requires policymakers to take a stance about which products belong to which categories. Understanding how category-specific thresholds can affect the effectiveness of the policy is an important area for future research.

Finally, our results also shed light on the importance of combining alternative policies to tackle obesity. In categories such as chocolates and candy, in which all products receive labels and are known to be high in critical nutrients, food labels are less effective for improving diet quality. Also, other market imperfections, such as lack of self-control or time inconsistency, may induce consumers to not always choose food products that are best for them (Sadoff et al., 2020; Samek, 2019). In those cases, a better policy tool may be to implement sugar taxes (Barahona et al., 2023). Consequently, because labels and sugar taxes counteract different internalities, they should be seen as complementary policies rather than substitutes.

References

Aguilar, A., E. Gutierrez, and E. Seira (2021). The effectiveness of sin food taxes: evidence from mexico. *Journal of Health Economics 77*, 102455.

Alé-Chilet, J. and S. Moshary (2022). Beyond consumer switching: Supply responses to food packaging and advertising regulations. *Marketing Science*, Forthcoming.

Allais, O., F. Etilé, and S. Lecocq (2015). Mandatory labels, taxes and market forces: An empirical evaluation of fat policies. *Journal of Health Economics 43*, 27–44.

Allcott, H., R. Diamond, J.-P. Dubé, J. Handbury, I. Rahkovsky, and M. Schnell (2019). Food deserts and the causes of nutritional inequality. *The Quarterly Journal of Economics 134*(4), 1793–1844.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019a). Regressive sin taxes, with an application to the optimal soda tax. *The Quarterly Journal of Economics 134*(3), 1557–1626.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019b). Should we tax sugar-sweetened beverages? an overview of theory and evidence. *Journal of Economic Perspectives 33*(3), 202–27.

Araya, S., A. Elberg, C. Noton, and D. Schwartz (2022). Identifying food labeling effects on consumer behavior. *Marketing Science*, forthcoming.

Barahona, N., C. Otero, and S. Otero (2023). Equilibrium effects of food labeling policies. *Econometrica Forthcoming*.

Bernheim, B. D. and D. Taubinsky (2018). Behavioral public economics. *In: Handbook of Behavioral Economics—Foundations and Applications 1*(5), 381–516.

Bollinger, B., P. Leslie, and A. Sorensen (2011, February). Calorie posting in chain restaurants. *American Economic Journal: Economic Policy 3*(1), 91–128.

Briesch, R. A., W. R. Dillon, and E. J. Fox (2013). Category Positioning and Store Choice: The Role of Destination Categories. *Marketing Science 32*(3), 488–509.

Cawley, J. (2015). An economy of scales: A selective review of obesity's economic causes, consequences, and solutions. *Journal of health economics 43*, 244–268.

Dranove, D. and G. Z. Jin (2010, December). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature 48*(4), 935–63.

Dubois, P., P. Albuquerque, O. Allais, C. Bonnet, P. Bertail, P. Combris, S. Lahlou, N. Rigal, B. Ruffieux, and P. Chandon (2021). Effects of front-of-pack labels on the nutritional quality of supermarket food purchases: evidence from a large-scale randomized controlled trial. *Journal of the Academy of Marketing Science 49*(1), 119–138.

Dubois, P., R. Griffith, and M. O'Connell (2017, 04). The Effects of Banning Advertising in Junk Food Markets. *The Review of Economic Studies 85*(1), 396–436.

Dubois, P., R. Griffith, and M. O'Connell (2020, November). How well targeted are soda taxes? *American Economic Review 110*(11), 3661–3704.

Falbe, J., N. Rojas, A. H. Grummon, and K. A. Madsen (2015). Higher retail prices of sugar-sweetened beverages 3 months after implementation of an excise tax in berkeley, california. *American journal of public health 105*(11), 2194–2201.

Falbe, J., H. R. Thompson, C. M. Becker, N. Rojas, C. E. McCulloch, and K. A. Madsen (2016). Impact of the berkeley excise tax on sugar-sweetened beverage consumption. *American journal of public health 106*(10), 1865–1871.

Finkelstein, E. A., K. L. Strombotne, N. L. Chan, and J. Krieger (2011). Mandatory Menu Labeling in One Fast-Food Chain in King County, Washington. *American Journal of Preventive Medicine 40*(2), 122–127.

Griffith, R., M. O'Connell, and K. Smith (2017). The importance of product reformulation versus consumer choice in improving diet quality. *Economica 84*(333), 34–53.

Historia de la Ley 20.606 (2011). Sobre composición nutricional de los alimentos y su publicidad. http://www.bcn.cl/obtienearchivo?id=recursoslegales/10221.3/37370/1/HL20606.pdf.

Hut, S. and E. Oster (2022). Changes in Household Diet: Determinants and Predictability. *Journal of Public Economics 208*, 104620.

Kan, C., Y. L. Liu, D. R. Lichtenstein, and C. Janiszewski (2023). The Negative and Positive Consequences of Placing Nonpromoted Products Next to Promoted Products. *Journal of Marketing 87*(6), 928–948.

Kiesel, K. and S. B. Villas-Boas (2013). Can information costs affect consumer choice? Nutritional labels in a supermarket experiment. *International Journal of Industrial Organization 31*(2), 153–163.

Kluser, F. and M. Pons (2023). The Apple Does Not Fall Far From the Tree: Intergenerational Persistence of Dietary Habits . Mimeo.

Lee, M. M., J. Falbe, D. Schillinger, S. Basu, C. E. McCulloch, and K. A. Madsen (2019). Sugar-sweetened beverage consumption 3 years after the berkeley, california, sugar-sweetened beverage tax. *American Journal of Public Health 109*(4), 637–639.

Leeflang, P. S. H. and J. Parreo-Selva (2012). Cross-category Demand Effects of Price Promotions. *Journal of the Academy of Marketing Science 40*(4), 572–586.

Lim, J. H., R. Rishika, R. Janakiraman, and P. Kannan (2020). Competitive effects of front-of-package nutrition labeling adoption on nutritional quality: Evidence from facts up front–style labels. *Journal of Marketing 84*(6), 3–21.

Mathios, A. D. (2000). The impact of mandatory disclosure laws on product choices: An analysis of the salad dressing market. *The Journal of Law and Economics 43*(2), 651–678.

Moorman, C., R. Ferraro, and J. Huber (2012). Unintended Nutrition Consequences: Firm Responses to the Nutrition Labeling and Education Act. *Marketing Science 31*(5), 717–737.

NCHS (2018). National health and nutrition examination survey data, national center for health statistics. *U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.*

Pachali, M. J., M. J. Kotschedoff, A. van Lin, B. J. Bronnenberg, and E. van Herpen (2022). How do nutritional warning labels affect prices? *Working Paper*.

Quintiliano Scarpelli, D., A. C. Pinheiro Fernandes, L. Rodriguez Osiac, and T. Pizarro Quevedo (2020). Changes in nutrient declaration after the food labeling and advertising law in chile: a longitudinal approach. *Nutrients 12*(8), 2371.

Ravaioli, S. (2021). Coarse and Precise Information in Food Labeling. Mimeo.

Reyes, M., L. Smith Taillie, B. Popkin, R. Kanter, S. Vandevijvere, and C. Corvalán (2020). Changes in the amount of nutrient of packaged foods and beverages after the initial implementation of the chilean law of food labelling and advertising: A nonexperimental prospective study. *PLoS medicine 17*(7), e1003220.

Sadoff, S., A. Samek, and C. Sprenger (2020). Dynamic inconsistency in food choice: Experimental evidence from two food deserts. *The Review of Economic Studies 87*(4), 1954–1988.

Samek, A. (2019). Gifts and goals: Behavioral nudges to improve child food choice at school. *Journal of Economic Behavior & Organization 164*, 1–12.

Silver, L. D., S. W. Ng, S. Ryan-Ibarra, L. S. Taillie, M. Induni, D. R. Miles, J. M. Poti, and B. M. Popkin (2017). Changes in prices, sales, consumer spending, and beverage consumption one year after a tax on sugar-sweetened beverages in berkeley, california, us: A before-and-after study. *PLoS Medicine 14*(4), e1002283.

Taillie, L. S., M. Reyes, M. A. Colchero, B. Popkin, and C. Corvalán (2020). An evaluation of chile's law of food labeling and advertising on sugar-sweetened beverage purchases from 2015 to 2017: A before-and-after study. *PLoS Medicine 17*(2), e1003015.

Taylor, R., S. Kaplan, S. B. Villas-Boas, and K. Jung (2019). Soda wars: Effect of a soda tax election on soda purchases. *Economic Inquiry 57*(3), 1480–1496.

Temple, N. J. (2020). Front-of-package food labels: A narrative review. *Appetite 144*, 104485.

Unnevehr, L. J. and E. Jagmanaite (2008). Getting rid of trans fats in the us diet: Policies, incentives and progress. *Food Policy 33*(6), 497–503. Food Product Composition, Consumer Health, and Public Policy.

WHO (2018). Factsheet No. 311. *World Health Organization*.

Wisdom, J., J. S. Downs, and G. Loewenstein (2010). Promoting Healthy Choices: Information versus Convenience. *American Economic Journal: Applied Economics 2*(2), 164–178.

Zhu, C., R. A. Lopez, and X. Liu (2015). Information Cost and Consumer Choices of Healthy Foods. *American Journal of Agricultural Economics 98*(1), 41–53.

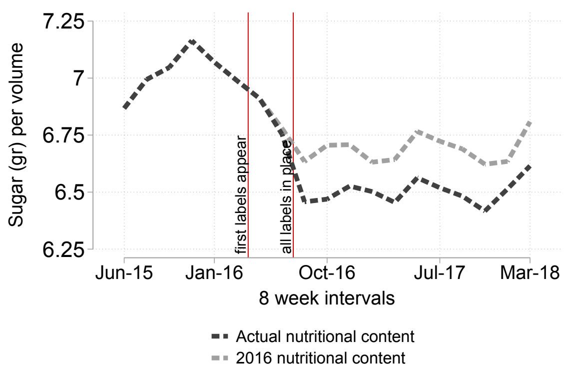
Online Appendix for:

On the Design of Food Labeling Policies

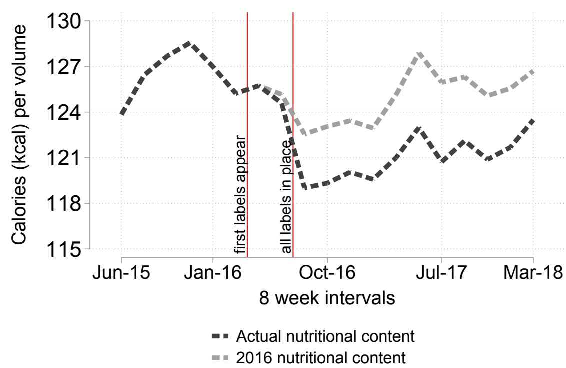
Nano Barahona Cristóbal Otero Sebastián Otero Joshua Kim

December 6, 2023

Appendix A: Additional Figures



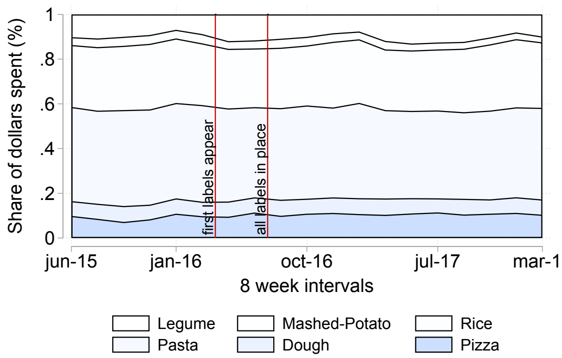
(a) Sugar intake



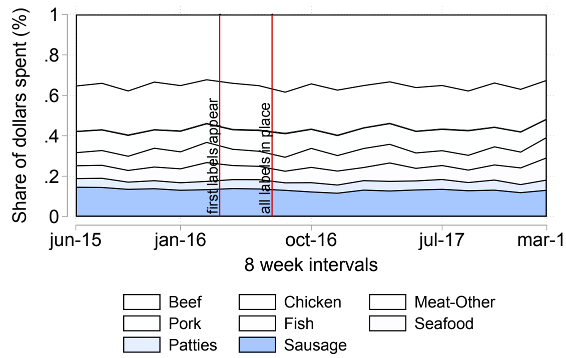
(b) Calorie intake

Figure A.1: Sugar intake per grams and milliliters consumed before and after the policy

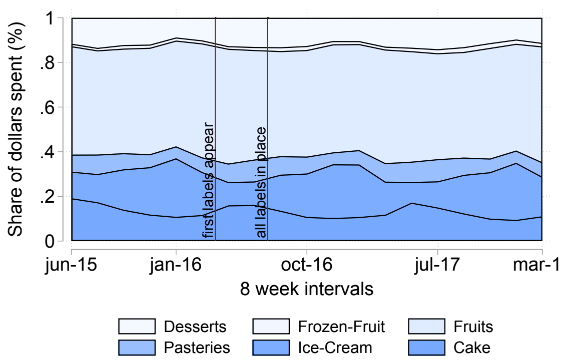
Notes: This figure shows the changes in nutritional intake per volume/mass of food products purchased at Walmart. For volume, we calculate the total amount of kilograms and total liters of products purchased at Walmart. We then divide the total intake of sugar by the total volume/mass of products. Measures of volume and mass of products are subject to measurement error from potential coding errors in package sizes.



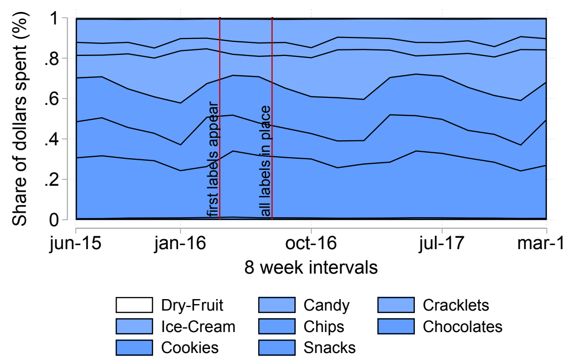
(a) Carbs



(b) Meat



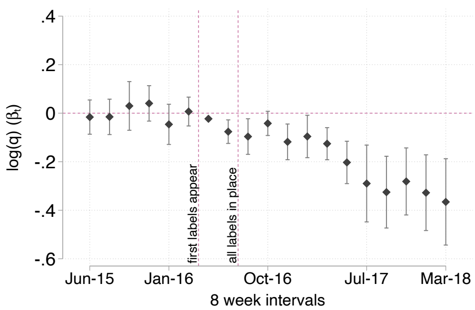
(c) Desserts



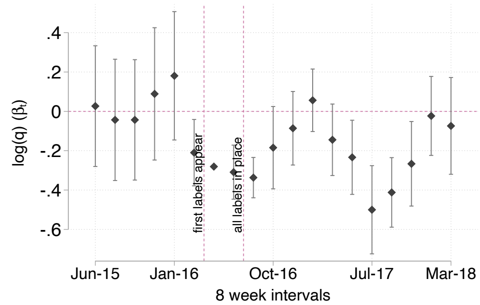
(d) Snacks

Figure A.2: Share of dollars spent across different categories

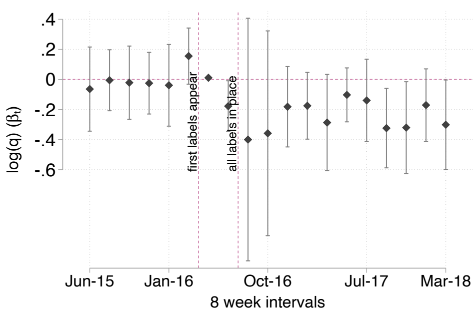
Notes: This figure shows the evolution of the share of dollars spent in each category within broader groups of products. Colors represent the share of labeled products within each category. White areas are categories in which no product received a label, and dark-blue areas (e.g., snacks) are categories in which all products received at least one label. We show that there are no differential changes in dollars spent between low-in-labels and high-in-labels categories.



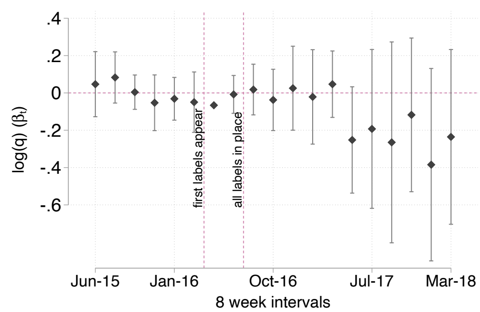
(a) Seasonings



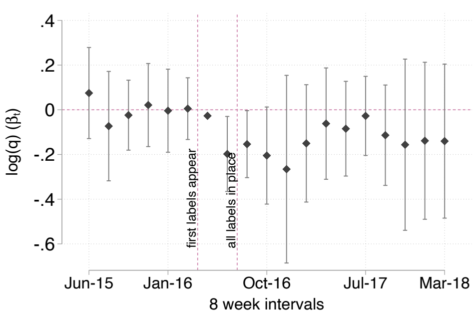
(b) Frozen Fruit and Pulp



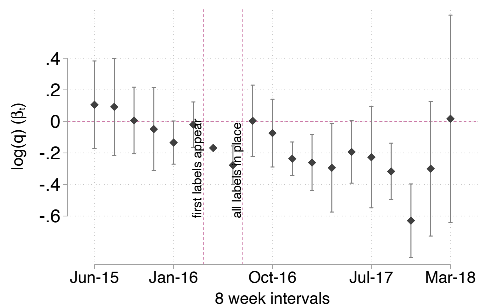
(c) International Cuisine



(d) Syrup and Honey



(e) Instant Noodles



(f) Instant Rice

Figure A.3: Estimates by category

Notes: This figure presents regression coefficient estimates. Each panel presents the β coefficients from Equation (2) for selected categories.

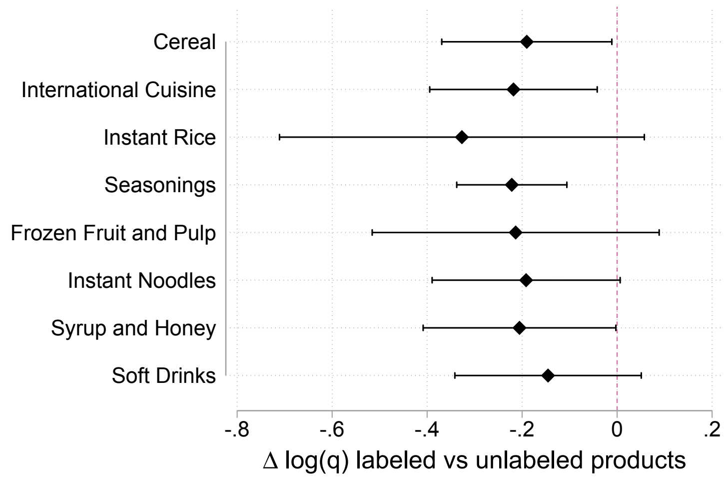
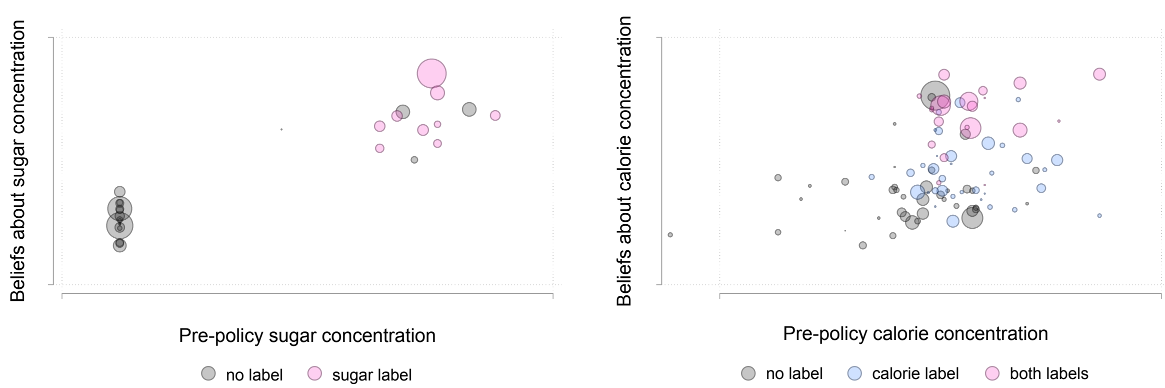


Figure A.4: Changes in demand for selected categories (without controlling for prices)

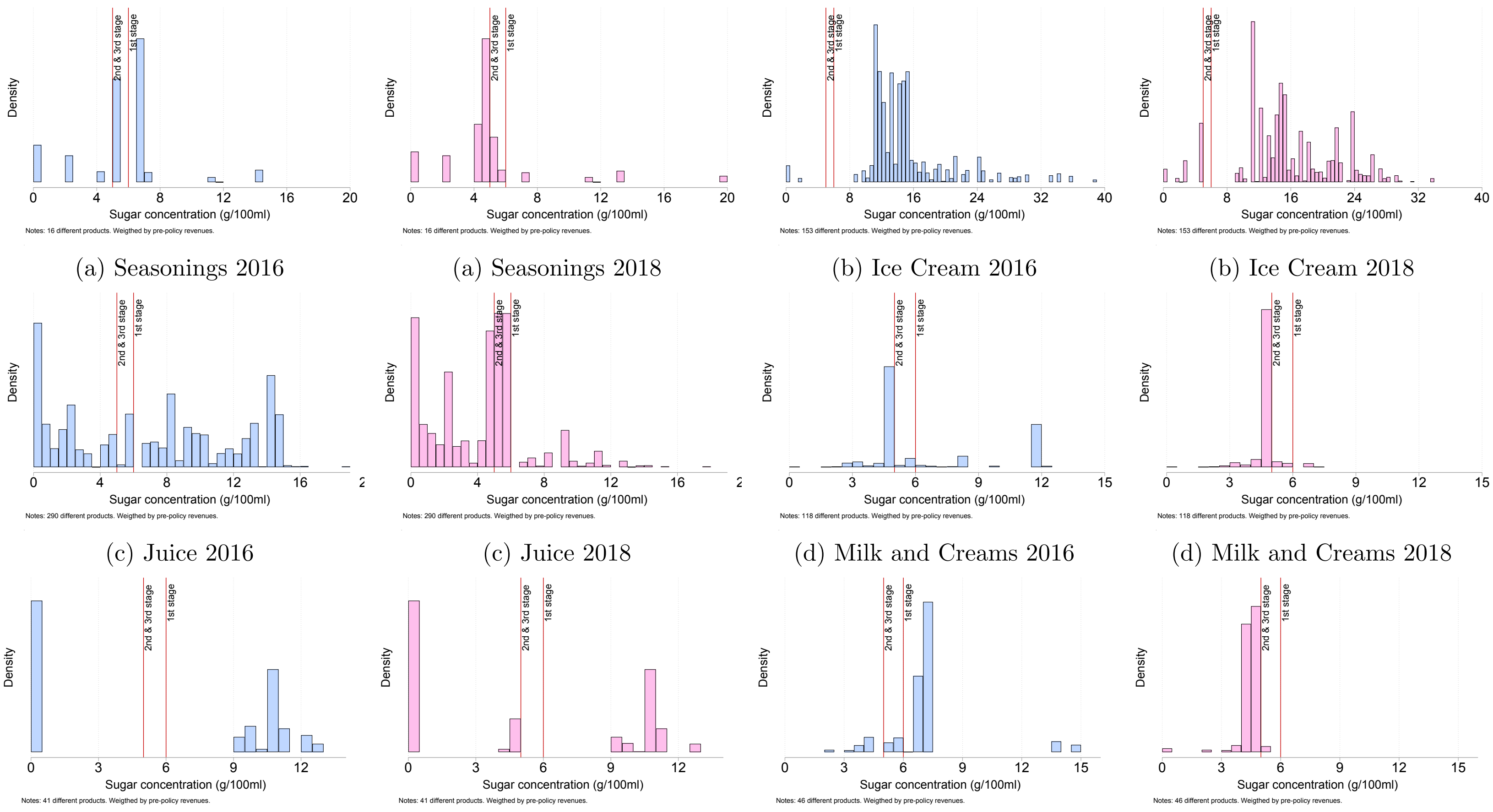
Notes: This shows the changes in equilibrium quantities between labeled and unlabeled products from estimating Equation (2) for different product categories after fixing $\gamma = 0$. The set of products in this analysis represents 11.9% of the pre-policy revenue of labeled products in the sample. We provide more details on sample selection in Appendix B.



(a) Soft Drinks (b) Cereals

Figure A.5: Correlation between beliefs about nutritional content and true nutritional content

Notes: This figure shows the first moments of beliefs about each product's nutritional content vs. its real nutritional content. Each circle corresponds to a different cereal, and its size represents the total revenue from that product in our sample period. Panel (a) focuses on the sugar concentration of soft drinks as measured by g sugar/g product and panel (b) on the caloric concentration of cereals, as measured by kcal/g product. Since we focus on the relative distance between survey responses for different products, we do not provide numerical labels for the x- and y-axes.



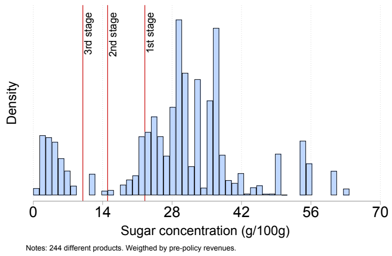
(a) Seasonings 2016 | (a) Seasonings 2018 | (b) Ice Cream 2016 | (b) Ice Cream 2018

(c) Juice 2016 | (c) Juice 2018 | (d) Milk and Creams 2016 | (d) Milk and Creams 2018

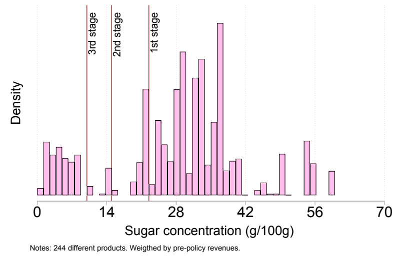
(e) Soft Drinks 2016 | (e) Soft Drinks 2018 | (f) Yogurt 2016 | (f) Yogurt 2018

Figure A.6: Distribution of sugar content pre- and post-policy for liquids in selected categories

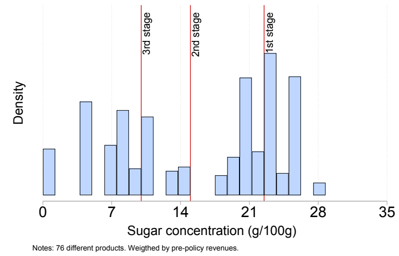
Notes: This figure plots the distribution of sugar concentration for liquid products in different categories before and after the policy implementation. Observations are weighted by pre-policy revenue.



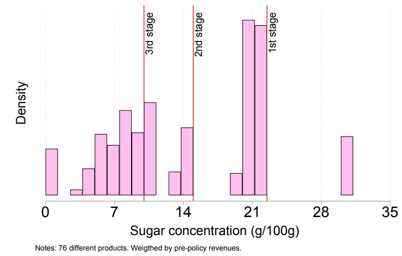
(g) Cookies 2016



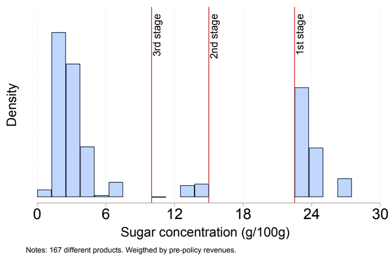
(g) Cookies 2018



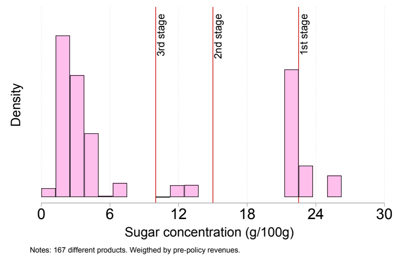
(h) Desserts 2016



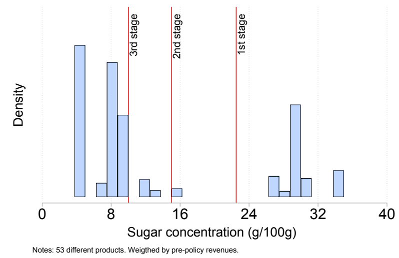
(h) Desserts 2016



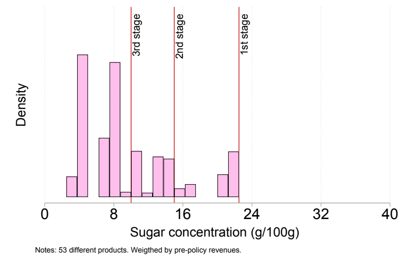
(i) Condiments 2016



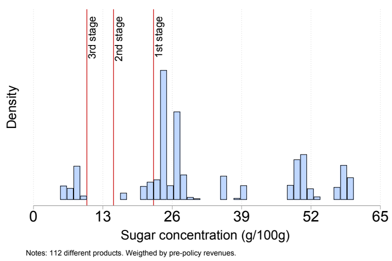
(i) Condiments 2016



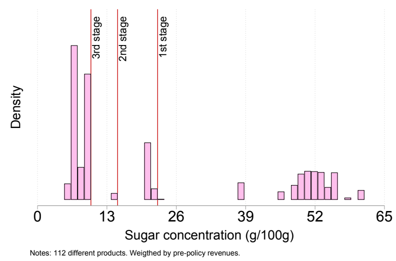
(j) Frozen Fruit and Pulp 2016



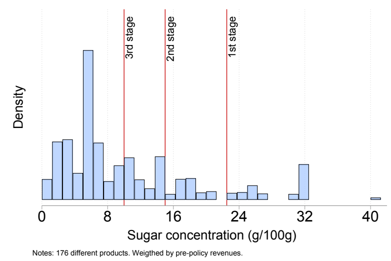
(j) Frozen Fruit and Pulp 2018



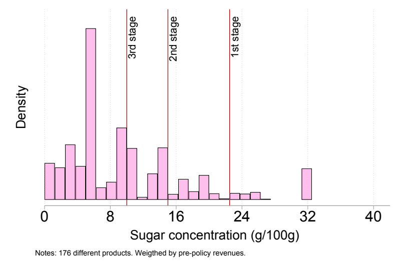
(k) Jam 2016



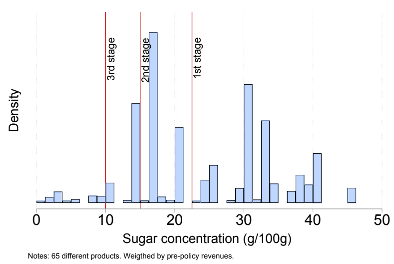
(k) Jam 2018



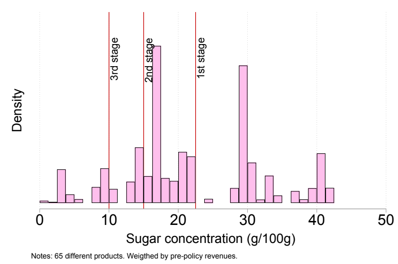
(l) Soup 2016



(l) Soup 2018



(m) Cereal 2016



(m) Cereal 2018

Figure A.7: Distribution of sugar content pre- and post-policy for solids in selected categories

Notes: This figure plots the distribution of sugar concentration for solid products in different categories before and after the policy implementation. Observations are weighted by pre-policy revenue.

Table A.1: FOP warning-label policies

Country	Status	Year	Critical nutrients			
			Sugar	Sodium	Fat	Calories
Chile	Implemented	2016	✓	✓	✓	✓
Peru	Implemented	2019	✓	✓	✓	
Israel	Implemented	2020	✓	✓	✓	
Mexico	Implemented	2020	✓	✓	✓	✓
Uruguay	Implemented	2021	✓	✓	✓	
Argentina	Implemented	2022	✓	✓	✓	✓
Brazil	Implemented	2022	✓	✓	✓	
Colombia	Approved	2021	✓	✓	✓	
Venezuela	Approved	2021	✓	✓	✓	
Canada	Approved	2022	✓	✓	✓	
Antigua and Barbuda	In discussion	-	✓	✓	✓	
Bahamas	In discussion	-	✓	✓	✓	
Barbados	In discussion	-	✓	✓	✓	
Costa Rica	In discussion	-	✓	✓	✓	✓
Dominica	In discussion	-	✓	✓	✓	
El Salvador	In discussion	-	✓	✓	✓	✓
Guatemala	In discussion	-	✓	✓	✓	
Guyana	In discussion	-	✓	✓	✓	
Haiti	In discussion	-	✓	✓	✓	
India	In discussion	-	✓	✓	✓	✓
Jamaica	In discussion	-	✓	✓	✓	
Panama	In discussion	-	✓	✓	✓	✓
Paraguay	In discussion	-	✓	✓	✓	✓
Saint Kitts and Nevis	In discussion	-	✓	✓	✓	
Saint Lucia	In discussion	-	✓	✓	✓	
Suriname	In discussion	-	✓	✓	✓	
Trinidad and Tobago	In discussion	-	✓	✓	✓	

Notes: The table shows the introduction and discussion of mandatory FOP warning-label policies around the world. The table includes countries with laws or government resolutions implemented, approved but not implemented, or under discussion. Countries with discussions of the topic that do not have a specific government plan or law proposal to establish mandatory FOP warning labels are not included. For countries with approved or implemented policies, "Year" indicates the date the policy was approved or the first implementation stage began. Some of these policies distinguish between total fat, saturated fat, and trans fat; we group all of them together for expositional purposes.

Table A.2: Chilean Food Act thresholds

	Solids			Liquids		
	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3
Energy (kcal/100g)	350	300	275	100	80	70
Sodium (mg/100g)	800	500	400	100	100	100
Sugar (g/100g)	22.5	15	10	6	5	5
Saturated fat (g/100g)	6	5	4	3	3	3

Notes: The table shows the level of calories, sodium, sugar, and fat at which a product would need to be labeled. The policy was implemented in three stages, with each stage setting stricter thresholds. Stages 1, 2, and 3 took place in June 2016, June 2018, and June 2019.

Appendix B: Sample Selection in our Within-Category Substitution Analysis

In this appendix we discuss the sample selection process for the categories in our analysis of within-category substitution. We also provide basic descriptive statistics regarding sample coverage.

B.1. *Selection of Categories*

For the purpose of estimating the impact of labels on consumer demand, we first need to define the set of products contained in a given category. The ideal definition of a category for this exercise meets three criteria: (a) products are sufficiently similar, such that consumers would consider substituting from one to another as a result of the regulation; (b) there is sufficient variation in terms of the share of products that receive a label; and (c) for the purpose of estimating a differences-in-differences model, we would need unlabeled and labeled products within a category to follow similar pre-trends in the absence of the policy.

Examining the categories that meet these conditions in our data is not straightforward. First, the product categories in our data are defined by Walmart for administrative and internal processes, and in many cases, they include products that are not necessarily substitutes. Second, about 35% of total revenue comes from products that belong to a category with significant variation in exposure to labels (as defined by having less than 90% of labeled and unlabeled products). Third, most products within Walmart's categories do not follow parallel trends. To address these issues, we selected and combined certain categories, and within those, we restricted our analysis to products that have the potential to function as close substitutes. We also visually inspected and kept those in which labeled and unlabeled products followed similar pre-trends.

B.2. *Sample Coverage*

In Table B.1, we show the share of the revenue covered by the categories included in our demand-side analysis. Column (1) displays the share of total revenue represented by each category provided by Walmart. This column contains the universe of food products purchased by individuals from our consumer panel and adds up to 100%. We group a set of categories for which most products are either below or above the policy threshold in the post-policy period and label them "Mostly unlabeled" and "Mostly labeled." Together, these two groups represent close to 63% of the revenue and include categories such as fruit, meat, salads, candy, and chocolate. Another 30% of total revenue corresponds to other categories in which labeled and unlabeled products did not follow similar pre-trends. Some products in these categories include pastry, bakery products, cold cuts, and biscuits. Our selected categories cover the remaining 5.7% of total revenue.

Table B.1: Selected categories used to study the impact of food labels on consumer demand

	Market share	Share labeled	Market share within labeled products
	(1)	(2)	(3)
Included	5.7	47.7	11.9
Cereal	1.4	62.7	3.6
Frozen Fruit and Pulp	0.1	11.2	0.1
Instant Noodles	0.1	47.3	0.1
International Cuisine: Mexican food	0.1	27.6	0.1
Instant Rice	0.1	27.1	0.1
Seasonings	0.5	57.5	1.2
Soft Drinks	3.3	42.4	6.5
Syrup and Honey	0.1	42.4	0.2
Not Included	94.3	21.2	88.1
Mostly unlabeled	57.4	1.2	1.6
Mostly labeled	6.3	99.2	25.8
Others	30.6	46.9	60.7
Total	100	22.8	100

Notes: Column (1) presents the share of total revenue for each category provided by Walmart. This column contains the universe of food products purchased by individuals from our consumer panel and adds up to 100%. Column (2) presents the share of labeled products within each of the categories. Column (3) presents the share of total revenue for labeled products.

Column (2) reports the share of products, weighted by revenue, that received a warning label within a given category. Column (3) reports the share of total revenue for products that received a label. When focusing on labeled products, our working sample comprises around 12% of the pre-policy revenue of labeled products in the total sample.

Appendix C: Sample Selection in our Product Reformulation Analysis

In this appendix we discuss the sample selection process for the categories in our analysis of product reformulation. We also provide basic descriptive statistics regarding sample coverage.

C.1. *Selection of Categories*

Organizing the categories for this exercise requires different criteria from those used to study demand substitution in Appendix B. We focus on categories in which the distribution of sugar and calories is not entirely to the left of the regulatory threshold. Naturally, unlabeled products do not face any incentives to change their nutritional content. We also dropped categories with products that were too far to the right and for which it was not feasible to modify the nutritional content up to the threshold level.

C.2. *Sample Coverage*

In Table C.1, we show the share of revenue covered by the categories included in the supply-side analysis. Column (1) reports the share of total revenue for each category provided by Walmart. This column contains the universe of food products purchased by individuals from our consumer panel and adds up to 100%. We group a set of categories for which most products lie either below or above the policy threshold in the pre-policy period and label them "Mostly below" and "Mostly above." Together, these two groups represent close to 67% of revenue and include categories such as fruit, meat, salads, candy, and chocolate. Another 17.5% of total revenue is for categories with products that are exempted from the regulation (e.g., nuts) or products for which we are missing the pre-policy nutritional content. Our selected categories cover the remaining 15.5% of total revenue.

Column (2) reports the share of products, weighted by revenue, that are above the sugar threshold in the pre-policy period within a given category. Column (3) reports the share of total revenue for all products that are above the sugar threshold in the pre-policy period. When focusing on products with the potential to bunch, our working sample comprises around 54% of the pre-policy revenue for them.

Table C.1: Selected categories used to study the impact of food labels on product reformulation

	Market share (1)	Share above the sugar threshold before the policy (2)	Market share within products above the sugar threshold (3)
Included	15.5	63.2	53.6
Cereal (g)	1.4	43.5	3.1
Cookies (g)	2.1	77.3	10.7
Desserts (g)	0.6	31.2	1.3
Condiments (g)	0.4	30.9	0.9
Seasonings (ml)	0.1	49.0	0.1
Frozen Fruit and Pulp (g)	0.1	28.4	0.2
Ice Cream (ml)	0.9	98.2	6.0
Jam (g)	0.4	83.2	2.2
Juice (ml)	2.6	54.0	9.4
Milk and Creams (ml)	2.5	92.0	5.4
Soft drinks (ml)	3.3	53.7	11.5
Soup (g)	0.6	11.4	0.2
Yogurt (ml)	0.5	83.0	2.6
Not Included	84.5	10.0	46.4
Mostly below	63.8	0.3	1.0.1
Mostly above	3.2	98.0	17.2
Others	17.5	29.3	28.1
Total	100	18.28	100

Notes: Column (1) presents the share of total revenue for each category provided by Walmart. This column contains the universe of food products purchased by individuals from our consumer panel and adds up to 100%. Column (2) presents the share of labeled products within each of the categories. Column (3) presents the share of total revenue for labeled products.