Identifying Food Labeling Effects on Consumer Behavior

Sebastián Araya, Andrés Elberg, Carlos Noton and Daniel Schwartz*

December 13, 2021

We examine a large-scale mandatory food labeling regulation to identify its effects on consumer behavior. We take advantage of exogenous variation in product-labeling status from the gradual and asynchronous introduction of labeled products on store shelves many weeks before the legal deadline. We combine individual-level scan data from a large retailer with on-the-shelf information on the actual warning-label status for breakfast cereals, chocolates, and cookies. Warning labels decrease demand and purchase probabilities in the cereal category, and this effect is larger on medium-low socioeconomic groups. We find inconclusive results of the warning label on chocolates and cookies. Overall, results suggest that the warning label effect is consistent with information disclosure influencing consumers' choices when the advertised information is unexpected.

Keywords: Food Labeling, Consumer Behavior, Nutritional Information, Point-of-Sale Advertising

JEL Codes: D12, I12, I18

*Araya: Department of Industrial Engineering, University of Chile, Beauchef 851, Santiago, Chile, searaya@ing.uchile.cl. Elberg: Pontifical Catholic University of Chile, School of Management, Vicuña Mackenna 4860, Macul, Santiago, Chile, aelberg@uc.cl. Noton: Department of Industrial Engineering University of Chile, Beauchef 851, Santiago, Chile, cnoton@dii.uchile.cl. Schwartz: Department of Industrial Engineering, University of Chile, Beauchef 851, Santiago, Chile, daschwar@dii.uchile.cl. We thank comments from Bryan Bollinger, Rosario Macera, Marcelo Olivares and Arjen van Lin. Schwartz acknowledges financial support by the Complex Engineering Systems Institute (ANID PIA/APOYO AFB180003) and ANID Fondecyt 1191745. Noton acknowledges financial support from Conicyt (Fondecyt No. 1170029) and the Institute for Research in Market Imperfections and Public Policy, ICM IS130002. Elberg acknowledges financial support from ANID Fondecyt 1180301. This research was partially supported by the computing infrastructure of the NLHPC (ECM-02).

1 Introduction

Obesity has rapidly become a first-order public health concern for governments around the world.[1,2] One approach that has gained prominence in recent years to help curb the global obesity epidemic is the use of nutrition labeling. As the scientific evidence linking obesity -and associated chronic diseases- to dietary habits has mounted, the World Health Organization (WHO) has forcefully advocated the use of nutrition labeling schemes and the provision of nutritional information as leading strategies to improve healthy food choices (WHO, 2004). In line with the WHO's recommendations, many countries have required food providers, such as supermarkets, to disclose calorie and nutritional content information (Hawkes, 2004, WHO, 2004). However, a potential drawback of this approach is that consumers may misunderstand or misuse nutritional information impeding effective communication (Cowburn and Stockley, 2005). To overcome this problem, some retailers have voluntarily provided simplified information on healthful products to persuade customers interested in improving their food choices at the point-of-sale.[3] More broadly, several countries are moving towards mandatory simple front-of-package labeling, focusing on unhealthful products, to change shoppers' behavior.

Despite its importance, identifying the effect of food labeling on consumer behavior has proved elusive to date. Bleich et al. (2017), for instance, review 53 studies on the impact of calorie labeling on consumer behavior and conclude that the lack of statistical power and strong designs challenge clear conclusions about the effects of calorie labels.[4] Moreover, the fact that most regulations are implemented at a single point in time poses additional challenges to pin down the effect of front-of-package labeling. For instance, before-after estimations need the unobserved components of consumer behavior to be time-invariant (Ippolito and Mathios, 1995, Dumanovsky et al., 2011; Elshiewy and Boztug, 2018, Nikolova and Inman, 2015; Taillie et al., 2020b; Correa et al., 2020). Also, comparing regulated markets with other geographic locations (as control) requires unobservables to have a parallel trend across markets (Elbel et al., 2009, Bollinger et al., 2011, Finkelstein et al., 2011, Hobin et al., 2017).

We take advantage of the gradual implementation of a comprehensive and mandatory food labeling regulation recently introduced in Chile to identify its effects on consumer behavior. The regulation was prominently featured in the international press (Jacobs (2018); Boseley (2020)) and described as "the world's most ambitious attempt to remake a country's food culture", and "could

[1] According to the Global Burden of Disease Study IHME (2013), the number of obese and overweight individuals rose by 28 percent in adults and 47 percent in children worldwide over the last 30 years.

[2] Cawley (2015) and Malnick and Knobler (2006) document the adverse consequences of obesity and other nutrition-related diseases on health and economic outcomes.

[3] The "Guiding Stars" system has been implemented by a few grocery stores in North America.

[4] See Kiszko et al. (2014), Harnack and French (2008), Taillie et al. (2020a), and Ikonen et al. (2020) for additional systematic reviews with similar findings; all reviews conclude that more research is needed to determine whether nutrient warning labels work with actual purchases.

be a model for how to turn the tide on a global obesity epidemic" by the New York Times. The new regulation established that products exceeding certain thresholds of critical nutrients should display mandatory warning labels by the end of June 2016. However, food suppliers gradually introduced the warning labels in different retail stores a few months before the regulation came into force. During this period, supermarkets began selling labeled products driven by stock availability in each store, and suppliers were delivering labeled and unlabeled products in a seemingly unplanned manner.

We collected daily data on the label status of specific products (at the Universal Product Code (UPC)-level), and observe substantial variation in labeled and non-labeled products across time and supermarket stores, allowing us to overcome identification problems present in some of the previous literature. We combine the label information with individual-level transaction data from a big-box retailer. We focus on three categories in which a large proportion of products was *expected* to be labeled and affected by the regulation: breakfast cereals, chocolates, and cookies. Our scan data come from all transactions in the loyalty card records of the retailer.

We conduct several analyses. First, we leverage the staggered rollout in a diff-in-diff approach using product-level panel data, in which the warning label varies by UPC, store, and day. We estimate the warning level effect on purchased quantity and category share. Second, we use household-level panel data to examine heterogeneous treatment effects based on two dimensions: Socioeconomic status (SES) and the presence of small children in the household. We also provide evidence on the exogeneity of the labeling rollout as a robustness check, and we perform a placebo test to support our identification strategy. We also examine changes in prices to rule out alternative explanations. Finally, we use a hierarchical discrete choice model to assess the policy's effectiveness in substituting purchases of unhealthy products. We compare demand patterns under the labeling policy with those predicted by our model in a counterfactual scenario of labels removal. We further estimate the ad-valorem tax rate that would have equivalent effects on demand as the labeling policy.

Across analyses, we find that shopper responses to the warning labels vary widely between product categories. In the breakfast cereal category, the warning labels reduce the purchased volume in 6.2 percentage points, and the purchase probability by 5.8 percent. In chocolates and cookies, we find inconclusive evidence to determine substantial effects of the regulation. This suggests that the warning labeling effect may be consistent with information disclosure being more effective when it adds new insights to the consumer initial information set (Loewenstein et al. (2014)).[5] Furthermore, in the breakfast cereals category, our estimates from the household analysis indicate that lower-income consumers and families with children are sensitive to warning labels. These two groups are generally highlighted as a priority to tackle the obesity pandemic worldwide (CDC, 2019; Loring and Robertson, 2014). Our demand model finds that, in breakfast

[5] Along the same lines, Hobin et al. (2017) find that people prefer more nutritious cereals after a nutritional guiding-stars system was implemented, but this does not hold for snacks.

cereals, that warning label effect is explained by a large share of substitution from unhealthy to healthy products and, to a lower degree, to a reduction of purchase in the food category.

Our paper contributes to a growing literature studying the role of interpretive nutritional information on consumer behavior and point-of-sale advertising on several dimensions.[6] First, we contribute to the statistical identification of nutritional warning label effects. As mentioned above, most research on nutritional labels has used a diff-in-diff strategy (before-after approach and treatment-control geographic locations) exploiting the label implementation in a specific time or territory. In our setting, the staggered rollout of warning labels affected products and stores differently. The mix of labeled and not labeled UPCs in the stores was different over time. Furthermore, the mandatory nature of the regulation ensured that customers could not systematically avoid labels by switching stores. Attempting to address the identification concerns, other authors have implemented experiments to assess the effects of food labeling. Kiesel and Villas-Boas (2013) conduct a study in which they manipulate the information content of nutritional shelf labels in one product category (microwave popcorn) across five treatment stores selected by the supermarket chain and one "synthetic control group".[7] While the experimental setting allows the authors to compare labeled versus unlabeled microwave popcorn, the selection on unobservables to create a control store remains a potential issue. In our study, stores were not participating in a voluntary experiment but under a mandatory regulation, which allows us to provide evidence on consumers' in-store purchase behavior relevant for countries planning or introducing similar compulsory nutritional labels. We also add to this research by considering other relevant food categories from managerial and health policy perspectives. Downs et al. (2009) conducted two experiments showing people a menu with calorie information. They find a small effect of calorie information, which may even backfire for dieters. One concern is that consumers may be aware of their participation in a study, potentially driving their attention to the new nutritional information. Previous research has shown that awareness of participation may change people's behavior (Schwartz et al., 2013). Our field setting, combined with the use of vast amounts of transactional data, avoids potential biases from surveys and laboratory experiments as it captures the normal shopping behavior of consumers during the introduction of an exogenous change in information.

Also, thanks to the staggered rollout of warning labels occurring months ahead of the date the law entered into force, we are able to separately identify the impact of warning labels from the effect of other policies included in the law –which entered into force concurrently with the warning label system– such as the ban on advertising directed to children and the prohibition to display cartoon characters in the package of products exceeding the regulatory thresholds. Thus, our paper can inform marketers, industry participants and policymakers about the specific effects of nutritional warning labels on purchase behavior.

[6] See Griffith and Nevo (2019) for a survey of the quantitative marketing literature on nutrition labeling.

[7] Kiesel and Villas-Boas (2013) randomly assigned different "low" tag labels (a combination of low calorie, low fat, and low trans-fat) to each store. The authors use a synthetic control group to address the store selection issue, as the supermarket chain did not provide information on how they selected the five treated stores.

Second, the mandatory nature of the policy we study allows us to overcome some challenges facing previous work on the effects of interpretive warning labels. Compared to other nutritional label schemes that require people to interpret and understand the information (e.g., calorie posting), warning labels advertise distinct information levels and provide an interpretation about which food is unhealthful.[8] Similar to our work, Bollinger et al. (2019) estimated the effect of an interpretive nutritional label, the Guiding Stars system, on purchase behavior in supermarkets' point-of-sale. However, the authors found that consumers did not understand nor were sufficiently aware of the Guiding Stars, a voluntary system, making it difficult to examine the label's effect. The mandatory law affected all products and stores and had a simple and publicized design that ensures widespread awareness, understanding, and consideration from consumers (CERET, 2016; Valdebenito et al., 2016).[9] In effect, the Chilean warning label system has become a prominent standard implemented or discussed in several countries (e.g., Israel, Canada, Mexico, Brazil, South Africa, and Peru) – our findings help stakeholders evaluate a crucial piece of the legislation that has attracted most of the attention worldwide.

Finally, our paper contributes to the debate on mandatory information disclosure and its potentially heterogeneous effect across different segments of the population (Cawley et al., 2016). Finding effective measures to close the gap in nutritional inequality has proved to be difficult.[10] We find that the labeling policy may help modify the purchasing behavior of lower socioeconomic consumers –who tend to suffer more from obesity– in specific food categories.

The remainder of this paper is organized as follows. Section 2 describes the nutritional warning information we study, institutional details, and summary statistics of our supermarket data. Section 3 presents our analysis based on the panel data of products and households. Section 4 presents our analysis using a structural demand model. Section 5 discusses the overall results and concludes.

[8]The marketing literature distinguishes between reductive (e.g., calorie posting) and interpretive (e.g., warning labels) information (Ikonen et al., 2020), with the former considered as more complex and time consuming for consumers. Examining the US regulation on calorie posting for restaurant chains, Bollinger et al. (2011) find that mandatory calorie posting causes average calories to decrease by 6 percent in Starbucks, comparing New York City, where a calorie posting law was first implemented, with other cities. Our paper extends this research by examining advertising information different in nature in addition to the identification procedure.

[9]For example, a survey with more than 3,000 customers at the exit of supermarket stores indicated that 73 percent of consumers identified products with the new food labeling before the law came into effect (CERET, 2016).

[10]For instance, Allcott et al. (2019) found that product availability does not explain the substantial purchasing differences of healthful products across income groups in the US.

2 Data and Institutional Background

2.1 The Chilean Law of Food Labeling and Advertising and its gradual implementation

In the last few years, Chile introduced groundbreaking changes to its legislation regulating nutritional food labeling. The new regulatory framework was broadly aimed at improving point-of-sale nutritional information using simple interpretive front-of-package labeling.[11] Under the new regulations, pre-packaged food products whose contents of four critical nutrients –sugar, sodium, saturated fats, and calories[12] exceed certain thresholds must display standardized black labels warning that the product contains excessive levels of the corresponding critical nutrients.[13]

According to the regulation, the warning labels must be prominently displayed on the front pack of the product. They take the form of octagons, resembling a black stop sign, displaying the legend *High in* followed by the name of the critical nutrient being exceeded.[14] Figure 1 presents the labels introduced by the law. The regulation is precise about the size of the warning labels and the position they must occupy to ensure saliency to the public. For instance, according to the law, a product which exceeds a critical nutrient limit and whose front pack exceeds 300 square centimeters (approximately 0.32 square feet) must include a warning label of dimensions 3.5 by 3.5 centimeters (about 1.38 by 1.38 inches). The law divided products into solids and liquids and specified the thresholds for labeling a product in terms of a fixed quantity of the product (100 grams for solids and 100 ml. for liquids).[15]

The Chilean Law of Food Labeling and Advertising established a three-stage process over which products would be progressively labeled as "High in" a critical nutrient. The initial phase began on June 27 of 2016, one year after the official order specifying the details of the new regulation was published. More stringent thresholds were mandated to be gradually introduced in June 2018 and June 2019. In this paper, we focus on the impact of the nutritional labels introduced during the first phase of the process.[16]

An international comparison puts Chile among the early adopters of a mandatory front-of-

[11]The law only targets processed packaged products and not bulk goods or unpackaged food such as bread.

[12]While calories are not, strictly speaking, a nutrient, we refer hereafter to all four food components regulated by the law (i.e., sugar, sodium, saturated fats and, calories) as nutrients for expositional convenience.

[13]Also, the new legislation regulated the advertising of the labeled products and their sales in schools. In particular, advertising of unhealthy tagged products targeting children under age 14 years was prohibited as was the sale of these products in or within 100 meters of a school.

[14]All the black stop signs included the name of the Ministry of Health. Mentioning the institution backing the nutritional message has been found to enhance the warning (Feunekes et al., 2008).

[15]For further details on the design and threshold considered see Corvalán et al. (2013); Reyes et al. (2019).

[16]The thresholds for solid (liquid) products over the initial phase were defined as 350 (100) for calories; 800 (100) for sodium; 22.5 (6) for sugars; and 6 (3) for saturated fats.

pack nutrition labeling law, an ambitious policy intervention that is being increasingly considered by other countries worldwide (Hawkes, 2010, NYT, 2018; TheGuardian, 2020). For example, Canada has begun discussing the adoption of a mandatory front-of-pack nutrition labeling system which, according to the initial specifications set by the Canadian Ministry of Health, would include several elements contained in the Chilean law.[17] Among the countries that have already implemented mandatory front-of-pack nutrition labeling systems are Bolivia, Ecuador, Peru, Uruguay, Israel and Mexico.[18]

The actual implementation of the new regulation plays a key role in our empirical strategy. The law mandating the introduction of the warning labels was approved in June 2012, but its implementation required the completion of several administrative and legal procedures. The legislation was finally enacted in April 2015 and entered into force on June 27th of 2016.

There was an initial period of confusion about whether the stock of unlabeled products exceeding the limits of critical nutrients would be allowed to be exhibited on store shelves after the June 2016 deadline. The authorities ruled that all products "High in" some nutrients would have to display the warning labels by June 27th of 2016 regardless of their manufacturing date. Stores that failed to comply with the new regulations by the deadline would be subject to fines. This clarification prompted large retailers to demand delivery of labeled products several months in advance of the legal deadline. This process resulted in some products displaying the black warning label(s) in some stores but not in others.

Our empirical strategy exploits this gradual implementation of the warning labels. Since retail stores received labeled products before the deadline set by the law, we can observe at a given point in time a product displaying a warning label in one store and the same product in a similar outlet without the warning label.[19] This overlap of labeled and unlabeled products changing over time, coupled with observations of purchasing behavior at the UPC-store level, allows us to measure the impact of the food warning labeling on consumer behavior.

One potential concern is that the assignment of labeled products to retail outlets was manipulated by manufacturers based on sales or prices, being considered endogenous to consumers. From several interviews we conducted with large suppliers of products directly affected by the regulations, we learned that it was logistically impractical for them to determine which specific stores would end up receiving the labeled products. Nevertheless, we conduct an econometric analysis to test this potential concern in Subsection 3.2. We introduce our data and discuss the

[17] In a recent stakeholder engagement meeting organized by Health Canada, the authority required stakeholders to submit possible front-of-package nutrition symbols, which complied with three criteria included in the Chilean law. The three principles are: (1) follow the "high-in" approach; (2) focus on the three nutrients of public health concern (sugars, sodium, and saturated fats); and use only black and white colors (HC, 2017).

[18] In other nations, graphical nutrition labeling schemes are applied voluntarily. A pioneering intervention along these lines is the traffic light system implemented in the UK. The system was born as an initiative of the industry and has replicated by some retailers in France and Portugal (Hawkes (2010)).

[19] We identify a product based on its Universal Product Code, UPC.

observed implementation in the next subsection below.

2.2 Data Description

We partnered with a large chain of supermarkets in Chile to study the impact of the nutrition labeling law on purchasing behavior.[20] We were able to measure whether specific UPCs displayed warning labels on the shelves of six supermarket stores located in the two most populated regions of Chile over a period of gradual and asynchronous introduction of warning labels in supermarket stores. Our team of research assistants visited the stores before the legal deadline, during May-July 2016, took pictures, and recorded the label status of each UPC and the specific type of warning label it displayed. On average, each store was visited 40 times over the period in which the warning-label status of a given UPC exhibited variation across stores.[21]

We combine our collected data on the presence of warning labels with consumer-level point-of-sale data, which include all items in consumers' shopping baskets, the prices paid for each item, and the date and time of the transaction. Our scan-data identify individual consumers using customer membership in the retailer loyalty program and the *unregistered* consumers. According to the retailer, purchases made through its loyalty program account for approximately 80 percent of its total revenues. Our consumer-level data also contains gender, age, and socioeconomic status (SES). The retailer classifies a customer into one of five SES categories (ABC1, C2, C3, D, and E) based on the specific street block where the customer resides using Census data. Also, our dataset includes historical data extending back to early 2015 with purchases made by the same set of customers. We use data from May to July 2015 for placebo tests.

We focus on three product categories, as a large fraction of their products was targeted by the regulation: breakfast cereals, chocolates and cookies.[22] As an illustration of the variation we observe, Figure 2 shows the evolution of warning labels per category in the six stores included in our sample for the top-selling 20 products in each category. As expected, there is an upward trend in the number of labeled UPCs over time across all stores and categories. There is considerable variation in the food labeling implementation across products, stores, and time.[23]

[20] Our data comes from one of the top-three supermarket chains. The three largest chains in Chile account for more than 90% of the market.

[21] Our data also include transactions between June 27 and July 22, 2016, when the law had already entered into effect.

[22] We also collected data on juices and started collecting data on yogurt. However, only one product for juices and none for yogurts were labeled under the current law. Soft drinks are another relevant category, but these products were all labeled several weeks before we began collecting data on whether products displayed the new label (i.e., several months before the law came into force).

[23] Only as an example, using the top UPC (in terms of sales) that exceeded the regulatory thresholds in the breakfast cereals category we observe that in store #1, this UPC displayed warning labels on May 4th, and then continued that way. In store #2, this UPC displayed warning labels from May 10th, but only for a few days. It started displaying warning labels again on June 3rd. In store #3, it displayed warning labels from May 17th. In store #4, it displayed warning labels for the first time on May 10th, then on the third and fourth week of May, it displayed warning labels

Additionally, Figure 3 shows the number of days in advance of the legal deadline when the warning labels were implemented for each for the same top 20 products in each category. The figure displays average and standard deviation across stores of the number of days ahead of the deadline when the warning label was introduced.[24] Importantly, the charts in Figure 3 show no clear pattern linking market shares with the timing of the introduction of warning labels across stores. In Subsection 3.2, we perform a formal econometric test to verify whether the time of label implementation or the prices can be explained by pre-regulation sales, prices or quantities.

Table 1 shows summary statistics for our data. Panel A presents details of product-level panel such as the number of products, average price, and the average of products sold per day-store. Panel B describes the household-level panel reporting the number of households, their average (and median) number of trips as well as their purchased quantity. Panel C shows the number of labeled UPCs per category and their respective fraction of products and market shares in July 2016 (i.e., when the law came into force). For the last panel, we observe that almost all cookies and chocolates ended up with 3 labels (stop signs with "high in") .

The product-level panel used in Section 3 aggregates transactions, for each store-day, from registered and unregistered households and considers all the UPC-store-day observations with an average of more than six transactions per day per store to ensure that the product was available to consumers as stockouts are unobservable, and research assistants could only get information from products on the shelves. Using this threshold, the set of products represents a very large portion of the universe of products in the entire period in all stores (between 71 and 76 percent). In each observation (UPC-store-day), we observe whether the product was labeled. We also use category-specific household panel data to examine heterogeneous effects. In this case, we built the entire set of available UPCs for each consumer in a store visit. To control for household fixed effects, we use the category-specific sample of registered consumers with at least six trips in 2015 and another six in 2016. Finally, Section 4 estimates a discrete choice demand model using the same set of households and the top ten products in each category.

3 Product Level Analysis

We first leverage the gradual rollout using a staggered diff-in-diff strategy at the product level.[25] We estimate the warning label effect using a linear regression model, including the available UPCs

only half of the days. Moreover, this top UPC displayed no warning labels from June 3rd to June 16th, and then it kept displaying warning labels. In stores #5 and #6, there was not much variation. The UPC displayed warning labels almost every day.

[24]Within a category, the figure shows the products ordered by their market shares, with product 1 being the largest market share product and the last product exhibiting the smallest market share among the selected products.

[25]In contrast to a standard diff-in-diff approach, in our setting a UPC can change its status as "treatment" or "control" across stores and over time.

in each store per day.[26] Each observation is at UPC-store-day level, indicating whether the product was labeled, the number of stop-sign labels, its presence on transactions, volume sold, and the mean price. We use the following specification:

$$Y_{jst} = \beta P_{jst} + \gamma L_{jst} + \delta_{js} + \xi_t + \lambda_t + \varepsilon_{jst} \tag{1}$$

where the dependent variables for UPC j at store s on day t, Y_{jst}, can be (1) the percentage of transactions of the product j within its category, or (2) the log quantity sold of product j in that day-store combination.[27] L_{jst} indicates whether product j was labeled on day t at store s. This equation also include other explanatory variables such as UPC mean price (P_{jst}) at the store-day level, product-store fixed effects (δ_{js}) to control for baseline differences across stores (e.g., layout, store-specific promotions, and space on the shelf), day-of-the-week dummy variables (ξ_t) and week fixed effects (λ_t) to control for purchase patterns on different days of the week and week-specific shocks, respectively. We compute two-way clustered standard errors by UPC and by store (Bertrand et al., 2004). We use this specification for each food category.[28]

3.1 Main Labeling Effect

Table 2 presents the estimates of the average treatment effect of the warning label in the breakfast cereals category. We present different specifications including different sets of covariates. We begin by noting that the estimated price coefficient is consistently negative in both panels. More importantly, we find that the warning label effect is consistently negative across all specifications. Estimates in Column (3) of Panel A, using the full set of covariates from Equation (1), indicate that the volume of breakfast cereals is reduced by 6.2 percentage points (CI95% [-10.9 pp., -1.5 pp.]) when a product is labeled. To get a sense of this effect, we observe that, on average, each cereal sold 20 units per store-day during the time of the study. The warning level effect reduced this number by between 1 and 2 units in each store per day. Also, based on Panel B, when a product is labeled, it is less likely to be present in a ticket that includes any cereal by 0.11 percentage points (CI95% [-0.15 pp.,-0.001 pp]). Because in breakfast cereals, a particular UPC will be in a transaction in 1.9% of the tickets that contain any cereal, this number goes down to 1.79% when the UPC is labeled (i.e., a 5.8% reduction).

Table 3 and Table 4 present the analysis for chocolates and cookies, respectively. In both cases, for the change in volume, the magnitude of the coefficient associated with the warning label is

[26] Since products had to be available at the store to be inspected on whether they were labeled, we considered products present in more than six transactions (regardless of the purchased quantity) in each store-day combination. Results were very similar when the number of transactions needed was lower or higher.

[27] We use $\log(Q_{jst} + 1)$.

[28] In Appendix A, we also use block bootstrapping and results are very similar, as well as more saturated models including product×time fixed effects, although this latter adds many singletons considering that our database aggregates all the transactions at the product-store level.

positive but not sizably different from zero (CI95% [-2.5 pp., 5.7 pp.] and CI95% [-2.3 pp., 6.0 pp.], respectively). Therefore we cannot reject the hypothesis that γ = 0, that the warning label changes demand in these cases. Similarly, for the percentage of transactions, the 95% confidence intervals overlap with zero leading to inconclusive evidence of the warning label effect in these categories, [-0.11 pp., 0.13 pp.] for chocolates and [-0.03 pp., 0.10 pp.] for cookies. Since a particular chocolate UPC is sold 28 times, on average, and a cookie 31 times, the point estimates would indicate a variation in half of a unit in both categories.[29] For the change in percentage of transactions, considering that a particular chocolate product is in 1.1% of the tickets with chocolates, and a cookie in 1.4% of the ones with cookies, the point estimates represent a variation of 1% and 2%. These results suggest that the effect of the nutritional label has a moderate magnitude for cereals and is small and inconclusive (given the imprecision of the estimates) for cookies and chocolates. In this regard, Appendix B compares the estimates across the three categories, finding that the estimates for quantity sold of cereals are significantly smaller than those for chocolates and cookies, with no sizable difference between the last two (for the percentage of transactions, the results were similar, except between cereals and chocolates in which the difference was not sizably different from zero at the 5% significance level).

We also examine the effect of the number of warning labels estimating Equation (1) but using three dummy variables for one, two, and three warning labels instead of a single indicator L_{jst}. In Table 5, we find that in breakfast cereals, the magnitude of the coefficient associated with one, two, or three warning labels are very similar, but showing wider standard errors compared to the previous analysis, especially for the three-warning label due to fewer products with this number of warning labels in this food category (12% of all cereals). In the chocolate category, we find positive coefficients, and none are sizably different from zero at a 10% significance level, and we need to consider that only 9% of the chocolate products ended up with a one-warning label. For cookies, there are negative and positive estimates, this latter for three-warning labels, but again none of them are sizably different from zero at a 10% significance level (most of the cookies ended up with three-warning labels). The estimates in these products are consistent with the intuition that "unhealthy" warning labels may be interpreted as "tasty indicators" (Ikonen et al. (2020), Raghunathan et al. (2006)). Similarly, Kiesel and Villas-Boas (2013) found a negative effect of a "low-fat" nutritional label on microwave popcorn. They argue that a trade-off between taste and nutritional content might be an explanation for their finding.

[29] Even a negative effect small as 2.5 percentage points would indicate a less than one unit variation in these categories.

3.2 Pricing

To examine whether pricing behavior was affected by the gradual implementation of the warning labels we estimate the following regression:

$$P_{jst} = \gamma L_{jst} + \delta_{js} + \xi_t + \lambda_t + \varepsilon_{jst} \tag{2}$$

Using the same notation of Equation (1), we seek to test whether prices were endogenous to the label treatment.

Table 6 shows that for cereals and cookies is unlikely that prices were affected by the warning label, or its effect is tiny, during the implementation period (95%CI [-0.03, 0.03] and 95%CI [-0.01,0.01], respectively). As a reference, CLP 5 (from the coefficient γ = 0.005, in thousands of CLP) were less than 1 cent in US Dollars at the time of the study.[30] In the case of chocolates, there was a difference of CLP 31 (95%CI [-0.05,-0.01]), but this effect was minimal (equivalent to 4.6 cents in US Dollars), as the average mean price of chocolates was CLP 1,405 (SD = 912). These results indicate that price promotions were not associated with the early implementation of the warning labels.

3.3 Labeling Timing and Placebo Tests

Identification of the food labeling effect relies on the staggered implementation of the labels conducted by suppliers at different times and stores. We first test whether the warning label implementation was associated with products' sales. For example, it is plausible that products which were sold faster were replaced by labeled products at the beginning of the implementation period. We examine this issue by regressing the number of days in advance of the official deadline that the product was labeled in a given store for the first time on either sales, volume, or price using previous aggregated sales from May-July 2015.[31] This analysis is conservative because it assumes that when a UPC is labeled in a given store, it remains labeled. However, we often observe a product's labeling status being reversed from labeled to unlabeled in the same store.

The estimates in Table 7 show that in none of the categories considered we find a meaningful correlation between the timing of a product's first labeling and its previous sales, quantity sold or price. All estimates are quite small and not sizably different from zero at the 10% significance level (almost all p-values are greater than 0.5), and only in the case of cookies, there is a negative coefficient indicating that more expensive cookies were labeled earlier. However, the corresponding effect size is very small: Cookies that cost CLP 1,000 more were labeled 5-6 days before on average (95%CI [-8.3, -2.4]), and the average price of cookies was CLP 846.[32]

[30] The average exchange rate in May-Jul 2016 was USD 1 = CLP 673.5.

[31] For products that did not exist in 2015, sales from May 2016 were used.

[32] Table A.3 in the Appendix section shows this analysis using UPC fixed effect and the conclusions remain un-

As an additional robustness check of our results, we conduct a placebo test and artificially introduce warning labels over a period predating the "intervention period" when the food labels were actually implemented. Our falsification test uses data for the same months for which we observe the actual rollout of the warning labels, but a year before when no warning labels had been introduced yet.[33] Table 8 presents the estimates of the placebo warning labels on consumer choice using the same Equation (1) for each product category. All warning label coefficients are not sizably different from zero in any product category, even at the 10% significance level, indicating an unlikely placebo warning label effect.[34]

3.4 Heterogeneous Effects and Household level Analysis

In order to study heterogeneity of the warning label effect, we estimate our main regression using household-level data. This approach complements the product-level analysis by examining whether the warning labeling effects depend on the socio-economic status and the presence of children in the household - both key variables for a nutritional labeling policy. We use a panel of households covering May, June, and July of 2015 and 2016.[35]

We estimate the main equation, extending the notation to include a household subscript i:

$$Y_{ijst} = \beta P_{jst} + \gamma L_{jst} + \eta_i + \nu_j + \zeta_s + \xi_t + \lambda_t + \varepsilon_{ijst} \tag{3}$$

where Y_{ijst} is a purchase indicator equal to 1 if household i purchased product j at time t in store s (0 otherwise). As the main explanatory variables, we have the warning label, L_{jst}. The specification also controls for price (P_{jst}) and household (η_i), product (ν_j), store (ζ_s), day-of-the-week (ξ_t), and week (λ_t) fixed effects. We compute two-way clustered standard errors by household and UPC, analogous to the product-level analysis.[36]

To build the category-specific household panel, we consider consumers with at least six trips (two per month approximately) in 2015 and six other trips in 2016.[37] For the product list, we consider the same UPCs as in the previous analysis.[38]

---

changed.

[33] According to the supermarket chain and press reports of the time (Mostrador, 2016), manufacturers began delivering products carrying the warning labels in March 2016.

[34] Additionally, Appendix E shows the 2015 trends of "never labeled" versus "labeled products at some point." However, these trends are difficult to interpret because "treatment" and "control" groups change over time (a UPC could be a labeled product and then reversed to be unlabeled in the same store).

[35] We use households that have purchased in a category in both years to identify individual households using the retailer loyalty card information. The main analyses use 2016 as it was done in the product-level analysis, leaving 2015 to include past purchases information used in the heterogeneity analysis.

[36] Results are very similar if adding week-UPC fixed effects (Table C2 Appendix C).

[37] We explore a different number of visits, and the conclusions remain unchanged.

[38] We drop transactions in the 99 percentile of the purchased quantities per trip to ensure we consider households

First, we use a socio-demographic characterization of the consumers. This measure (created by the retailer) combines the consumer address with census data to input the average income and other socio-economic attributes at the block level, as described in Section 2. Thus, we estimate the warning label effect using the sub-samples of the top and the remaining groups, denoted as "ABC1" and "No ABC1," respectively.

Table 9 shows the estimates for the two sub-samples for the breakfast cereals, chocolates, and cookies categories. In the cereal category, we find that the estimate of the warning label effect is almost three times larger for the low/middle income households relative to the top income group. The ABC1 group reduced the likelihood of purchasing by 0.11 percentage points (95%CI [-0.25 pp., 0.03 pp.]), and the rest by 0.30 percentage points (95%CI [-0.54 pp., -0.06 pp.]). However, the confidence intervals overlap, making any difference between these two groups unclear. Similar to the previous analyses with chocolates and cookies, the socio-economic sub-samples in these product categories do not show conclusive evidence on the warning label effect.[39]

As the second dimension of heterogeneity, we study households that are likely to have children. We have to infer this status from their purchase decisions, and thus, we chose a subset of cereals that, given their flavors and box design, are clearly intended for children (e.g., Froot Loops, Chocapic and Trix). We defined as a household with children those registered consumers that purchased these selected UPCs in the year 2015. We expect that their purchase behavior in 2016 will shed light on their response to food labeling.

Table 10 shows the estimates using the parent groups defined above. Our estimates suggest that the parent sub-sample negative response is slightly stronger in the breakfast cereal category, although their 95% confidence overlap with zero and their regions also overlap between the two groups. In particular, households that are more likely to have children have an estimate of -0.17 percentage points (p = 0.09 with 95%CI [-0.36 pp., 0.03 pp.]). In the other product categories, the estimates show inconclusive evidence of the label and are quite similar regardless the presence of children. We explore alternative definitions for households with children with similar conclusions.

In Appendix C, we also show the results of this household-level analysis for the average household – i.e., using the whole sample. As a consistency check, results are remarkably similar to the product-level analysis. In the same Appendix, we present a placebo test similar to the one conducted in the previous subsection, but using the household-level panel data, finding very small label coefficients and not sizably different from zero.[40]

---

and not wholesalers.

[39]To have a reference, the average baseline probability of purchase a UPC in cereals, chocolates and cookies are 1.9%, 1.1%, and 1.4%, respectively.

[40]We also examined households' decisions regarding purchased quantity conditional on purchase. However, the standard errors are very large – the number of observations is 0.1% of the observations when using all available products in the dataset. In addition, quantities are difficult to include to measure the label effect as a product may have different volumes and even be sold in packs, and they would be identified as different UPCs.

Finally, at the product and household levels, both analyses identify an average treatment effect of the warning label, assuming a homogeneous effect, as a standard policy evaluation. However, these analyses cannot determine whether consumers substitute healthy for unhealthy products or whether they prefer an outside option. The following section uses structural modeling to study the warning label effect accounting for the substitution patterns across UPCs and conduct a counterfactual analysis.

4 Demand Model

In this section, we use a structural demand model to evaluate the effectiveness of the warning-label policy. Unlike other policy options (e.g., nutrient taxes), which might not be exclusively motivated by health-related concerns,[41] the labeling policy we study was explicitly aimed at curbing the consumption of products the law identifies as unhealthful. Thus, to assess the effectiveness of the policy, we estimate a flexible discrete choice model –which allows for heterogeneous treatment effects while accounting for substitution patterns and point of sales factors (e.g., pricing)– and use it to compare the predicted demand patterns under the labeling policy and those predicted under a counterfactual scenario where no warning labels are implemented. This scenario allows examining the effect of the labeling policy, observing substitutions between healthy and unhealthy products and inside and outside a product category, and for different subpopulations. In addition to estimating the changes in demand patterns under label removal, we estimate a tax equivalent to the labeling policy; that is, we compute the ad valorem tax rate that would have to be applied on objectionable products to mimic the demand patterns predicted by our model under the actual implementation of the policy.

4.1 Model

Our setting considers a consumer facing a choice between J inside products in a category and an outside option (labeled $j = 0$). The inclusion of a no-purchase alternative allows us to investigate the effects of warning labels on category contraction or expansion. We assume that household choices are driven by product prices and labels, in addition to time-invariant store-product attributes and other time-specific product characteristics. Importantly, whether a product displays a warning label is included as an attribute that may impact households' choices. More formally, we specify the following conditional indirect utility function:

$$\begin{aligned} u_{ijst} &= \alpha_{ij} + \beta_i p_{jst} + \gamma_i L_{jst} + \theta_{jt} + \delta_{js} + \varepsilon_{ijst}, \qquad && j = 1, ..., J \\ u_{ijst} &= \varepsilon_{ijst}, && j = 0 \end{aligned} \tag{4}$$

[41] Seiler et al. (2021) studied the impact of a tax on sweetened beverages implemented in Philadelphia, which was primarily aimed at increasing tax revenue.

where p_{jst} is the price of product j faced by household i on a visit to store s at time t, L_{jst} is a dummy variable which takes the value 1 if product j displays a warning label at store s and time t.[42]

The product-week fixed effects, θ_{jt}, may capture national marketing campaigns and other product-specific activities that change over time but are common across stores. For example, we could account for advertisement in a particular brand of Easter eggs, and still, identify the warning-label effects in that weekend as long as we have stores with and without the warning labels for that specific product. The product-store fixed effects, δ_{js}, capture factors such as the time-invariant location characteristics that affect products differently such as the consumer demographics. Finally, ε_{ijst} is an i.i.d extreme value type I error term. We normalize the deterministic component of the utility of the outside good to zero.

The key parameter in the household's indirect utility function is the warning label coefficient, γ_i, that captures the intensity of the household's like/dislike for the presence of (one or more) warning labels on the front pack of product j. Our model is agnostic about the specific mechanisms driving the household's marginal utility associated to the warning labels. For instance, the household's taste for the warning labels might be a function of its concern for healthful eating, information about the product's excessive content of specific critical nutrients being disclosed by the warning labels, or concern about the product flavor communicated through the warning label. The coefficient might also be sensitive to how well informed a given household is with regard to the product's critical nutrients (or overall healthful status) prior to facing the warning labels. Thus, the coefficient might be more negative, for instance, if the presence of warning labels in a product changes prior expectations.

Our choice to introduce the effects of warning labels as a simple dichotomous variable –instead of accounting for the different permutations of warning label types– is grounded on the limited variation in warning label combinations we observe in the data and previously shown. For instance, among the top 10 breakfast cereal products (by market share), only one of them is required to display both a 'high in' calories and a 'high in' sugars label (Nestlé Trix), the rest of them being required to display only a calorie label. Similarly, we observe no variation in label combinations among the top 10 products in the chocolates and cookies categories. Thus, it turns out to be unfeasible to separately identify the effects of alternative warning label combinations from the effects of labels on individual products.

Our specification allows for the alternative-specific constants α_{ij} as well as the sensitivities to warning labels γ_i, and prices β_i, to vary over households. In principle, we could allow the taste coefficients associated to the warning labels to vary across alternatives as well, so as to account for a differential impact of warning labels across different products (e.g., the warning labels for

[42] While in principle the warning label type (i.e., high in sugars, calories, sodium or saturated fat) might affect choice, we observe almost no variation on these dimensions across products within a given category (e.g., all cereals with one label display a high in calories label).

specific products might be more informative for certain households). However, the fact that some of the products (i.e., the healthy ones in the choice set) do not get labeled over the whole sampling period (and beyond) precludes us from identifying alternative-specific warning label coefficients.

A usual concern in demand estimation is the potential for price endogeneity. In our setting, however, prices are identical for consumers in the same store (no coupons are used by the retailer). Moreover, store managers have little discretionality to set prices as the retailer follows a mostly national pricing policy. In our data, cross-store price variation within a category accounts for between 1.2 percent (chocolates and cookies) and 2.2 percent (breakfast cereals) of total price variation. To the extent that the retailer sets prices based on unobservables (to the researcher), our structural model would still be subject to price endogeneity when examining substitution patterns. We mitigate the problem by using weekly-product intercepts (θ_{jt}) to control for weekly product-specific characteristics, as suggested in Chintagunta et al. (2005).[43] The fact that estimated price elasticities are within the range of previously reported estimates (see Table 11) suggests that price endogeneity (which typically leads to an attenuation bias)[44] might not be a major issue in our setting.

We adopt a hierarchical Bayesian approach (Rossi et al., 2005) to model taste heterogeneity between households. Letting $\Theta^i = \{\beta_i, \gamma_i, \{\alpha_{ij}\}_{j=1}^{j=J}\}$ be a vector containing the random coefficients in the model, we specify first-stage normal priors for $\{\Theta^i\}_{i=1}^N$:

$$\mathbf{\Theta^i} \sim N(\boldsymbol{\mu}, \mathbf{\Sigma}) \tag{5}$$

We denote by $\sigma = (\sigma_\beta, \sigma_\gamma, \sigma_{\alpha_{1,\ldots,}\sigma_{\alpha_J}})$ the population standard deviations (e.g., $\sigma_\beta = \sqrt{\Sigma_{11}}, \sigma_\gamma = \sqrt{\Sigma_{22}}$) and allow $\mathbf{\Sigma}$ to be non-diagonal. We follow Rossi et al. (2005) and specify normal and inverse-Wishart priors on $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$, respectively, at the second stage.

As discussed in the reduced form analysis above, the identification of the warning label preference parameters comes from the staggered roll out of warning labels across stores. Identification of the γ_i parameters relies on the variation of the warning over time and in different stores. For a given product in given day we observe stores where the product displays the warning label as well as stores where the same product does not. Differences in consumer purchasing probabilities between and within stores allow us to identify the warning-label effect.

Identification of the price coefficients, β_i, relies on the observed price variation across time, stores and products. In effect, we observe price promotions (i.e., temporary price reductions) that differ primarily over time and across products.

[43] Rossi (2014) presents a critical assessment of alternative strategies based on IV and control function approaches.

[44] In principle, the bias arising from a correlation between prices and an unobserved error component can go either way. In their meta-analysis on the determinants of price elasticities, Bijmolt et al. (2005) find that the elasticity is larger in magnitude if endogeneity is accounted for (-3.74) than when it is not (-2.47)

4.2 Estimation and Results

We estimate the model at the category level (i.e., breakfast cereals, cookies, and chocolates), as we did in the previous section. In each category, we select the ten most important products (at the UPC level) by market share and define the outside option as visits to the supermarket store which do not include purchases of any of the products in the category. We group other less popular products in a category into two composite goods aggregating, each one of them, products displaying and not displaying warning labels in a given purchase occasion (i.e., a particular store-day can have different of these composited products). As is standard in the literature (e.g., Dubé et al., 2010, Simonov et al., 2020), we use Monte Carlo Markov Chain (MCMC) methods to estimate our demand model. Given diffuse standard priors for the parameters, we take draws through a Gibbs sampler with a Metropolis-Hastings step (Train (2009)).[45]

Table 11 reports summary statistics for the posterior mean of the first-stage prior for the warning label and price coefficients. The table also reports price elasticities and warning labels semi-elasticities to facilitate comparisons across categories. Warning label semi-elasticities are computed as the difference between choice probabilities obtained when a given targeted product is labeled according to the actual label implementation –while all other products remain unlabeled– and the corresponding choice probabilities when none of the products are labeled over the duration of the data.

The results are broadly consistent with our findings from the previous section. We find a significantly negative effect of warning labels in the breakfast cereals category. The point estimate (i.e., the mean of the mean posterior) for the warning label parameter equals -0.642, and the 95% credible interval ranges between -1.07 and -0.207. The associated average semi-elasticities, across "unhealthy," unlabeled products, equal -36.6% with a 95% credible interval ranging between -45.9% and -25.8%. We stress that these numbers represent an upper bound to the consumer response to the warning labels as they result from comparing a situation in which a single focal product is labeled with a situation in which none of the products displays a warning label. In contrast, we do not find significant results for the effects of warning labels in chocolates and cookies. The point estimates equal 0.104 and 0.004 for chocolates and cookies, respectively, with 95% credible intervals ranging between -0.055 and 0.081 in the case of chocolates, and between -0.032 and 0.054 in the case of cookies. Thus, we cannot reject the hypothesis that warning labels are ineffective in the case of these two categories.

As expected, point estimates for the price coefficients and own-price elasticities are negative across all three categories, and the associated 95% credible intervals do not include zero in any of the three categories. Own-price elasticities are within the range of previously reported figures: Approximately -3.02, -3.05 and -4.10 for cereals, chocolates and cookies, respectively. For instance,

[45]We use 800,000 draws and observe convergence after 400,000 draws (warm-up). To assess convergence of the posterior distribution, we visually inspect the sequences of draws and experiment with adding more draws to ensure estimates stability.

Bijmolt et al. (2005) report an average price elasticity of -3.74 across several categories and Nevo (2001) finds price elasticities on the order of -3 for the breakfast cereal category, in line with our mean estimate for the same category.

The results also suggest that the covariance between price and warning-label sensitivity is only relevant in the breakfast cereal category. The point estimate equals -0.481 with a 95% credible interval of [-0.794, -0.186], suggesting that more price sensitive individuals tend display a lower distaste for the warning labels.[46] In the case of the chocolates and cookies categories we cannot reject the hypothesis that the price and warning-label coefficients are uncorrelated as the 95% credible intervals for the covariance between the coefficients include zero in both cases.

Figure 4 examines the distribution of warning-label taste coefficients across households for the three product categories. We observe important differences in distribution of distaste coefficients across categories. Warning label coefficients show substantial heterogeneity across households in the breakfast cereals category with most of the probability mass concentrated below zero. In the chocolate category, in contrast, while the distaste for warning labels varies to an important extent across households, most of the probability mass is concentrated above zero indicating that the majority of households derives positive utility from the presence of warning labels. Finally, in the cookies category we observe a smaller degree of heterogeneity with a similar mass of probability above and below zero with the distribution peaking in the negative region.

4.3 Impact of the Labeling Policy on the Demand for Unhealthful Products

We use our estimates to assess the effects of the labeling policy on the demand for healthful and unhealthful products. We focus on the breakfast cereal category, as it is the single category in which we consistently find significant warning-label demand effects. While our results are inconclusive regarding the effects of the policy on cookies and chocolates, the average demand effects of warning labels in these two categories are substantially less negative than those we find for the cereal category (they are either close to zero or lie in the positive range). Thus, we argue that the demand effects of warning labels in the breakfast cereal category are likely to provide us with an upper bound on the effectiveness of the policy among the categories we study.

We compare the demand for cereals predicted by our model with the predicted demand in a counterfactual scenario where the labeling policy is not implemented (i.e., all labels from unhealthful products are removed). Since the goal of our counterfactual analysis is to assess the effectiveness of the policy in shifting demand away from products mandated to display warning labels, we focus the analysis on the last weeks of the data, after July 1st, when the law already had

[46] At first glance, this negative correlation may seem to contradict the finding that the wealthiest (ABC1) group was less sensitive to the warning label in the household-level analysis. However, both sets of estimates are not strictly comparable. On the one hand, the structural model considers a subset of products and transactions and identifies the parameters based on purchase behavior only. On the other hand, the reduced form estimates use all transactions and products and rely on the retailer's socio-economic categorization. For instance, several price-sensitive consumers in the ABC1 group could explain this result.

come into force and labels were implemented in almost all top products (99.4%).

The results, presented in Table 12, show that the demand for unhealthy cereal products contracts substantially under the actual labeling policy compared to a scenario in which no policy is enacted. We observe a similarly strong demand contraction across all top cereal products. On average, column (1) indicates that the share of top labeled products shows a reduction of 16.7% relative to a situation where no warning labels are imposed on unhealthy products. Part of that lower demand for labeled products goes to non-labeled products within the cereal category –the shares of non-targeted products goes up by approximately 9.1%– and part goes to the outside option –the share of the outside alternative increases by approximately 7.7%. Therefore, consumers both substituted unhealthy cereal products with more healthy ones, as defined by the law, and reduced their purchase in the breakfast cereal category.

We further investigate how the impact of removing the warning labels varies across different consumer segments. As in the previous section, we focus on two groups based on socio-economic status and the presence of children in the household. We find that the impact of the warning label policy is similarly strong across all market segments (see Columns (2)-(5) in Table 12).

Tax-Equivalent to the Labeling Policy. To provide a further perspective on the impact of the labeling policy, we compute the counterfactual ad-valorem tax that would have to be levied on labeled products in order to replicate the demand configuration predicted by our model under the labeling policy. As before, we focus our analysis on the last four weeks of the data when labels have already been implemented across the top products and stores. We compute the percentage increase in prices of unhealthy cereals (i.e., those exceeding the regulatory thresholds) in a scenario characterized by the absence of labels which would most closely match –in a minimum mean square error sense– the predicted demand when labels are in place across 180 different store-time combinations. We find that a tax of 15.3% levied on unhealthy, "high-in" cereals would have the same effects as the warning label policy on cereal demand. This reinforces the conclusion that the policy was effective in shifting demand away from unhealthy cereals. By way of comparison, Seiler et al. (2021) find that a tax of 1.5 cents per ounce on sweetened beverages in Philadelphia leads to a 34% price increase and a 46% reduction in purchased quantities. Therefore, our findings provide a tax equivalent of the warning label that may yield a moderate effect, considering also, it may require a relevant degree of salience, as the labels have (Chetty et al., 2009).

5 Conclusions

Providing consumers with simplified nutritional information is an increasingly favored policy option to induce healthier food choices (Hawkes (2010)). In this paper, we study the effects of a comprehensive nutrition labeling law enacted in Chile, which mandated the introduction of front-of-pack labels warning of the high levels of calories, sugars, sodium, and saturated fats contained in frequently-bought packaged goods. There was a strong and divided reaction in the industry as the law added uncertainty on how consumers would respond to the new labeling. Whether consumers react to the implemented warning label regulation has profound managerial implications for the food industry, and policy lessons for the health authorities worldwide as the Chilean regulation has been singled out as an ambitious policy and followed closely by several countries.

A distinctive feature of our empirical setting is the rich variation we observe in the display of warning labels across time and stores. This variation allows us to overcome a traditional challenge of identifying the effects of nutrition labeling policies on consumer choices when they are fully implemented in a single period.[47]

Our analysis focused on three product categories in which most products were affected by the regulation. Our results indicate that consumers reduced their purchases of labeled breakfast cereals while the noisy estimates provide inconclusive evidence for cookies and chocolates. This study does not examine the underlying mechanisms that explain the observed purchase patterns despite the more substantial effect on cereals being consistent with labels changing consumer decisions when providing new or unexpected information about foods' nutritional content.

Furthermore, the warning labels not only convey information (MinSal (2018)) but they also attempt to nudge healthy behavior. Thus, our results per category are in line with research using more informative labels on purchases (Hobin et al., 2017) and consumers' intentions (Ikonen et al., 2020). However, there are other differences across these categories; the regulation severity allowed for non-labeled breakfast cereal products (one-third approximately), whereas almost all UPCs in the chocolate and cookie categories ended up tagged as unhealthful products, and the inclination for *taste* may affect more products that are considered indulgent (Ikonen et al., 2020).

Our results also suggest that purchases by medium socioeconomic groups and families with children are susceptible to be modified by the provision of simplified nutritional information. These findings are highly relevant for policymakers who typically target both groups, given their higher risk of developing obesity (especially given the alarming obesity rates among children (IHME (2013))). The effectiveness of the Chilean warning-label policy among medium socioeconomic households could be driven by the fact that, in our setting, prices did not play a significant role as unlabeled and labeled cereals displayed similar price levels. Hence, substituting away from labeled breakfast cereals was not seriously affecting household expenditures.

While our empirical approach allows us to identify consumer responses to nutrition labeling in natural market environments, there are some limitations. First, our study focuses on a single retail chain. To the extent that purchasing behavior and, in particular, the response to interpretative nutritional information may be different in other retailers, our results cannot be extrapolated

[47] In the online Appendix D, we show the warning labeling effect by comparing purchases after one week prior to the law came into force with the same period in 2015 as control. This is, without taking advantage of the staggered model and performing a *what-if* pre-post analysis between labeled and unlabeled UPCs. Indeed, the results are very different for the cereal and cookies categories.

to the population at large. We should emphasize, however, that our analysis used stores from the two most populated regions of the country, and the focus on one retail chain does not compromise the internal validity of our findings. A second limitation is that we quantify the short-run impact of the intervention over the first few months of its introduction. Hence, we are unable to capture learning effects that may be taking place over a longer time horizon. However, this long-term effect may include many other elements other than the warning labels that could have affected food purchases. The law included other components beyond the warning label, such as prohibitions to sell labeled products in schools, bans on the advertising of tagged products targeting children, and the removal of cartoons from cereal boxes (NYT (2018)), which were all implemented after June 2016. Even though all these changes were relevant for this particular policy (Taillie et al. (2020b)), they do not allow separately identifying the effect of the warning labels on consumer choices. Moreover, the evidence showed that suppliers did not change the product formulation before the law came into force, which they did afterwards (Kanter et al. (2019); Reyes et al. (2020); Barahona et al. (2020); Alé-Chilet and Moshary (2020)) as well as their prices (Pachali et al., 2020). Therefore, the present study manages to identify the impact of the warning label on purchase patterns. Our results stress that warning labels might not be sufficient to curb the purchase behavior of unhealthful products, and additional measures, such as banning unhealthy products from schools, seem to complement the disclosure of food information. To tease apart the successful channels required to change food purchase behavior is necessary to understand the consumer responses to the warning labels. The effect of this type of information will help assess one of the main components of nutritional labeling policies and an essential marketing strategy for point of sale advertising used in retail chains worldwide.

References

Alé-Chilet, J. and Moshary, S. (2020). Beyond consumer switching: Supply responses to food packaging and advertising regulations. *Available at SSRN 3678744*.

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

Barahona, N., Otero, C., Otero, S., and Kim, J. (2020). Equilibrium effects of food labeling policies. *Available at SSRN 3698473*.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, 119(1):249–275.

Bijmolt, T., Van Heerde, H. J., and Pieters, R. G. (2005). New empirical generalizations on the determinants of price elasticity. *Journal of Marketing Research*, XLII:141–156.

Bleich, S., Economos, C., Spiker, M., Vercammen, K., VanEpps, E., Block, J., Elbel, B., Story, M., and Roberto, C. (2017). A systematic review of calorie labeling and modified calorie labeling interventions: Impact on consumer and restaurant behavior. *Obesity*, 25(12):2018–2044.

Bollinger, B., Hammond, D., Hobin, E., Liebman, E., and Sacco, J. (2019). Do educational campaigns for on-shelf nutritional labeling work? *Available at SSRN 3500981*.

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

Boseley, S. (2020). Chile's drastic anti-obesity measures cut sugary drink sales by 23%. *The Guardian*.

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

Cawley, J., Hanks, A. S., Just, D. R., Wansink, B., et al. (2016). Incentivizing nutritious diets: A field experiment of relative price changes and how they are framed. *NBER Working Paper 21929*.

CDC (2019). Childhood Obesity Facts.

CERET (2016). Medicion de calidad de servicio en la industria del retail supermercados. *Centro de Estudios del Retail (CERET), Depto Ing Industrial, U de Chile.*

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–77.

Chintagunta, P., Dubé, J.-P., and Goh, K. Y. (2005). Beyond the endogeneity bias: The effect of unmeasured brand characteristics on household-level brand choice models. *Management Science*, 51(5):832–849.

Correa, T., Reyes, M., Taillie, L. S., Corvalán, C., and Dillman Carpentier, F. R. (2020). Food advertising on television before and after a national unhealthy food marketing regulation in chile, 2016–2017. *American Journal of Public Health*, (0):e1–e6.

Corvalán, C., Reyes, M., Garmendia, M. L., and Uauy, R. (2013). Structural responses to the obesity and non-communicable diseases epidemic: the chilean law of food labeling and advertising. *Obesity Reviews*, 14(S2):79–87.

Cowburn, G. and Stockley, L. (2005). Consumer understanding and use of nutrition labelling: a systematic review. *Public Health Nutrition*, 8(1):21–28.

Downs, J. S., Loewenstein, G., and Wisdom, J. (2009). Strategies for promoting healthier food choices. *The American Economic Review Papers and Procedings*, 99(2):159–164.

Dubé, J.-P., Hitsch, G., and Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. *RAND Journal of Economics*, 41(3):417–445.

Dumanovsky, T., Huang, C. Y., Nonas, C. A., Matte, T. D., Bassett, M. T., and Silver, L. D. (2011). Changes in energy content of lunchtime purchases from fast food restaurants after introduction of calorie labelling: cross sectional customer surveys. *The British Medical Journal*, 343:d4464.

Elbel, B., Kersh, R., Brescoll, V. L., and Dixon, L. B. (2009). Calorie labeling and food choices: a first look at the effects on low-income people in new york city. *Health Affairs*, 28(6):w1110–w1121.

Elshiewy, O. and Boztug, Y. (2018). When back of pack meets front of pack: How salient and simplified nutrition labels affect food sales in supermarkets. *Journal of Public Policy & Marketing*, 37(1):55–67.

Feunekes, G. I., Gortemaker, I. A., Willems, A. A., Lion, R., and Van Den Kommer, M. (2008). Front-of-pack nutrition labelling: testing effectiveness of different nutrition labelling formats front-of-pack in four european countries. *Appetite*, 50(1):57–70.

Finkelstein, E. A., Strombotne, K. L., Chan, N. L., and Krieger, J. (2011). Mandatory menu labeling in one fast-food chain in King County, Washington. *American Journal of Preventive Medicine*, 40(2):122–127.

Griffith, R. and Nevo, A. (2019). Marketing and public policy. In Dubé, J.-P. and Rossi, P. E., editors, *Handbook of the Economics of Marketing, Vol. 1*, chapter 10, pages 553–596. North-Holland.

Harnack, L. J. and French, S. A. (2008). Effect of point-of-purchase calorie labeling on restaurant and cafeteria food choices: A review of the literature. *International Journal of Behavioral Nutrition and Physical Activity*, 5(1):51.

Hawkes, C. (2004). Nutrition labels and health claims: the global regulatory environment. In *Nutrition Labels and Health Claims: the Global Regulatory Environment*. WHO.

Hawkes, C. (2010). *Government and Voluntary Policies on Nutrition Labelling: A Global Overview*, chapter 4, pages 37–58. The Food and Agriculture Organization of the United Nations and Woodhead Publishing Series in Food Science, Technology and Nutrition.

HC (2017). Front-of-package nutrition labelling: September 18, 2017 Stakeholder Engagement Meeting. *Ministry of Health, Canada*.

Hobin, E., Bollinger, B., Sacco, J., Liebman, E., Vanderlee, L., Zuo, F., Rosella, L., L'abbe, M., Manson, H., and Hammond, D. (2017). Consumers' response to an on-shelf nutrition labelling system in supermarkets: Evidence to inform policy and practice. *The Milbank Quarterly*, 95(3):494–534.

IHME (2013). The global burden of disease: Generating evidence, guiding, policy. *Institute for Health Metrics and Evaluation*.

Ikonen, I., Sotgiu, F., Aydinli, A., and Verlegh, P. W. (2020). Consumer effects of front-of-package nutrition labeling: An interdisciplinary meta-analysis. *Journal of the Academy of Marketing Science*, 48(3):360–383.

Ippolito, P. M. and Mathios, A. D. (1995). Information and advertising: The case of fat consumption in the united states. *The American Economic Review Papers and Proceedings*, 85(2):91.

Jacobs, A. (2018). In a sweeping war on obesity, Chile slays Tony the tiger. *New York Times*.

Kanter, R., Reyes, M., Vandevijvere, S., Swinburn, B., and Corvalán, C. (2019). Anticipatory effects of the implementation of the chilean law of food labeling and advertising on food and beverage product reformulation. *Obesity Reviews*.

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

Kiszko, K. M., Martinez, O. D., Abrams, C., and Elbel, B. (2014). The influence of calorie labeling on food orders and consumption: A review of the literature. *Journal of Community Health*, 39(6):1248–1269.

Loewenstein, G., Sunstein, C. R., and Golman, R. (2014). Disclosure: Psychology changes everything. *Annual Review of Economics*, 6:391–419.

Loring, B. and Robertson, A. (2014). Obesity and Inequities. *World Health Organization*, page 48.

Malnick, S. D. and Knobler, H. (2006). The medical complications of obesity. *Journal of the Association of Physicians*, 99(9):565–579.

MinSal (2018). Informe de evaluación de la implementación de la ley sobre composición nutricional de los alimentos y su publicidad. Technical report, Chile. Ministerio de Salud. Subsecretaria de Salud Pública. División de Políticas Públicas Saludables y Promoción.

Mostrador (2016). Ley de etiquetados de alimentos. *Diario El Mostrador*.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342.

Nikolova, H. D. and Inman, J. J. (2015). Healthy choice: the effect of simplified point-of-sale nutritional information on consumer food choice behavior. *Journal of Marketing Research*, 52(6):817–835.

NYT (2018). In sweeping war on obesity, Chile slays Tony the tiger. *New York Times*.

Pachali, M. J., Kotschedoff, M. J., van Lin, A., Bronnenberg, B. J., and van Herpen, E. (2020). Do warning labels make products more differentiated and lead to higher prices? *Available at SSRN 3642756*.

Raghunathan, R., Naylor, R. W., and Hoyer, W. D. (2006). The unhealthy= tasty intuition and its effects on taste inferences, enjoyment, and choice of food products. *Journal of Marketing*, 70(4):170–184.

Reyes, M., Garmendia, M. L., Olivares, S., Aqueveque, C., Zacarías, I., and Corvalán, C. (2019). Development of the chilean front-of-package food warning label. *BMC public health*, 19(1):906.

Reyes, M., Smith Taillie, L., Popkin, B., Kanter, R., Vandevijvere, S., and Corvalán, C. (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.

Rossi, P. E. (2014). Even the rich can make themselves poor. *Marketing Science*, 33(5):655–672.

Rossi, P. E., Allenby, G. M., and McCulloch, R. (2005). *Bayesian Statistics and Marketing*. Wiley.

Schwartz, D., Fischhoff, B., Krishnamurti, T., and Sowell, F. (2013). The hawthorne effect and energy awareness. *Proceedings of the National Academy of Sciences*, 110(38):15242–15246.

Seiler, S., Tuchman, A., and Yao, S. (2021). The impact of soda taxes: Pass-through, tax avoidance, and nutritional effects. *Journal of Marketing Research*, 58(1):22–49.

Simonov, A., Dubé, J.-P., Hitsch, G., and Rossi, P. E. (2020). State-dependent demand estimation with initial conditions correction. *Journal of Marketing Research*, 57(5):789–809.

Taillie, L. S., Hall, M. G., Popkin, B. M., Ng, S. W., and Murukutla, N. (2020a). Experimental Studies of Front-of-Package Nutrient Warning Labels on Sugar-Sweetened Beverages and Ultra-Processed Foods: A Scoping Review. *Nutrients*, 12(2):569.

Taillie, L. S., Reyes, M., Colchero, M. A., Popkin, B., and Corvalán, C. (2020b). 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.

TheGuardian (2020). Chile's drastic anti-obesity measures cut sugary drink sales by 23%. *The Guardian*.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, 2nd edition.

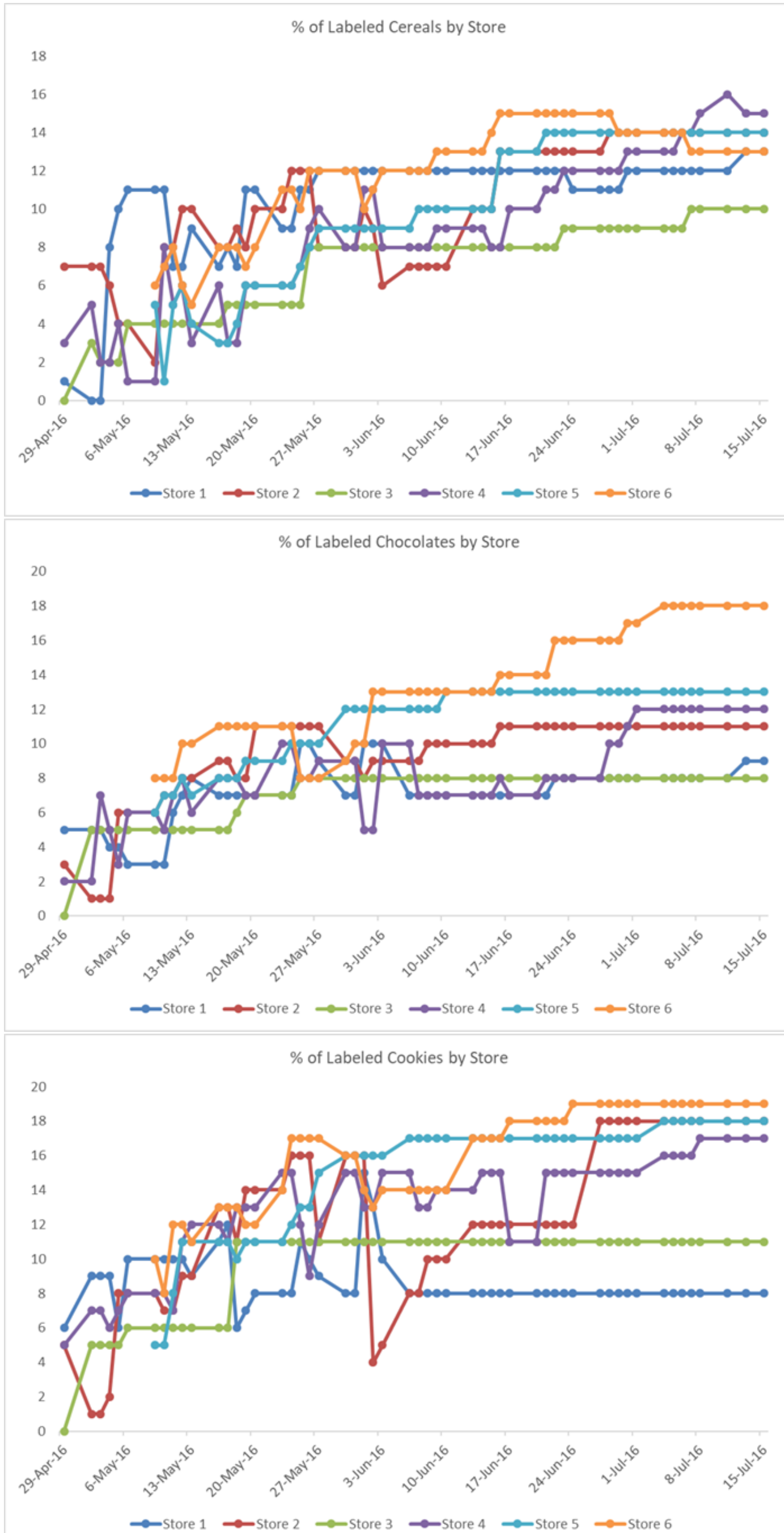
Valdebenito, M., Labrín, J. M., León, V., and Fierro, S. (2016). Descripción de las percepciones y actitudes de los/as consumidores respecto a las medidas estatales en el marco de la implementación del Decreto 13/15. Technical report, Ministerio de Salud (MINSAL) Chile.

WHO (2004). Global strategy on diet, physical activity and health. *World Health Organization Report*.

Figure 1: Warning Labels in Chile

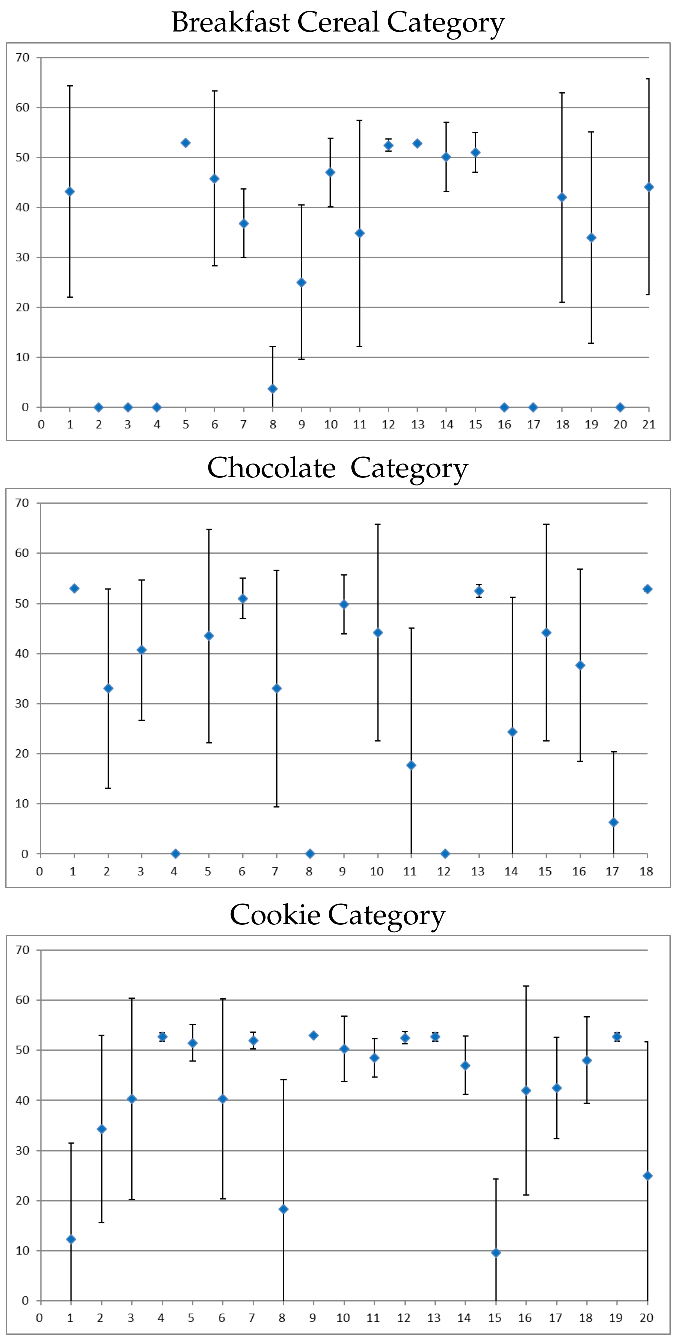
Notes: From left to right: High in Sugar, High in Calories, High in Saturated Fats and High in Sodium. At the bottom of each label it states Ministry of Health.

Figure 2: Evolution of the Number of Labeled products per store over time



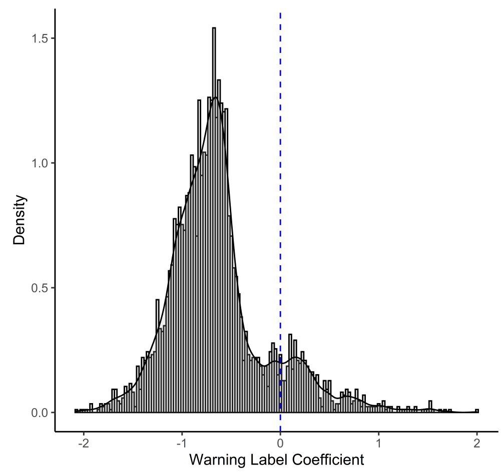
Notes: Y-axis is the number of labeled products (from the top-selling 20 products per category), X-axis is the time line in weeks. Different colors represent different stores.

Figure 3: Timing of Warning Label Implementation Across Stores

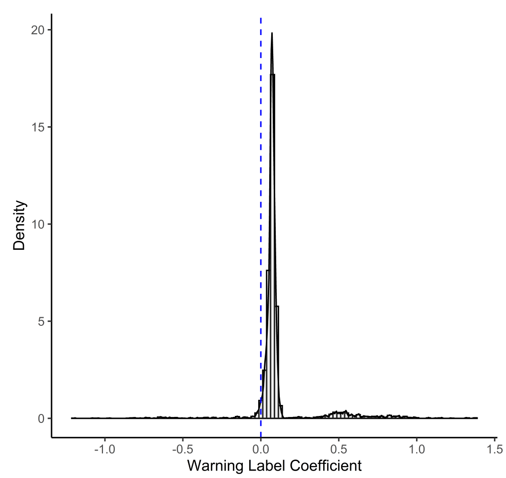


Notes: X-axis displays each of the top-selling 20 products per category. Y-axis is the number of days in advance of the actual implementation of the warning labels before the legal deadline. The blue dot represents the mean of the number of days in advance across stores for each product, and the error bars represent the corresponding variation (using the standard deviation). Dots at Y=0 correspond to unlabeled (healthful) products. Products are sorted by market share, being product 1 the UPC with the largest market share.

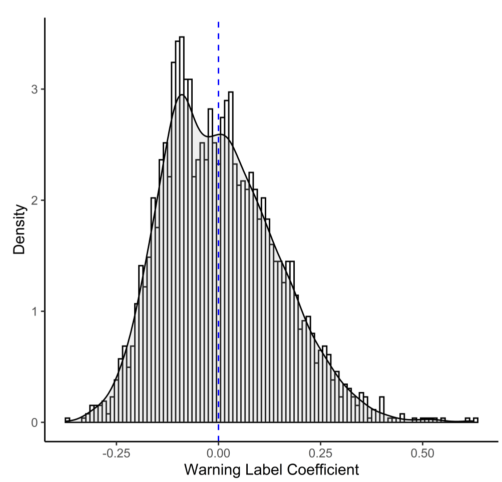
Figure 4: Distribution of Warning-Label Coefficients Across Households



(a) Breakfast Cereals



(b) Chocolates



(c) Cookies

Notes: The figure presents the distributions of household-specific taste coefficients for warning labels using a hierarchical Bayes model. Distributions based on 400,000 MCMC iterations. The number of households are 3,452, 3,671 and 2,623 for breakfast cereals, chocolates and cookies, respectively.

Table 1: Summary Statistics

Panel A. Product Statistics				
Category	Number of Products	Average Price	Price Std. Dev.	Av. Quantity per UPC-day-store
Breakfast Cereals	225	1715	878	19.5
Chocolates	308	1405	912	27.5
Cookies	423	846	623	30.8

Panel B. Household Statistics				
Category	Number of Households	Avg. Number of Trips	Median of Trips	Avg Quantity per UPC-day
Breakfast Cereals	4,801	25.2	21	1.4
Chocolates	16,117	25.1	21	1.4
Cookies	12,396	26.3	22	1.7

Panel C. Label Distribution					
	0	1	2	3	Total
Breakfast Cereals	75	52	71	27	225
UPC Share [%]	33.3	23.1	31.6	12.0	100
Market Share [%]	27.2	41.7	22.6	8.5	100
Chocolates	27	13	35	233	308
UPC Share [%]	8.8	4.2	11.3	75.7	100
Market Share [%]	3.1	1.8	7.2	87.9	100
Cookies	36	80	44	263	423
UPC Share [%]	8.5	18.9	10.4	62.2	100
Market Share [%]	1.1	12.2	4.5	82.2	100

Notes: Panel A presents details of product-level panel per category reporting number of products, average and standard deviation of price, and the average quantity sold per UPC-day-store. Prices in Chilean pesos (the average exchange rate in May-Jul 2016 was USD 1 = CLP 673.5). Panel B describes the household-level panel per category, reporting the number of households in the study period, their average and median number of trips, and their purchased quantity per UPC-day-household conditional on having had at least twelve trips in the data. Panel C shows the number of labeled UPCs per category and their respective fraction of products and market shares using UPC's number of labels in July 2016, after the law came into force. Note that not all products are sold in each store every day. Also no product has four labels in the categories we study.

Table 2: Label Effect on Breakfast Cereals at the Product-Level Analysis

Panel A. Dep Var: Log of Quantity			
	(1)	(2)	(3)
Label	-0.177***	-0.065**	-0.062**
	(0.037)	(0.021)	(0.018)
Price			-0.616***
			(0.147)
R-squared	0.574	0.638	0.672

Panel B. Dep Var: Percentage of Transactions			
	(1)	(2)	(3)
Label	-0.0011**	-0.0008*	-0.0007**
	(0.0004)	(0.0003)	(0.0003)
Price			-0.012**
			(0.004)
R-squared	0.568	0.571	0.605

Store-UPC FE	✓	✓	✓
Day-of-the-Week FE		✓	✓
Week FE		✓	✓
Number of Observations	12,891	12,891	12,891
Number of Stores	6	6	6
Number of Products	195	195	195

Notes: OLS estimates of Equation (1). As the dependent variable, Panel A uses the log of purchased quantities, while Panel B uses the percentage transactions in a given store-day combination containing UPC j. Two-way clustered-robust (at the store-UPC level) standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Label Effect on Chocolates at the Product-Level Analysis

Panel A. Dep Var: Log of Quantity			
	(1)	(2)	(3)
Label	0.013	0.039	0.016
	(0.029)	(0.020)	(0.016)
Price			-0.724***
			(0.114)
R-squared	0.715	0.732	0.751

Panel B. Dep Var: Percentage of Transactions			
	(1)	(2)	(3)
Label	0.0003	0.0003	0.0001
	(0.0009)	(0.0005)	(0.0004)
Price			-0.008*
			(0.004)
R-squared	0.748	0.749	0.756
Store-UPC FE	✓	✓	✓
Day-of-the-Week FE		✓	✓
Week FE		✓	✓
Number of Observations	12,094	12,094	12,094
Number of Stores	6	6	6
Number of Products	264	264	264

Notes: OLS estimates of Equation (1). As the dependent variable, Panel A uses the log of purchased quantities, while Panel B uses the percentage transactions in a given store-day combination containing UPC *j*. Two-way clustered-robust (at the store-UPC level) standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Label Effect on Cookies at the Product-Level Analysis

Panel A. Dep Var: Log of Quantity			
	(1)	(2)	(3)
Label	-0.015	0.019	0.019
	(0.023)	(0.018)	(0.016)
Price			-1.314***
			(0.264)
R-squared	0.662	0.683	0.701

Panel B. Dep Var: Percentage of Transactions			
	(1)	(2)	(3)
Label	0.0003	0.0004	0.0003
	(0.0003)	(0.0002)	(0.0002)
Price			-0.017**
			(0.006)
R-squared	0.630	0.631	0.642

Store-UPC FE	✓	✓	✓
Day-of-the-Week FE		✓	✓
Week FE		✓	✓
Number of Observations	24,925	24,925	24,925
Number of Stores	6	6	6
Number of Products	376	376	376

Notes: OLS estimates of Equation (1). As the dependent variable, Panel A uses the log of purchased quantities, while Panel B uses the percentage transactions in a given store-day combination containing UPC j. Two-way clustered-robust (at the store-UPC level) standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Number of labels

Panel A. Dep Var: Log of Quantity			
Dep Var: Price	Cereal	Chocolate	Cookies
One Label	-0.065*	0.178	-0.011
	(0.026)	(0.094)	(0.022)
Two Labels	-0.056*	0.005	-0.039
	(0.025)	(0.043)	(0.038)
Three Labels	-0.059	0.013	0.030
	(0.052)	(0.017)	(0.016)
Price	-0.616***	-0.724***	-1.313***
	(0.147)	(0.114)	(0.264)
R-squared	0.672	0.751	0.701

Panel B. Dep Var: Percentage of Transactions			
	Cereal	Chocolate	Cookies
One Label	-0.0006	0.0005	-0.00001
	(0.0004)	(0.0004)	(0.0003)
Two Labels	-0.0007	0.0001	-0.0004
	(0.0004)	(0.0003)	(0.0003)
Three Labels	-0.0009	0.0001	0.0005
	(0.0011)	(0.0005)	(0.0003)
Price	-0.012**	-0.008*	-0.017**
	(0.004)	(0.004)	(0.006)
R-squared	0.605	0.756	0.642
Number of Observations	12,891	12,094	24,925
Number of Stores	6	6	6
Number of Products	195	264	376

Notes: OLS estimates of Equation (1). As the dependent variable, Panel A uses the log of purchased quantities, while Panel B uses the percentage transactions in a given store-day combination containing UPC *j*. All specifications consider store-UPC fixed effects, fixed effects for each day of the week, and week fixed effects. Two-way clustered-robust (at the store-UPC level) standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Label Effect on Prices

Dep Var: Price	(1) Cereal	(2) Chocolate	(3) Cookies
Label	0.005	-0.031***	-0.0004
	(0.012)	(0.007)	(0.0039)
R-squared	0.957	0.974	0.985
Store-UPC FE	✓	✓	✓
Day-of-the-Week FE	✓	✓	✓
Week FE	✓	✓	✓
Number of Observations	12,891	12,094	24,925
Number of Stores	6	6	6
Number of Products	195	264	376

Notes: OLS estimates of Equation (2). Prices are in thousand of Chilean pesos. The average exchange rate in May-Jul 2016 was USD 1 = CLP 673.5. All specifications consider store-UPC, day of the week, and week fixed effects. Two-way clustered-robust (at the store-UPC level) standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Effect of Product Characteristics on Timing of Label Implementation

	(1)	(2)	(3)	(4)
Panel A. Cereals				
Previous Sales	0.774			0.179
	(1.983)			(2.915)
Previous Quantity		-0.0003		0.0003
		(0.003)		(0.004)
Previous Price			2.537	2.546
			(2.242)	(2.588)
Observations	220	220	220	220
Number of Products	80	80	80	80
Panel B. Chocolates				
Previous Sales	0.309			0.123
	(0.580)			(1.111)
Previous Quantity		0.0003		0.0002
		(0.0005)		(0.0011)
Previous Price			0.282	0.301
			(0.730)	(0.774)
Number of Observations	246	246	246	246
Number of Products	108	108	108	108
Panel C. Cookies				
Previous Sales	0.320			0.875
	(1.026)			(1.586)
Previous Quantity		0.0008		-0.00004
		(0.0006)		(0.0010)
Previous Price			-5.365***	-5.567***
			(1.522)	(1.724)
Number of Observations	616	616	616	616
Number of Products	199	199	199	199
UPC RE	✓	✓	✓	✓
Store FE	✓	✓	✓	✓

Notes: OLS estimates of Equation (1) when using as dependent variable the number of days in advance the product was first labeled in a store. Panel A, B and C uses the cereal, chocolate and cookie category, respectively. The explanatory variables are aggregated initial sales, quantities and mean price from each UPC-store; using data from May 2015 (or May 2016 for new products), before (or during the initial month) of the warning label implementation. Sales and prices in thousands and millions of Chilean pesos, respectively. All specifications consider UPC random effects and store fixed effects. Standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Placebo Label Effect

Panel A. Dep Var: Log of Quantity

	Cereal	Chocolate	Cookies
Label	-0.024	0.020	0.007
	(0.012)	(0.011)	(0.015)
Price	-0.830***	-1.070***	-1.453**
	(0.073)	(0.173)	(0.366)
R-squared	0.672	0.762	0.737

Panel B. Dep Var: Percentage of Transactions

	Cereal	Chocolate	Cookies
Label	-0.0005	0.0006	-0.0001
	(0.0004)	(0.0006)	(0.0004)
Price	-0.018***	-0.019*	-0.023**
	(0.003)	(0.009)	(0.008)
R-squared	0.564	0.752	0.731
Store-UPC FE	✓	✓	✓
Day-of-the-Week FE	✓	✓	✓
Week FE	✓	✓	✓
N. of Observations	10,895	11,141	24,362
Number of stores	6	6	6
Number of Products	165	234	327

Notes: OLS estimates of Equation (1). As the dependent variable, Panel A uses the log of purchased quantities, while Panel B uses the percentage transactions in a given store-day combination containing UPC j. All specifications consider store-UPC fixed effects, fixed effects for each day of the week, and week fixed effects. Two-way clustered-robust (at the store-UPC level) standard errors in parenthesis. P-values notation: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 9: Heterogeneous Effects on Purchasing Probability by Socioeconomic status

Dep Var: Purchase Indicator	Cereal		Chocolate		Cookies	
	ABC1	No ABC1	ABC1	No ABC1	ABC1	No ABC1
	(1)	(2)	(3)	(4)	(5)	(6)
Label	-0.0011	-0.0030*	0.0006	0.0004	0.0002	0.0005
	(0.0007)	(0.0012)	(0.0005)	(0.0007)	(0.0004)	(0.0005)
Price	-0.042***	-0.044***	-0.021***	-0.026***	-0.065***	-0.076***
	(0.008)	(0.008)	(0.005)	(0.006)	(0.010)	(0.012)
R-squared	0.013	0.016	0.016	0.021	0.016	0.018
N. of Observations	1,000,494	263,032	3,028,566	1,245,649	4,245,767	1,915,040
Number Households	3,302	1,153	10,141	4,821	7,545	3,893
Number of Products	224	225	308	308	423	423
Household FE	✓	✓	✓	✓	✓	✓
UPC FE	✓	✓	✓	✓	✓	✓
Store FE	✓	✓	✓	✓	✓	✓
Day-of-the-Week FE	✓	✓	✓	✓	✓	✓
Week FE	✓	✓	✓	✓	✓	✓

Notes: OLS estimates of Equation (3) using different subsamples of the household-level data based on socio-economic groups. Columns (1), (3), and (5) use households from the top income group, denoted ABC1, while Columns (2), (4), and (6) consider the rest of the households. Columns (1-2), (3-4) and (5-6) use the Cereal, Chocolate and Cookie category, respectively. The dependent variable is a purchase indicator for each UPC in a given household-store-day. All specifications consider household, UPC, store, week and day-of-the-week fixed effects. Two-way clustered standard errors at the household and product levels in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Heterogeneous Effects on Purchasing Probability by Children's presence

Dep Var: Purchase Indicator	Cereal		Chocolate		Cookies	
	Parent	No Parent	Parent	No Parent	Parent	No Parent
	(1)	(2)	(3)	(4)	(5)	(6)
Label	-0.0017+	-0.0011	0.0006	0.0005	0.0001	0.0004
	(0.0010)	(0.0007)	(0.0007)	(0.0006)	(0.0006)	(0.0004)
Price	-0.049***	-0.036***	-0.026***	-0.023***	-0.078***	-0.067***
	(0.009)	(0.006)	(0.006)	(0.005)	(0.012)	(0.010)
R-squared	0.016	0.015	0.020	0.017	0.023	0.015
N. of Observations	642,888	639,168	1,145,460	1,580,012	1,872,120	2,315,909
Number Households	2,142	2,341	3,966	5,471	3,302	4,256
Number of Products	225	225	308	308	423	423
Household FE	✓	✓	✓	✓	✓	✓
UPC FE	✓	✓	✓	✓	✓	✓
Store FE	✓	✓	✓	✓	✓	✓
Day-of-the-Week FE	✓	✓	✓	✓	✓	✓
Week FE	✓	✓	✓	✓	✓	✓

Notes: OLS estimates of Equation (3) using different subsamples of the household-level data based on the presence of children. Columns (1), (3), and (5) use households that have previously purchased children products, while Columns (2), (4) and (6) consider the rest of the households. Columns (1-2), (3-4) and (5-6) use the Cereal, Chocolate and Cookie category, respectively. The dependent variable is a purchase indicator for each UPC in a given household-store-day. All specifications consider household, UPC, store, week and day-of-the-week fixed effects. Two-way clustered standard errors at the household and product levels in parenthesis. P-values notation: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 11: Discrete Choice Model Estimates

Panel A. Breakfast Cereals

	Mean, μ		Standard deviation, σ		Avg. Elast./Semi-elast.	
Random parameter	Posterior Mean	95% CI	Posterior Mean	95% CI	Mean	95% CI
	(1)	(2)	(3)	(4)	(5)	(6)
Price	-1.6720	[-2.1093, -1.3096]	0.9743	[0.6619, 1.2300]	-3.0156	[-3.3829, -2.7363]
Label	-0.6417	[-1.0690, -0.2068]	1.1195	[0.8025, 1.4420]	-0.3655	[-0.4594, -0.2578]
Covariance ($\sigma_{\beta\gamma}$)	-0.4808	[-0.7936, -0.1861]				
Avg. log-likelihood	-21,457.2				Number of households	3,452
Null log-likelihood	-66,106.4				Number of trips	25,773

Panel B. Chocolates

	Mean, μ		Standard deviation, σ		Avg. Elast./Semi-elast.	
Random parameter	Posterior Mean	95% CI	Posterior Mean	95% CI	Mean	95% CI
	(1)	(2)	(3)	(4)	(5)	(6)
Price	-1.2439	[-1.4823, -0.9728]	0.3012	[0.2112, 0.3926]	-3.0462	[-3.1735 , -2.9360]
Label	0.1078	[-0.1846, 0.3452]	0.7054	[0.4082, 1.0341]	0.0250	[-0.0547, 0.0808]
Covariance ($\sigma_{\beta\gamma}$)	-0.0156	[-0.0930, 0.0570]				
Avg. log-likelihood	-15868.5				Number of households	3,671
Null log-likelihood	-56970.1				Number of trips	22,211

Panel C. Cookies

	Mean, μ		Standard deviation, σ		Avg. Elast./Semi-elast.	
Random parameter	Posterior Mean	95% CI	Posterior Mean	95% CI	Mean	95% CI
	(1)	(2)	(3)	(4)	(5)	(6)
Price	-3.6887	[-3.9447, -3.4346]	1.7519	[1.3203, 2.1480]	-4.1145	[-4.3067, -3.9378]
Label	0.0038	[-0.0844, 0.0991]	0.3956	[0.3019, 0.4998]	0.0111	[-0.0323 , 0.0542]
Covariance ($\sigma_{\beta\gamma}$)	-0.0013	[-0.2963, 0.2562]				
Avg. log-likelihood	-24,573.2				Number of households	2,623
Null log-likelihood	-60,889.3				Number of trips	23,739

Notes: Columns (1)-(4) present summary statistics for the mean posterior density of the warning label and price coefficients in the breakfast cereals, chocolates and cookies categories. Columns (5) and (6) present descriptive statistics on average own-price elasticities and warning labels semi-elasticities. Warning labels semi-elasticities indicate the percentage change in choice probabilities obtained from comparing a situation in which a given (eventually labeled) product is labeled according to the actual policy rollout and a situation in which no product is labeled throughout the period. We run the MCMC procedure for 800,000 iterations and use the last 400,000 for estimation. Elasticity and semi-elasticity estimates use 10,000 draws from the simulated posterior distributions. Following Equation (4), all estimations include product-store and product-week fixed effects, in addition to price and a warning label indicator.

Table 12: Effects of Warning Labels Removal in the Breakfast Cereal Category

Product Description	Predicted % Change in Demand All Customers (1)	ABC1 (2)	Non-ABC1 (3)	Parents (4)	Non-Parents (5)
Labeled Products					
Quaker Oatmeal Squares 370 g.	-0.1753	-0.1834	-0.1649	-0.1806	-0.1692
Nestlé Fitness & Yogurt 490 g.	-0.1583	-0.1637	-0.1524	-0.1605	-0.1555
Quaker Granola Honey & Almonds 380 g.	-0.1745	-0.1817	-0.1665	-0.1781	-0.1699
Nestlé Fitness Honey & Almonds 480 g.	-0.1679	-0.1741	-0.1605	-0.1703	-0.1650
Nestlé Trix 480 g.	-0.1753	-0.1812	-0.1690	-0.1791	-0.1710
Nestlé Cheerios Multigrain 400 g.	-0.1560	-0.1661	-0.1417	-0.1575	-0.1529
Composite Labeled	-0.1626	-0.1693	-0.1550	-0.1659	-0.1583
Mean	-0.1671	-0.1742	-0.1586	-0.1703	-0.1631
Unlabeled Products					
Nestlé Chocapic 800 g.	0.0829	0.0846	0.0821	0.0860	0.0798
Quaker Oatmeal 900 g.	0.0872	0.0891	0.0827	0.0953	0.0810
Nestlé Chocapic 600 g.	0.0873	0.0891	0.0862	0.0907	0.0838
Quaker Oatmeal Multiseed 800 g.	0.0806	0.0835	0.0750	0.0890	0.0732
Composite Unlabeled	0.1169	0.1192	0.1160	0.1206	0.1131
Mean	0.091	0.0931	0.0884	0.0963	0.0862
Outside option	0.0766	0.0782	0.0758	0.0792	0.0738

Notes: The table presents predicted changes in outcomes for top breakfast cereal products under alternative warning label schemes. Column (1) shows whether products were labeled under implemented regulation. Column (2) shows the difference between the predicted market shares under the observed warning labels and a counterfactual market share with no-warning labels. Column (3) shows the difference between the predicted market shares under the observed warning labels and a counterfactual market share with stricter warning labels, such that all cereals products included in the Table were required to display warning labels. Column (4) presents the predicted change in prices that would be required in a no-warning-label scenario to match the observed market shares under the actual regulatory thresholds.

Appendix (For Online Publication)

A Alternative Specifications

Table A.1: Product-Level Estimates using Block Bootstrapping (Wild) Standard Errors

Panel A. Dep Var: Log of Quantity			
	Cereal	Chocolate	Cookies
Label	-0.062**	0.016	0.019
	(0.019)	(0.017)	(0.016)
Price	-0.616***	-0.724***	-1.314***
	(0.143)	(0.113)	(0.268)
R-squared	0.672	0.751	0.701

Panel B. Dep Var: Percentage of Transactions			
	Cereals	Chocolate	Cookies
Label	-0.0007*	0.0001	0.0003
	(0.0003)	(0.0005)	(0.0003)
Price	-0.0123**	-0.0080**	-0.0170**
	(0.004)	(0.004)	(0.006)
R-squared	0.605	0.756	0.642
Store-UPC FE	✓	✓	✓
Day-of-the-Week FE	✓	✓	✓
Week FE	✓	✓	✓
Observations	12,891	12,094	24,925
Number of Stores	6	6	6
Number of Products	195	264	376

Notes: OLS estimates of Equation (1). As the dependent variable, Panel A uses the log of purchased quantities, while Panel B uses the percentage transactions in a given store-day combination containing UPC j. All specifications consider store-UPC fixed effects, fixed effects for each day of the week, and week fixed effects. Block Bootstrapping (Wild) Standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

Table A.2: Alternative Specifications to estimate Label Effects at the Product-Level (using different Fixed Effects)

Panel A. Dep Var: Log of Quantity									
		Cereals			Chocolate			Cookies	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Label	-0.036*	-0.085*	-0.054**	0.013	0.052	0.020	0.037	0.026	0.024
	(0.018)	(0.036)	(0.016)	(0.032)	(0.039)	(0.016)	(0.023)	(0.016)	(0.018)
Price	-1.082***	-0.803**	-0.624***	-3.444**	-0.665***	-0.742***	-2.511***	-1.750***	-0.963***
	(0.161)	(0.226)	(0.143)	(0.876)	(0.138)	(0.121)	(0.447)	(0.332)	(0.228)
R-squared	0.826	0.800	0.679	0.869	0.878	0.753	0.830	0.828	0.711
Panel B. Dep Var: Percentage of Transactions									
		Cereals			Chocolate			Cookies	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Label	-0.0003	-0.0006	-0.0008*	-0.0004	0.0007	0.0001	0.0007	0.0008**	0.0004
	(0.0004)	(0.0007)	(0.0003)	(0.0009)	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0002)
Price	-0.022**	-0.020**	-0.013**	-0.069*	-0.005*	-0.009*	-0.052***	-0.020**	-0.017**
	(0.007)	(0.007)	(0.004)	(0.031)	(0.003)	(0.004)	(0.011)	(0.006)	(0.006)
R-squared	0.786	0.759	0.607	0.865	0.932	0.757	0.791	0.850	0.642
Observations	10532	10829	12891	8938	10085	12094	20725	21704	24925
Number of Products	6	6	6	6	6	6	6	6	6
Number of Stores	144	176	195	155	225	264	246	339	376

Notes: OLS estimates of Equation (1) using the product-level data. As the dependent variable, Panel A uses the log of purchased quantities, while Panel B uses the percentage transactions in a given store-day combination containing UPC *j*. Different columns use alternative set of fixed effects. Columns (1), (4) and (7) use UPC-date and store-UPC fixed effects. Columns (2), (5) and (8) use UPC-store-week and day-of-the-week fixed effects. Columns (3), (6) and (9) use store-UPC and date fixed effects. Two-way clustered-robust (at the store-UPC level) standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Effect of Product Characteristics on Timing of Label Implementation (using Fixed Effects)

	(1)	(2)	(3)	(4)
Panel A. Cereals				
Previous Sales	-2.436			-2.346
	(2.398)			(3.639)
Previous Quantity		-0.003		-0.000
		(0.003)		(0.005)
Previous Price			-7.606	-9.464
			(24.380)	(24.642)
R-squared	0.135	0.132	0.129	0.136
Number of Observations	220	220	220	220
Number of Products	80	80	80	80
Panel B. Chocolate				
Previous Sales	-0.134			0.253
	(0.858)			(1.661)
Previous Quantity		-0.000		-0.000
		(0.001)		(0.001)
Previous Price			2.904	-0.591
			(56.448)	(57.958)
R-squared	0.095	0.095	0.095	0.095
Number of Observations	246	246	246	246
Number of Products	108	108	108	108
Panel C. Cookies				
Previous Sales	0.346			1.185
	(1.232)			(1.941)
Previous Quantity		-0.000		-0.001
		(0.001)		(0.001)
Previous Price			-4.420	-5.433
			(33.527)	(33.706)
R-squared	0.137	0.137	0.137	0.138
Number of Observations	616	616	616	616
Number of Products	199	199	199	199
UPC FE	✓	✓	✓	✓
Store FE	✓	✓	✓	✓

Notes: OLS estimates of Equation (1) when using as dependent variable the number of days in advance the product was first labeled in a store. Panel A, B and C uses the cereal, chocolate and cookie category, respectively. The explanatory variables are aggregated initial sales, quantities and mean price from each UPC-store; using data from May 2015 (or May 2016 for new products), before (or during the initial month) of the warning label implementation. Sales and prices in thousands and millions of Chilean pesos, respectively. All specifications consider UPC and store fixed effects. Standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

B Across Categories Comparison

Table B.1: Comparing Label Estimates across Categories

Dep Var:	Log of Quantity (1)	Percentage of Transactions (2)
Label	-0.119***	-0.0010**
	(0.020)	(0.0003)
Label x Choco	0.168***	0.0013
	(0.037)	(0.0009)
Label x Cookies	0.141***	0.0014**
	(0.017)	(0.0004)
Price	-0.779***	-0.011**
	(0.126)	(0.003)
Observations	49,910	49,910
R-squared	0.712	0.678
Number of Stores	6	6
Number of Products	835	835

Notes: OLS estimates of Equation (1) when pooling the data across categories and interacting the label dummy with category dummies with cereal breakfast as the baseline. As the dependent variable, Column (1) uses the log of purchased quantities, while Column (2) uses the percentage transactions in a given store-day combination containing UPC j. All specifications consider store-UPC fixed effects, fixed effects for each day of the week, and week fixed effects. Two-way clustered-robust (at the store-UPC level) standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

C Additional Household-level Analysis

Table C.1 shows the warning label estimates at the household level on purchase probability. Columns (1) to (3) show the label effect for the breakfast cereals, chocolates, and cookies categories, respectively. Consistent with the product-level analysis, we find a negative impact of the warning label on cereals. The label estimate in Column (1) implies a 0.14 percentage points demand decrease due to the warning label. Thus, relative to the baseline purchase probability, there is a 7.2% reduction (from 1.94% to 1.80%). This estimate is quite close to the one found in the product level analysis. For the other two product categories, consistent with the product-level analysis, we find inconclusive evidence of the warning label effect. Table **??** repeats the analysis using a different specification. Finally, similarly to the product-level analysis, we check our identification strategy by implementing a placebo test at the household level. The estimates are presented in Table **??** and show that the label coefficients are very small and not sizably different from zero in all product categories.

Table C.1: Label Effect at the Household-Level Analysis

	(1)	(2)	(3)
Dependent Variable: Purchase indicator	Cereal	Chocolate	Cookies
Label	-0.0014**	0.0005	0.0002
	(0.0007)	(0.0006)	(0.0004)
Price	-0.043***	-0.023***	-0.069***
	(0.008)	(0.005)	(0.010)
R-squared	0.014	0.018	0.017
N. of Observations	1,330,702	4,523,844	6,600,255
Number of Households	4,801	16,117	12,396
Number of Products	225	308	423

Notes: OLS estimates of Equation (3) using the household-level data, in which the dependent variable is a purchase indicator for each UPC in a given household-store-day. Specifications consider household, UPC, store, week and day-of-the-week fixed effects. Two-way clustered-robust (at the household-UPC level) standard errors in parenthesis. P-values notation: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table C.2: Label Effect at the Household-Level Analysis(using UPC-Time Fixed Effects)

	(1)	(2)	(3)
Dependent Variable: Purchase indicator	Cereal	Chocolate	Cookies
Label	-0.0025**	0.0006	0.0001
	(0.0011)	(0.0007)	(0.0004)
Price	-0.112***	-0.060***	-0.150***
	(0.018)	(0.013)	(0.022)
R-squared	0.019	0.021	0.020
Number of Observations	1,330,702	4,523,844	6,600,255
Number of Households	4,801	16,117	12,396
Number of Products	225	308	423

Notes: Specifications consider household, store, day-of-the-week and UPC-week fixed effects. Two-way clustered-robust (at the household-UPC level) standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

Table C.3: Placebo Test at the Household-Level Analysis

	(1)	(2)	(3)
Dependent Variable: Purchase indicator	Cereal	Chocolate	Cookies
Label	-0.0003	-0.0002	-0.0002
	(0.0009)	(0.0013)	(0.0005)
Price	-0.071***	-0.046***	-0.067***
	(0.013)	(0.012)	(0.023)
R-squared	0.017	0.022	0.017
Number of Observations	1,044,256	4,080,412	6,139,354
Number of Households	4,760	16,130	12,320
Number of Products	180	269	364

Notes: OLS estimates of Equation (3) using data of year 2015 and inputting the gradual label implementation that took place in 2016. The dependent variable is a purchase indicator for each UPC in a given household-store-day. Specifications consider household, UPC, store, week and day-of-the-week fixed effects. Two-way clustered-robust (at the household-UPC level) standard errors in parenthesis. P-values notation: *** p<0.01, ** p<0.05, * p<0.1.

D Before-After Estimates

Table D.1: Product-Level Estimates assuming full label implementation at the legal deadline.

Categories	Cereal	Chocolate	Cookies
	(1)	(2)	(3)
Panel A. Dep Var: Log of Quantity			
Label	-2.59***	0.375	0.204*
	(0.556)	(0.233)	(0.095)
Price	-0.684***	-0.567***	-0.508***
	(0.095)	(0.089)	(0.098)
Panel B. Dep Var: Percentage of Transactions			
	(1)	(2)	(3)
Label	-0.057***	-0.004	0.003*
	(0.013)	(0.006)	(0.001)
Price	-0.012***	-0.006**	-0.005**
	(0.002)	(0.002)	(0.002)
Store-UPC FE	✓	✓	✓
Day-of-the-Week FE	✓	✓	✓
Week FE	✓	✓	✓
Number of Observations	8,278	8,270	17,566

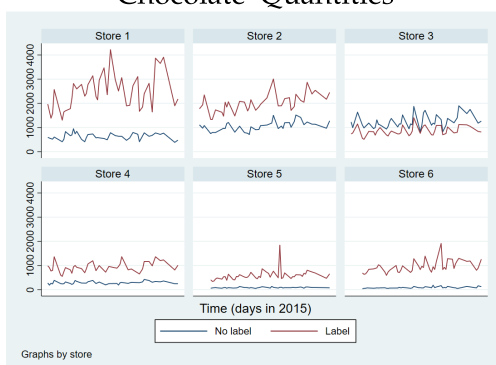
Notes: OLS estimates of Equation (1) using data one week after the law came into force in 2016 and the same period in 2015 as control (i.e., neglecting the gradual label implementation that took place in 2016). As the dependent variable, Panel A uses the log of purchased quantities, while Panel B uses the percentage transactions in a given store-day combination, respectively. All specifications consider store-UPC fixed effects, fixed effects for each day of the week, week, and year fixed effects. Two-way clustered-robust (at the store-UPC level) standard errors in parenthesis. P-values notation: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

E Parallel Trends of Treated and Untreated Products

Figure E.1: 2015 Quantities of Labeled and Unlabeled products by store

Breakfast Cereal Quantities

Chocolate Quantities



Cookie Quantities

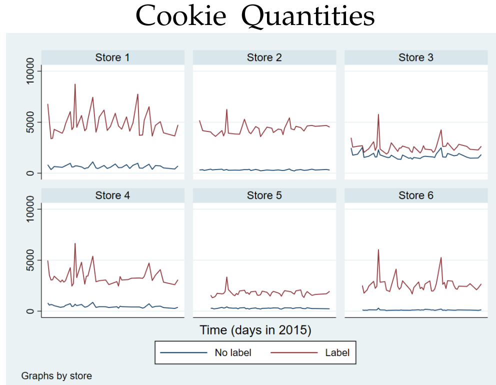


Figure E.2: 2015 Sales of Labeled and Unlabeled products by store