Strategic Adaptation to Future Demand Shifts: Evidence from Nutritional Labeling

Or Avishay-Rizi
Universitat Pompeu Fabra and BSE

Oren Reshef
Washington University in St. Louis

August 1, 2025

Abstract

How well can firms predict and respond to future shifts in consumer demand? We investigate this question by examining the effects of a nationwide labeling policy for unhealthy food products. Our analysis uses a novel dataset of hand-collected information over several years, supplemented with administrative data. We first show that sales of products with warning labels dropped sharply, with significant variation across product categories and consumer segments. To avoid the labels, firms strategically adjusted their product designs so that reformulated products fell just below the labeling threshold. Importantly, firms targeted their reformulation efforts toward products that would have been most adversely affected by the labels—both within the most impacted categories and among products sold to the most responsive consumers. Moreover, larger and more successful firms were better at anticipating these demand shifts, which translated into improved performance.

Keywords: Firm Behavior, Product Reformulation, Front-of-Package Labels

1 Introduction

Understanding the dynamics of consumer demand and adjusting offerings accordingly is crucial for a firm's long-term profitability. In many markets, demand for products changes frequently. While it is often implicitly assumed that firms can strategically respond to these changes, recent research indicates that firms tend to misprice their products in both dynamic and stable environments (DellaVigna and Gentzkow, 2019). The challenge firms face exacerbates with product design decisions, which typically involve costly, time-intensive, and irreversible investments. Moreover, firms must make these critical decisions well in advance of demand shocks, relying on their expectations of future market conditions. Firms that can better identify and adapt to demand shifts may secure a significant competitive advantage over their competitors, which may be one channel driving the persistent differences in size and profitability across firms (Syverson, 2011; Bloom et al., 2013). Nonetheless, there is limited research examining how accurately firms can predict and respond to these anticipated changes.

In this paper, we empirically explore whether firms are able to correctly anticipate and respond to shifts in consumer demand, and examine which types of firms are better at achieving this. Our empirical setting is an Israeli labeling policy aimed at addressing growing obesity rates through the promotion of healthy food choices. The policy required producers to post threshold-based front-of-package warning labels (FoPL) on all unhealthy food products.[1] The paper focuses on whether firms preemptively prepared for the new warning labels, their ability to optimize product reformulation choices ahead of implementation, and which types of firms are most effective at adjusting their offerings.

We use both administrative and detailed, hand-collected data to estimate the responses of

[1]A similar labeling policy in Chile is studied by Alé-Chilet and Moshary (2022) and Barahona et al. (2023), both of whom examine consumer responses and firms' reformulation in the breakfast cereal market. Our study builds on this literature by extending the analysis to multiple product categories and linking firms' strategic adaptation to consumer responses.

consumers and firms. Our findings confirm that consumers are affected by these labels, resulting in decreased purchases of products labeled as unhealthy. However, there is significant heterogeneity in responsiveness across different food categories and consumer segments. On the other side of the market, firms have successfully anticipated and adapted to shifting demand. Even before the new labeling came into effect, firms strategically reformulated products to narrowly meet the criteria required to avoid the warning labels, often bunching just above the minimal standards. Moreover, the heterogeneous responses from consumers created incentives for firms to prioritize the reformulation of certain products over others, specifically products that would have been most negatively impacted by labeling. These strategic responses were primarily driven by larger, more successful firms, indicating their superior ability to predict and respond to demand shifts in advance.

The empirical analyses rely on two primary data sources: The first consists of nutritional data on nearly 4,000 pre-packaged food products across 39 categories, collected both before and after the implementation of the policy. This data was hand-collected by a team of research assistants at various supermarkets over a period of more than three years. Secondly, we supplement the nutritional data with scanner sales data encompassing almost all Israeli grocery retail stores.

The nutritional labels are based on a binary threshold system, focusing on three nutrients that have been identified as unhealthy: sodium, saturated fat, and sugar. The sharp nature of this labeling system, combined with our detailed nutritional and sales data, enables us to estimate the responses on both sides of the market. Our analysis begins by focusing on consumer responses to the new labels, then examines how firms respond and make decisions regarding product reformulation. In practice, unlike price adjustments, which can be made relatively quickly (Pachali et al., 2023), product reformulation requires planning, implementation, and running through existing inventory. Thus, firms need to be forward-looking, basing their decisions on anticipated shifts in consumer demand, even before observing the demand response.

On the demand side, we find that consumers respond to the labels by reducing their purchases of labeled products. In particular, products with one or more warning labels see a 16% reduction in demand compared to non-labeled products, even when accounting for differences in underlying nutritional content. We also document substantial heterogeneity in the effect of the labels across different food categories. For example, products with high-sugar labels experience significant decreases in demand in the yogurt category but not in the chocolate category. Lastly, although we do not have data on individual-level purchases, the negative effect of labels on consumer demand appears to be driven by consumers of lower socioeconomic status, as inferred from the average demographics of the areas surrounding store locations.

Next, we estimate producers' responses to the new labeling scheme. We implement a Difference-in-Differences research design to compare the reformulation efforts of Israeli firms to those of international producers. The underlying intuition is that Israeli firms, who primarily operate in the Israeli market, are more directly and substantially affected by the policy, whereas the Israeli market constitutes only a small share of total sales for most international firms, giving them weaker incentives to respond to local regulatory changes. We find that affected firms respond to the policy by reformulating existing food products in order to avoid the warning labels. For instance, the likelihood that a food product receives the high-sugar and high-sodium warning labels decreases by approximately 2 and 5.5 percentage points, respectively. Notably, because product design changes take time to implement, many firms do not wait until the policy is enacted; instead, they begin reformulating products ahead of the implementation of new labeling policy in order to avoid the negative effects of the label.

Firms are strategic in their responses. First, their reformulation efforts are concentrated almost exclusively on products that are just above the regulatory threshold. Moreover, firms reduce the levels of targeted nutrients by the minimal amount necessary to evade the label, often resulting in products bunching just below the threshold. Secondly, firms adeptly anticipate heterogeneities in consumer demand and adjust their reformulation decisions accordingly. We find that they successfully target products in categories where labeled items would experience the most significant declines in demand. For instance, in categories where consumers largely avoid labeled products, we estimate a 5.1 percentage point increase in the likelihood of a product being reformulated to avoid the high-sugar label. Conversely, in categories where consumers appear less affected by the labels, the impact on the likelihood of reformulation is negligible. Thirdly, firms are also able to accurately target products consumed by specific consumer segments. In particular, we find that products popular in areas with label-responsive consumers are more likely to be selected for reformulation.

Having demonstrated that firms are highly strategic in responding to labeling policies, we examine whether certain types of firms are better able to accurately target their reformulation efforts. We find that larger, more successful firms, as measured by their past revenue, are more strategic in their responses. For example, successful firms are more likely to target products just above the threshold for reformulation and are more inclined to bunch at minimum standards. These firms are also highly effective at targeting specific food categories and consumer segments. For instance, considering the high-sugar label, we find that a 1% increase in firm size, measured by pre-regulation revenue, is associated with a 2.3% greater likelihood of differentially targeting affected categories compared to the average supplier. Similarly, larger firms are more effective at targeting responsive consumer segments—a 1% increase in firm size is associated with a 2.4% greater likelihood of differentially targeting affected consumer segments. Taken together, these results are consistent with a relationship between a firm performance and its ability to adapt to changing market conditions.

Literature review Our paper contributes to a growing literature on firm decision-making processes in dynamic markets. A large body of research focuses on learning by doing in production (De Loecker, 2013; Levitt et al., 2013), as well as in bidding, advertising, estimating consumer biases, coordinating prices, managerial decision-making, and assessing demand (Hitsch, 2006; Doraszelski et al., 2018; Byrne and de Roos, 2019; Yang, 2020; Tadelis et al., 2023).[2] Recent work also highlights firms' limitations in responding to demand, especially in pricing decisions (DellaVigna and Gentzkow, 2019; Strulov-Shlain, 2021; Hortaçsu et al., 2024). A smaller body of literature focuses on the demand for new products and product attributes (Ellickson et al., 2019; Cao and Zhang, 2021). In particular, Huang et al. (2022) finds that firms entering a new market are successful in learning demand and gradually set more profitable prices until converging to prices consistent with (static) profit maximization. We extend existing literature, which primarily focuses on firms iteratively learning to set prices, by studying a unique case in which firms make product design choices without immediate feedback on the outcomes of their actions.

This paper also contributes to our understanding of the large and persistent heterogeneity in firms' performance over time. A vast literature in managerial strategy focuses on the theoretical mechanisms and strategies driving competitive advantage (Porter, 1985; Barney, 1991). Our work relates most closely to the work on dynamic capabilities (Teece et al., 1997), which generally studies the ability of organizations to modify and recombine their resources. In economics, special focus has been given to heterogeneity in production (e.g, Bernard et al., 2022; Hsieh and Klenow, 2009; Syverson, 2004, 2011), and managerial capabilities and practices (Bloom and Van Reenen, 2007, 2010; Goldfarb and Xiao, 2011). We offer a new mechanism for firm success: the differential ability to predict and adjust to trends in consumer demand. Our empirical results are consistent with Hortaçsu and Puller (2008), who find that larger firms actions are closer to profit maximization than smaller firms, though the latter improve over time.

The usage of labeling to promote healthy nutrition, in addition to alternative policies such as taxation (Allcott et al., 2019; Seiler et al., 2021), has been growing in recent years. For

[2]For a comprehensive review on firms' learning we refer the reader to Aguirregabiria and Jeon (2020).

example, the introduction of non-GMO labels has resulted in a rise in consumer demand (Adalja et al., 2023; Kim et al., 2022). In Israel, the display of signs indicating the price difference between in-store prices and lower prices paid by others has led to a decrease in demand (Ater and Avishay-Rizi, 2022). In addition, Chilean Food Act, which introduced a similar set of labels, has been the subject of various studies. Araya et al. (2022) found that labels reduce demand for breakfast cereals but not for chocolates or cookies, with a greater impact on medium-low socioeconomic groups. Pachali et al. (2023) observed increased prices for labeled products due to product differentiation. Alé-Chilet and Moshary (2022) revealed that reformulation reinforces the intended effects of the policy, particularly by reducing the calorie content of cereal purchases. Barahona et al. (2023) demonstrated that consumers switch from labeled to unlabeled products, a pattern primarily driven by the misconception that certain products are healthy. Barahona et al. (2025) discuss the effectiveness of food labeling policies in a broader context.

Finally, we also contribute to the literature on labeling and the effects of new information on market outcomes. Extensive research examines how detailed product information, such as seller ratings, influences consumer demand and producer incentives, particularly in online markets (Jin and Kato, 2006; Hui et al., 2016; Luca, 2016; Tadelis, 2016). Labeling may also provide simplified or coarse information across various contexts, potentially capturing consumer attention while diverting focus from more precise but complex information (Davis and Metcalf, 2016; Houde, 2018; Rao and Ursu, forthcoming).

The remainder of the paper is organized as follows: Section 2 outlines the setting, provides additional institutional context, and describes the main datasets used in our analysis. Section 3 describes consumer responses to the labeling, including heterogeneity across different food categories and consumer demographic. Section 4 examines producers' responses, detailing the research design and the main results. Section 5 explores the implications for consumers health and the competitive environment. Finally, Section 6 concludes.

2 Setting & Data

The study focuses on a large-scale food labeling policy change in Israel. In April 2016, in order to address growing health costs and overweight rates in Israel (over 60% of adults), the Israeli Ministry of Health appointed a committee to promote healthy nutrition. The primary recommendation of the committee was to mandate manufacturers to add threshold-based front-of-package labels (FoPL) on unhealthy products, with the goal of assisting consumers in choosing healthier food products. Specifically, the labels are based on a binary threshold system targeting three nutrients that have been identified as unhealthy: sugar, sodium, and saturated fat. Every product with nutrient levels exceeding each and any of these thresholds would receive a nutrient-specific warning label; meaning that a single product can receive up to three labels.

In order to give manufacturers time to properly prepare, the policy, enacted in December 2017, rolled out in two phases (January 2020 and 2021), with thresholds becoming more stringent in the second phase. The labels and nutrient thresholds by phase are presented in Table 1. Once the policy was implemented, all pre-packaged food products were required to carry the labels in a specific location, size, and color. The regulation was inspired by policies introduced in Chile which enforced similar threshold levels, and to some extent, also the voluntary labeling scheme adopted in European countries.[3] Importantly, even prior to the new regulation, all pre-packaged food in Israel carried a detailed nutritional label on the back of the package, similar to the US. The back-of-the-package nutritional label includes ingredients, allergens, and a complete list of nutritional values.[4] Thus, the main goal of the policy was to highlight and increase the salience of certain unhealthy nutrients, rather than

[3]Unlike the Chilean policy, the Israeli regulation does not impose any restrictions on advertising labeled products. This distinction is particularly significant as it helps isolate the potential effect of labeling on demand from the effect of advertising (Dubois et al., 2018).

[4]The list of nutritional values includes sodium, saturated fat, calories, carbs, etc. Prior to the regulation, manufacturers were not obligated to disclose sugar content; however, it turns out that disclosing the sugar content was common practice.

provide new information.

Table 1: Thresholds by Nutrient

	Nutrient	2020	2021
	Sugar	13.5g	10g
	Sodium	500mg	400mg
	Saturated Fat	5g	4g

Notes: The table shows the labeling thresholds for sugar, sodium, and saturated fat per 100g in the first and second stages of the policy implementation. Labeling is mandatory for products that exceed these thresholds.

2.1 Data

The introduction of mandatory FoPL provides an institutional shock that affected both the demand and supply sides of the Israeli food market. To study firm and consumer responses to the introduction of FoPL, we use two main data sources: nutritional information and sales data. The data are described in Table 2.

Nutritional Information Tracking product nutrition information over time is essential for our research design. Unfortunately, no centralized database exists for the nutritional content of Israeli products. To address this gap, we manually gathered an extensive panel of nutritional data from product packages over three years, beginning shortly after the announcement of the labeling policy in mid-2017 and concluding in 2021. This effort was led by research assistants—primarily undergraduate students from two major Israeli universities—who visited multiple branches of the largest grocery retailer across Israel's major cities, typically every one to two months.

During each store visit, the team meticulously recorded information on all food products from the 39 distinct categories participating in the study and matching items by barcode. To ensure data accuracy and reliability, the assistants took detailed photographs of both the

Table 2: Descriptive Statistics

Characteristic	Count	
Categories	39	
Items	2328	
Producers	197	
Israeli Producers	104	
Soon to be Labeled Items	1561	
Israeli Items Reformulated	137 of 1007	
International Items Reformulated	7 of 554	
	Average	**Standard Deviation**
Producers in Category	9.1	5.2
Items in Category	59.7	38.9
Products per Producer	13.1	42.9
Producer Sales (Million Shekels)	3.73	18.27

Notes: The table presents descriptive statistics for items observed in 2019, and reformulations refer to the reduction of at least one critical nutrient to fall below the thresholds in 2020.

back and front packaging of each product. These photographs were subsequently digitized by another team, who manually input the data into our system. The integrity of the digitized data was periodically verified through random checks.

Following digitization of the data, we implemented several processes to standardize and clean the data. We restricted the sample to products that were observed at least three times to ensure robustness. A significant challenge was dealing with missing data, often due to inventory constraints in the store or errors by assistants, such as unclear photographs. However, we note that nutritional content generally remained stable over time. For example, for each nutrient separately, the median product in our dataset showed no nutritional changes throughout the study period, with the maximum change recorded as 4. Thus, to interpolate missing information, if a product was observed on two different occasions with the same nutritional value, we assumed continuity of that value for all intervening periods. When changes in nutritional content were detected, we coded the date of change as the midpoint between the two observation dates.[5] A comprehensive account of our data cleaning process is provided in Appendix A. Our final sample of nutritional information consists of 3,844 unique products and 38,724 product-visit observations.

Sales data The second data source is monthly scanner data for the years 2019 and 2021, obtained from Storenext.co.il.[6] The data consist of price and sales data from 150 stores of the 12 largest Israeli retailers. Although the identity of each individual store is obfuscated to protect sensitive business information, we observe measures of consumer demographics in the area for each store. We aggregate the data to the product-chain-city-month level, i.e., for each product we observe retailer-level monthly sales in each of Israel's 10 largest cities. After merging the sales data with our manually-collected nutritional information, the final sample consists of an unbalanced panel of 3,551 products and 6,211,612 product-chain-city-month observations.

3 Analysis of Consumer Response

We begin by briefly examining the effect of the label on consumer purchase behavior, following the analysis presented in Barahona et al. (2023). We use the panel structure of our data to estimate the effect on product-level monthly purchases by comparing labeled and unlabeled products before and after the reform. To account for unobserved differences in preferences for different products and nutritional compositions, we flexibly control for nutrient content as well as prices. We also include a myriad of fixed effects to isolate the effects

[5] For products for which sugar appears only after the reform, we impute the amount of sugar as equal to the post amount, only if the values of all other nutrients remain unchanged. We deem this imputation innocuous as sugar and carbs levels are closely related, and thus, any change in sugar is bound to affect the amount of carbs in the product.

[6] Storenext.co.il collects scanner data from thousands of stores throughout Israel, and is the main source of scanner data for academia and industry, similar to Nielsen scanner data in the U.S.

of the labeling while holding fixed other factors. Formally, we estimate:[7]

$$\log\left(q_{irct}\right) = \beta \mathrm{Post}_t \times \mathrm{FoPL}_{int} + \theta \mathrm{FoPL}_{int} + \tag{1}$$
$$\eta \log(\hat{p}_{its}) + f(nutrient_{int}) + \mu_i + \alpha_{rc} + \gamma_t + \epsilon_{irct},$$
$$\text{with} \quad f(nutrient_{int}) = \psi_1 Sodium_{int} + \psi_2 Sugar_{int} + \psi_3 SatFat_{int}$$

where i, r, c, t, n are indices for item, retailer, city, time, and nutrient, respectively; q is quantity sold; $Post$ is an indicator for post period; and $FoPL_{int}$ indicates whether, at time t, nutrient n of product i exceeds the relevant regulatory threshold. To separate the effect of the labels from other, secular trends in consumer preferences for health, our main specification linearly controls for preferences for nutritional values, though the main results remain essentially unchanged by adding higher-order polynomials or allowing for time changing preferences (Appendix Table A1). We also include product, city-retailer, and time fixed effect. Finally, since prices, denoted by p_{its}, are potentially endogenous, we instrument for price using prices in other stores of the same retailer (Nevo, 2001; DellaVigna and Gentzkow, 2019). Standard errors are clustered at the product-level, corresponding to the level at which treatment is assigned (Abadie et al., 2023). Finally, following Barahona et al. (2023) observations are weighted by product-store pre-policy revenue, which ensures that we capture the average policy effect across the population, weighted towards the effects on larger, product-store pairs. in our main specification, we exclude the second half of each year because many manufacturers began offering new and labeled products several months before the policy was formally introduced. However, the results are robust to the inclusion of the full year (Appendix Table A2).

[7]This specification estimates the labels effect on consumer behavior, potentially simultaneously reducing demand for labeled products and increasing demand for unlabeled products. Hence, the coefficient of interest, θ, should be interpreted as capturing the *difference* between labeled and unlabeled products, as opposed to the counterfactual where labeling does not exist.

The results are presented in Table 3. Across specifications, we find economically meaningful and statistically significant negative effects of the warning labels on product sales. These effects are especially pronounced for sugar and sodium. Our findings echo the results in Alé-Chilet and Moshary (2022) and Barahona et al. (2023), who also find significant reductions in the demand for labeled products.

Column (1) presents the estimated effect of receiving at least one label on product sales in 2020. Products that received one or more labels experience, on average, a decrease of 16% in sales, compared to products that did not receive any warning labels. As presented in Column (2), receiving one additional warning label (with a maximum of three labels), reduces sales by approximately 10%. Column (3) breaks down the effect of labels on sales by nutrient. High-sugar label products experienced the largest decrease in sales, 25.5%, and high-sodium products experienced a decrease of 11%. The coefficient on saturated fat is smaller in magnitude, positive, and statistically insignificant.[8]

The results remain qualitatively unchanged and statistically significant across alternative specification, which include data from the full year, restricting attention to products just-above and below the label or unchanged products, etc. These robustness tests are detailed in Appendix B.

4 Analysis of Producer Response

4.1 Descriptive Evidence

We begin to examine firms response to the introduction of FoPL by evaluating the changes in the distribution of nutritional values following the new labelling scheme. Firms could have

[8]The weaker effects of the saturated fat label are rationalized by several interviews we conduct with Israeli nutritionists: On the demand-side, consumers are generally less familiar with this nutrient and find it difficult to understand how it affects product taste or future health outcomes.

Table 3: The Effect of FoPL on Quantities

Dependent Variable:	log(q)		
	(1)	(2)	(3)
Post × $\mathbb{1}$(has labels)	-0.161		
	(0.036)		
	[.001]		
Post × (# of labels)		-0.104	
		(0.022)	
		[.001]	
Post × $\mathbb{1}$(Sugar > 13.5)			-0.255
			(0.062)
			[.001]
Post × $\mathbb{1}$(Sodium > 500)			-0.113
			(0.051)
			[.027]
Post × $\mathbb{1}$(Sat. Fat > 5)			0.022
			(0.051)
			[.673]
log(p)	-2.099	-2.102	-2.106
	(0.094)	(0.095)	(0.093)
	[.001]	[.001]	[.001]
Observations	1,402,525	1,402,525	1,402,525

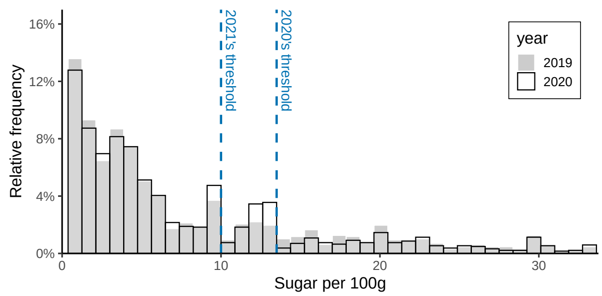
Notes: The table displays the estimation results for Equation 1. Column (1) presents the average effect of labels on quantity, comparing products with and without labels. Column (2), reports the linear relationship between the number of labels and their impact on demand. Column (3), shows the effect of each label on demand. The observations are weighted by pre-regulation product × store revenue. Additional covariates include thresholds dummies, nutrient content, product, city × retailer, and month-fixed effects. Standard errors are clustered at the product level and reported in parentheses. Exact p value reported in the square brackets

prepared for the introduction of the labels in several ways, such as product reformulation, retiring existing products, or introducing new products.[9] Figure 1 presents the raw changes in the distribution of nutritional values between 2019 (in gray) to 2020 (in white).[10] The distribution, even pre-regulation, is not smooth, as nutritional values tend to bunch at round numbers such as 500, 750, and 1000 for sodium; and 10, 20, and 30 for sugar. That said, following the introduction of the new FoPL in 2020, there appears to be significant excess

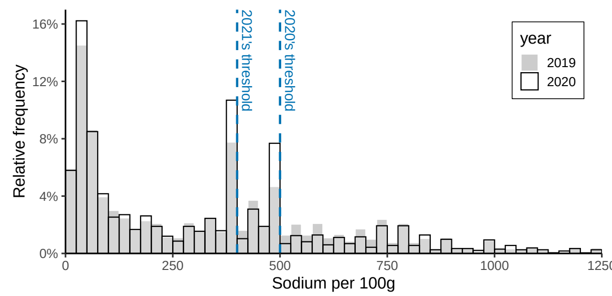
[9]Another potential response is readjusting prices. However, as we can see in Appendix Table A3, prices were affected to a lesser extent.

[10]Appendix Figure A1 presents similar results examining product changes from 2019 to 2021.

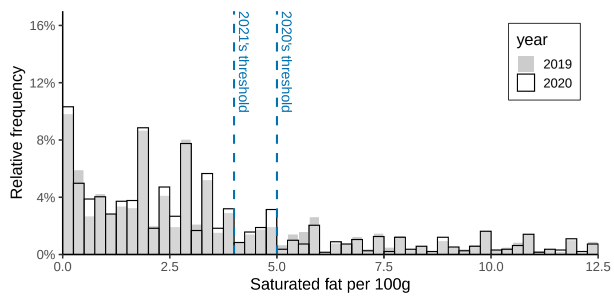
bunching just-below the cutoffs for both labels: 400 and 500 for sodium, and 10 and 13.5 for sugar. Notably there appears to be little to no increases in the density of any values below the threshold except those just-below the threshold. In the same vain, almost all of the soon-to-be-labeled 2019 products that are "missing" in 2020 (either through reformulation or exit), are initially located just-above the threshold. Similar to the demand estimation results, the effects are weaker for saturated fat.



Panel A: Sugar



Panel B: Sodium



Panel C: Saturated Fat

Figure 1: Distribution of Israeli Products by Nutritional Values

Notes: The figure illustrates the distribution of Israeli products based on their nutrient values. The vertical dashed lines represent the thresholds that were implemented at the start of 2020 and 2021. The distribution of products in 2019, indicated by the gray color, is contrasted with the overlayed distribution in 2020, depicted in white.

This set of results suggests that firms respond strategically to the policy: Rather than reducing unhealthy nutrients across the board, firms respond "locally," taking advantage of the discontinuous nature of the policy to avoid the labels while only making minimal changes to the product. There are several potential explanations as to why only marginal products are affected: One possibility is that making large changes to a product while maintaining its taste or consistency, is prohibitively costly.

This pattern in firm responses also suggests the reductions are not driven by general changes in consumer preferences towards healthier nutrition. If observed changes in product offerings were the results of secular trends in demand for nutrients or changes in the cost of providing healthier products, then we would expect to see changes across the entire distribution. However, as presented in Figure 1, nearly all changes in product offerings are concentrated in a narrow band around the cutoff. The proportion of reformulated products drops sharply as the original nutritional values grow further away from the threshold, becoming null for products with nutritional values well above the threshold. These finding are consistent with producers strategically avoiding the new warning labels.

4.2 Research Design

Next, we formally estimate firms' responses to the introduction of FoPL using a Difference-in-Differences research design. We define the treated group to be Israeli producers, who primarily operate locally and thus have strong incentives to respond to the new labeling. In contrast, international producers are unlikely to change their products in response to the Israeli regulation, given that Israel, with a population of fewer than 10 million, constitutes a small market relative to the rest of the world.

To illustrate, consider one product: ketchup, sold in Israel by the large local producer Osem and the globally recognized brand Heinz. While both brands are popular in Israel, the Israeli market represents 85% of Osem's global ketchup sales, compared to less than 0.5% for Heinz. These stark differences suggest that the two producers faced very different incentives to respond to the new labeling policy. Indeed, Osem introduced a low-sodium version before the policy's implementation, whereas Heinz stated it would continue to follow its international standard and would not modify its original recipe.[11] To assess the validity of using

[11] We calculated shares based on information from https://en.globes.co.il/en/

international producers as the control group, we replicate the analysis shown in Figure 1, but this time focus solely on international products. Figure A2 presents changes in the distributions of nutritional values from 2019 to 2020 for international products. Unlike the Israeli products, there is no evidence of strategic reformulation or bunching among international products, confirming that they serves as a valid control group.

The main Difference-in-Differences specification is thus given by:

$$y_{it} = \beta \text{Post}_t \times \text{IL}_i + \mu_i + \gamma_t + \epsilon_{it}, \tag{2}$$

where i and t are indices for item and time, y_{it} is an indicator for whether a product is reformulated below a threshold, for a given warning label,[12] IL_i is an indicator for an Israeli producer, and $Post_t$ is an indicator for 2020. Finally, μ_i and γ_t represent product and time fixed effects, respectively. In all regressions, we cluster the standard errors at the producer-category level to allow producers' decisions to be correlated within a product category.

Because product design decisions unfold over time and new product batches do not appear in stores precisely when the policy takes effect, it is difficult to assess the parallel trends assumption using pre-policy trends. Nevertheless, as we show below, the discontinuous nature of the policy and the fact that producer responses are concentrated almost exclusively around labeling thresholds suggest that the observed effects are driven by the labeling scheme itself rather than by other unobserved factors.

Finally, product fixed effects restrict the analysis to products that were present both before and after the regulation was implemented. Consequently, this specification captures on

article-osem-sells-less-than-strauss-but-earns-more-1000968195, https://www.ynet.co.il/economy/article/byncm8zto, and https://www.kraftheinzcompany.com/company.html.

[12] Our main specification focuses on the introduction of the 2020 policy, with results for 2021 presented in the appendix. However, many firms began reformulating products in 2020 to avoid the stricter 2021 threshold. Accordingly, we code the outcome variable as 1 once a product is reformulated—either by moving from above to below the 2020 threshold, or from above to below the 2021 threshold. Importantly, all of our main results are robust to excluding changes related to the 2021 threshold, including specifications that entirely remove those observations.

Table 4: The Effect of FoPL on Reformulations

Dependent Variable:	Below the thresholds		
Nutrient:	Sugar	Sodium	Sat. Fat
	(1)	(2)	(3)
Post × IL	0.020	0.055	0.015
	(0.009)	(0.014)	(0.008)
	[.036]	[.001]	[.056]
Observations	4,396	5,196	5,192

Notes: The table presents the likelihood of reformulations that firms undertake to avoid warning labels. Formally, the table summarizes the estimation results of Equation 2 using all the products available in both 2019 and 2020.For instance, Column (1) shows that the proportion of unlabeled Israeli products increased by 2.6 percentage points due to the policy. All regressions include product and year fixed effects. Standard errors are clustered at the category × producer level and reported in parentheses. Exact p value reported in the square brackets

reformulation rather than entry or exit. We separately study entry and exit decisions below — in general, we find some evidence of strategic entry, but no clear evidence on product exit.

4.3 Results

Table 4 presents the estimates of suppliers' responses, as described in Equation 2, essentially estimating the likelihood of the reformulation below the threshold for existing products. Across nutrients, FoPL have a significant effect on the probability that a product will be reformulated to levels below the warning label. For example, through reformulation, the proportion of Israeli products below the sodium (sugar) threshold increases by 5.5 (2) percentage points compared to the change observed for international products.

We again observe that the results for saturated fat are weaker, which has been a consistent trend throughout our analyses. From our discussions with several Israeli nutrition designers, we understand that saturated fat is generally challenging to manipulate. Unlike sugar and sodium, which are typically added to the products, saturated fat is inherently produced during the manufacturing process. For this reason, we exclude saturated fat from our analysis of heterogeneous firm responses. However, for completeness, we present these analyses in Appendix B.

These results are robust to multiple alternative specifications, including focusing on the transition to the more stringent threshold from 2020 to 2021, focusing solely on whether a product is reformulated below the 2020 threshold (Appendix Table A4), and setting placebo thresholds (Appendix Table A5).

To assess the validity of our research design, we examine whether the observed differences between Israeli and international producers are driven by efforts to avoid the warning label, rather than to other potentially unobserved factors. The results are presented in Appendix Table A6. Panel A presents a placebo analysis in which we restrict the sample to products that were reformulated but did not cross the labeling threshold—-that is, reformulations that did not result in the removal of the warning label. Panel B reports the probability that a reformulated product ends up below the threshold, conditional on having been reformulated.

As Panel A shows, in stark contrast to Table 4, Israeli producers do not exhibit increased reformulation activity when those changes do not lead to the avoidance of a warning label. In fact, several estimates are negative. This pattern is consistent across all three nutrients. Finally, Panel B demonstrates that, conditional on reformulation, Israeli producers are substantially more likely than international producers to reduce nutrient levels below the threshold. Taken together, these findings indicate that the observed behavioral differences are driven by strategic efforts to avoid the warning label, rather than by broader, policy-unrelated reformulation trends.

Another potential strategy for firms to respond to the FoPL policy is by discontinuing existing products above the threshold or by launching new products below it. To examine the effect on the extensive margin, we conduct an additional analysis of product entry and exit. More specifically, we estimate a model to derive likelihood that a product is introduced or

Table 5: The Effect of FoPL on Entry & Exit

Dependent Variables:	Exit			Enter		
Nutrient:	Sugar	Sodium	Sat. Fat	Sugar	Sodium	Sat. Fat
	(1)	(2)	(3)	(4)	(5)	(6)
Above 2020 thresholds	0.017	0.016	0.051	-0.110	-0.123	-0.029
	(0.042)	(0.029)	(0.028)	(0.052)	(0.045)	(0.023)
	[.683]	[.582]	[.071]	[.036]	[.006]	[.212]
Sample:	2019	2019	2019	2020	2020	2020
Observations	2,360	2,818	2,816	2,685	3,089	3,087

Notes: This table presents estimates of product entry and exit. Columns (1)–(3) use all products sold in 2019 and measure the likelihood that products above the thresholds exited the market (i.e., were not sold in 2020). Columns (4)–(6) use all products sold in 2020 and measure the likelihood that products above the thresholds were new products (i.e., were not sold in 2019).All specifications include category × producer fixed effects, and above threshold indicator. Standard errors are clustered at the category × producer level and reported in parentheses. Exact p value reported in the square brackets

discontinued above or below the labeling threshold.

The results are presented in Table 5. The effects on product entry are presented in Columns (1) through (3). We find that new products are significantly less likely to have nutritional values above the label threshold. For example, there is an 11% decrease in the probability that a new product will exceed the the 2020 sugar threshold. The effects on product exit, shown in Columns (4) through (6), are significantly weaker. Although there is a modest increase in the likelihood of a labeled product exiting the market, these effects are smaller in magnitude and are not statistically significant. A possible explanation for this could be that producers are more reluctant to discontinue successful products before they can assess the impact of the label on consumer demand.

4.3.1 Are Producers Responding Strategically?

In this section, we examine whether producers strategically focus their reformulation efforts by focusing on particular sets of products. We first examine whether firms target products that are close to the nutritional label threshold and reduce nutrient levels by the minimum amount required to avoid the label. Next, we test whether producers base their efforts on specific food categories and consumer segments — primarily those most responsive to the new warning labels.

Strategic manipulation of nutritional values We begin by examining whether firms strategically reformulate products to avoid the new label. The analysis is presented in Figure 3. Panel A presents the likelihood of crossing the threshold, allowing the coefficients to vary non-monotonically by initial nutrient level bins in 2019. We find that products initially just above the threshold (in the fourth bin) demonstrate the most notable and statistically significant effects for both nutrients. Weaker effects are found for products in the third bin, which includes products just above the 2021 threshold. Panel B estimates the effect, allowing for heterogeneity by the nutritional levels where reformulated products eventually "end up." The most substantial effects, both in magnitude and statistical significance, are for products that end up in the third bin, just below the 2020 threshold. Again, we find smaller effects for the second bin, which captures products bunching just below the 2021 threshold.

This set of results echos the descriptive findings in Figure 1 and formalizes the intuition that firms respond strategically to the labeling by reformulating product just above the threshold and move towards bunching just below the threshold. Rather than reducing unhealthy nutrients across the board, firms respond "locally," taking advantage of the discontinuous nature of the policy to avoid the labels while making minimal changes to the product. One potential explanation is that making large changes to the product while maintaining its taste or consistency, is prohibitively costly.

Heterogeneity across categories In this section, we examine whether consumer demand is more responsive in certain categories. To this end, we estimate a version of Equation 1 that allows the effect of the label to vary by food category. The distributions of estimators by category are displayed in Figure 4, essentially capturing the average effect of the label

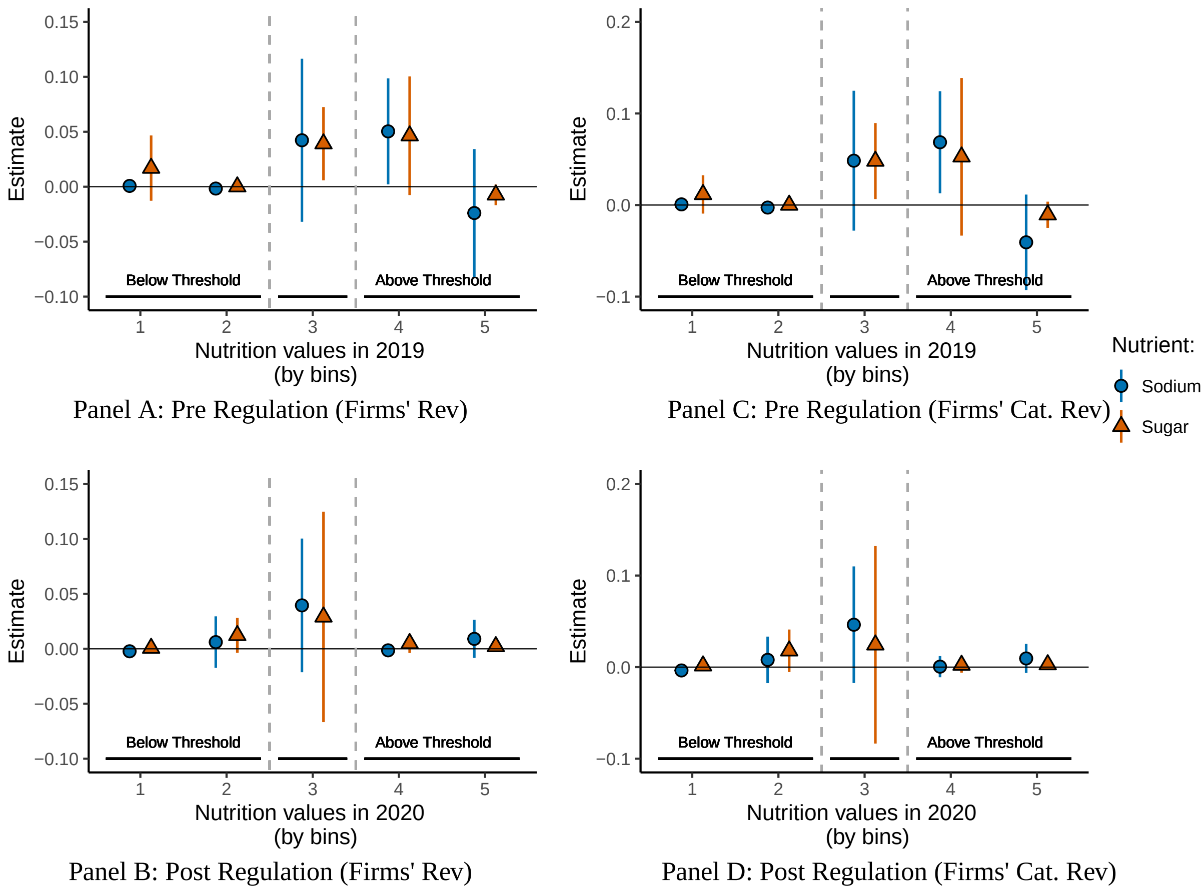
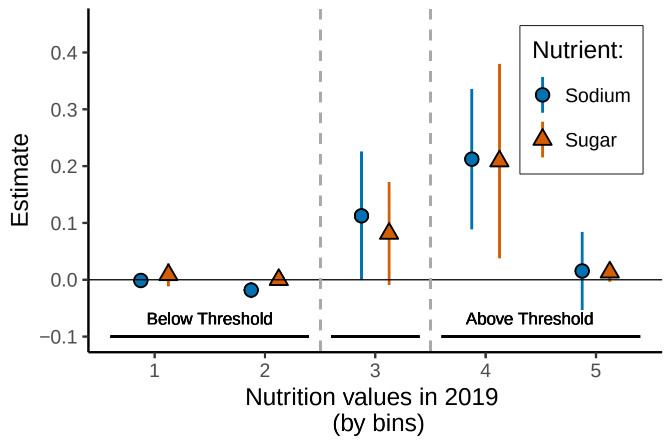
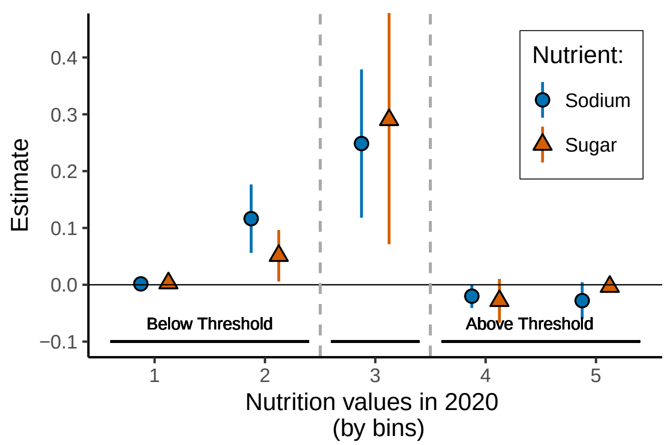


Figure 2: Likelihood of Crossing the Thresholds and firms' size, Based on Pre/Post Bins

Notes: The figure illustrates the relationship between reformulation and product nutrient levels before the regulation (Panel A) and after reformulation (Panel B). Each point in the figure corresponds to a specific coefficient of the pattern $\text{Post}_t \times \text{IL}_i \times \text{Bin}_i \times \log(\text{Size}_i)$. The vertical lines represent the 95% confidence intervals, which were calculated based on the standard error that we cluster on the $\textit{producer} \times \textit{category}$ level. Remarkably, the majority of products that cross the thresholds exhibit a distinct trend: they transition from being marginally above the thresholds in Panel A to just below the thresholds in Panel B. The bins are constructed according to the following criteria: Bin 1 represents 0-50% of the 2021 thresholds, Bin 2 represents 50-100% of the 2021 thresholds, Bin 3 represents the range between the 2021 thresholds and the 2020 thresholds, Bin 4 represents the range between the 2020 thresholds and 200% of the 2021 thresholds, and Bin 5 includes values that exceed 200% of the 2021 thresholds.



Panel A: Pre Regulation



Panel B: Post Regulation

Figure 3: Likelihood of Crossing the Thresholds, Based on Pre/Post Bins

Notes: The figure illustrates the relationship between reformulation and product nutrient levels before the regulation (Panel A) and after reformulation (Panel B). Each point in the figure corresponds to a specific coefficient of the pattern $\text{Post}_t \times \text{IL}_i \times \text{Bin}_i$. The vertical lines represent the 95% confidence intervals, which were calculated based on the standard error that we cluster on the *producer* × *category* level. Remarkably, the majority of products that cross the thresholds exhibit a distinct trend: they transition from being marginally above the thresholds in Panel A to just below the thresholds in Panel B. The bins are constructed according to the following criteria: Bin 1 represents 0-50% of the 2021 thresholds, Bin 2 represents 50-100% of the 2021 thresholds, Bin 3 represents the range between the 2021 thresholds and the 2020 thresholds, Bin 4 represents the range between the 2020 thresholds and 200% of the 2021 thresholds, and Bin 5 includes values that exceed 200% of the 2021 thresholds.

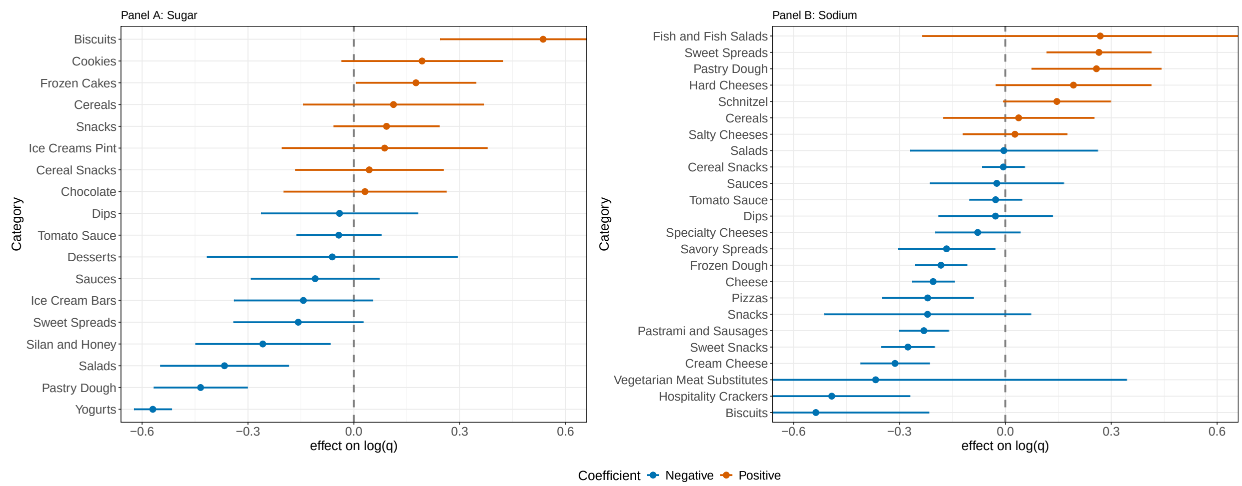


Figure 4: Heterogeneity in the Within Category Effect of FoPL on Quantities

Notes: The figure shows the within-category effect of front-of-pack labeling (FoPL) on consumer demand for different food categories. The y-axis shows the food category, and the x-axis shows the estimated effect of the label on log(q). The lines represent the 95% confidence interval using standard errors that are clustered on the product level. Panel A focuses on the effect of high-sugar labels, panel B on the effect of high-sodium labels. To estimate the within-category effect, we require both labeled and unlabeled products within each category. If a category lacks either labeled or unlabeled products, we cannot estimate the effect, resulting in variations in the number of coefficients across regressions.

separately for each food category. Generally, there is significant heterogeneity in the treatment effect on consumer demand across categories. For instance, in Panel A, which depicts the effect of the high-sugar label, this label shows no effect on the demand for products like cookies, ice cream, and chocolate. In contrast, we find a strong negative effect on the sales of products such as salads and yogurts.[13]

While identifying the mechanisms behind this heterogeneity is beyond the scope of this paper, the observed patterns anecdotally align with the model proposed by Barahona et al. (2023), where consumers respond more negatively when surprised to learn that a product is unhealthy. The finding also resembles the heterogeneity presented by Araya et al. (2022), where the labels affect demand in some categories but not in others.

Given these stark differences in consumer response across categories, we next estimate

[13]We note that for some categories the average effect is positive. However, this occurs only 3 out of 42 coefficients positive and statistically different than zero. Thus, we believe these effects are likely to be type I statistical errors rather than true effects.

whether firms are able to target the 'right' categories, i.e., categories in which consumers responded more negatively to the label. To this end, we estimate a version of Equation 2 that allows the reformulation decisions to vary by food category. In particular, we estimate:

$$
\begin{aligned}
y_{it} = \beta \text{Post}_t \times \text{IL}_i \times \mathbb{1}(\text{Aff.}_{\cdot i}) + & \\
\alpha_1 \text{Post}_t \times \text{IL}_i + \alpha_2 \text{Post}_t \times \mathbb{1}(\text{Aff.}_{\cdot i}) + \mu_i + \gamma_t + \epsilon_{it} &
\end{aligned}
\tag{3}
$$

Where all variables are as defined in Equation 2, and $\mathbb{1}(\text{Aff.}_{\cdot i})$ is an indicator variable for an affected category, defined as categories for which we estimate a negative effect of the label on consumer demand.[14] The main coefficient of interest is β, which captures differential likelihood of reformulation decisions to be concentrated in categories more affected by the label.

The results are presented in Panel A of Table 6, which shows the triple-differences specification. Across specifications, we find that firms effectively target products that are most negatively affected by the label. We generally estimate positive and (mostly) statistically significant increase in the likelihood of reformulating products in categories most affected by the label. In contrast, changes in the likelihood of reformulating products in unaffected categories are weak and not statistically significant. For example, when examining the likelihood of reformulations made to avoid receiving a high-sodium label in Column (2), we first see that Israeli firms barely adjust products in unaffected categories — the effect is weak and statistically insignificant. In contrast, Israeli firms are 7 percentage points more likely to target products in affected categories, and this effect is statistically significant. This set of results suggests that firms can correctly predict and target products in more affected categories.

[14] We experiment with alternative definitions of this variable, such as above- and below-median and the continuous point-estimate. The results, presented in Panel A of Appendix Table A7, remain qualitatively unchanged.

Table 6: Heterogeneity in the Effect of FoPL on Reformulations by Product Characteristics

Dependent Variable:	Below the thresholds	
Nutrient	Sugar	Sodium
	(1)	(2)
Panel A: Affected Categories		
Post × IL × Aff. Cat	0.051	0.070
	(0.027)	(0.028)
	[.059]	[.013]
Post × IL	0.005	0.024
	(0.003)	(0.015)
	[.097]	[.103]
Observations	4,396	5,196
Panel B: Affected Consumers		
Post × IL × Aff.Con	0.050	0.050
	(0.019)	(0.028)
	[.008]	[.072]
Post × IL	0.000	0.015
	(0.000)	(0.015)
	[.726]	[.318]
Observations	486	486

Notes: This table displays results from a triple-difference estimation, highlighting how firms target products or category reformulations in response to anticipated shifts in demand. Panel A examines reformulations for products within categories that experienced a decrease in demand ('affected categories'). Panel B examines reformulations of products popular among consumers who responded more to the regulation ('affected consumers'). Both panels show the heterogeneity in the effect of the regulation on reformulations of sugar, sodium, and saturated fat content, with the data covering the period from 2019 to 2020. In both panels, the coefficient Post × IL quantifies the regulatory impact on reformulations in 'unaffected' categories or items. Conversely, the coefficients Post × IL × Aff.Cat in Panel A and Post × IL × Aff.Con in Panel B evaluate the regulation's effect on 'affected' categories or consumers, respectively. In Appendix Table A9, we extend this analysis to other periods. All regressions include product and year fixed effects, as well as all lower-order interactions. In Panel A we cluster the standard errors at the category × producer level, and in Panel B we cluster the standard errors at the product level. Exact p value reported in the square brackets

One potential concern with our analysis is that demand and supply responses are determined simultaneously. For example, firms can self-select out of treatment by reformulating products to fall below the labeling threshold. In such cases, we only observe labels for firms that chose to retain their original formulations. As a result, our estimates may reflect a lower bound of the true demand response, which in turn could bias our classification of ”affected categories.”

To address this issue, we implement an instrumental variable (IV) strategy to instrument the labels in the post period using the pre-period nutritional values. We follow Wooldridge (2001, p. 621) and Angrist and Pischke (2009, p. 142), and estimate the probability that a product receives a label using a probit model, with instruments based on pre-regulation nutritional values and firm origin (Israeli vs. international). This approach mitigates simultaneity bias, as the instruments are predetermined and derived from pre-regulation data. For each nutrient, we estimate the following equation separately:

$$\begin{aligned}
&\Pr(\mathrm{N}_{n,i,post} \geq \mathrm{T}_n) = \beta_1 \mathbb{1}(\mathrm{N}_{n,i,pre} > \mathrm{T}_n) + \beta_2 \mathrm{N}_{n,i,pre} + \beta_3 \mathrm{N}_{n,i,pre} \mathbb{1}(\mathrm{N}_{n,i,pre} > \mathrm{T}_n) + \\
&\beta_4 1\mathbb{1}(\mathrm{N}_{n,i,pre} > \mathrm{T}_n) \times \mathbb{1}(\mathrm{IL}_i) + \beta_5 \mathrm{N}_{n,i,pre} \times \mathbb{1}(\mathrm{IL}_i) + \beta_6 \mathrm{N}_{n,i,pre} \times \beta_1 \mathbb{1}(\mathrm{N}_{n,i,pre} > \mathrm{T}_n) \\
&\times \mathbb{1}(\mathrm{IL}_i) + \delta_i + \alpha_{rc} + \gamma_t + \epsilon_{intcr},
\end{aligned} \tag{4}$$

$\mathrm{N}_{n,i,post}$ denotes the nutritional value of nutrient n of product i in period *post*, and T_n represents the threshold of nutrient n in 2020, we also include Producer$\times$ Category fixed effects δ_i,City$\times$ Retailer fixed effects αcr and time fixed effect γ_t . We note that Equation 4 does not include the full set of controls used in our demand specifications. Specifically, we cannot include product fixed effects in the first stage as they are perfectly collinear with pre-regulation nutritional values $\mathrm{N}_{n,i,pre}$.

We then use the fitted probabilities from this first-stage model as instruments for the label indicator in our demand regressions and re-estimate the heterogeneity analysis. Of the 66 category-nutrient pairs initially classified as affected, only 5 change sign under the IV correction. Using this IV-adjusted set of affected categories, we re-estimate all supply-side analyses. The results presented in Appendix Table A8 remain qualitatively unchanged, suggesting that simultaneity bias does not drive our main findings.

Heterogeneity by consumer demographics We next explore whether various consumer groups tend to differentially respond to the new label. Although we do not observe individual-level sales data, we are able to leverage demographic information at the store level to proxy for the type of consumer in each store. In particular, we observe the median salary, along with the proportions of (a) individuals with an academic background, (b) children in the population, and (c) Jewish-Sephardic households.[15] [16]

We first estimate whether these attributes moderate the effect of the label on consumer demand by estimating a version of Equation 1 allowing allowing for heterogeneous treatment effects by consumer demographics. The results are presented in Table 7. As shown in Column (1), the labeling has a stronger negative effect in neighborhoods characterized by less educated, lower-income, Jewish-Sephardic individuals with many children. These demographic attributes are generally correlated within the Israeli population, and hence it is difficult to interpret any of these variables as causal. Following the approach of Anderson (2008), we create a continuous index to capture the combined effect of all four measures. When constructing the index, we invert the signs of variables that have negative coefficients. Therefore, a high index value indicates households that are more educated, higher-income, primarily Ashkenazi Jewish households with fewer children. As is apparent in Column (2), high index values are associated with weaker effects of the new labeling, indicating a correlation between socioeconomic status and response to labeling.

[15] In order to protect the stores' identity, we do not observe continuous measure for these variables, but only their quartiles.

[16] The Jewish population in Israel can be roughly divided into Ashkenazi Jews, who migrated from Europe, and Sephardic Jews, who migrated from Asia and Africa. Ashkenazi Jews are the dominant socio-economic group in Israel. On average, Ashkenazi Jews have experienced greater socio-economic success in education, earnings, and criminal behavior. The gap had become as large as the Black-White gaps in the United States (Rubinstein and Brenner, 2014).

Table 7: Heterogeneity in the Effect of FoPL on Quantities by Demographics

Dependent Variable:	log(q)	
	(1)	(2)
$\mathbb{1}$(has labels) × Post	-0.247	-0.289
	(0.089)	(0.010)
	[.006]	[.001]
$\mathbb{1}$(has labels) × Post × Academic	0.021	
	(0.011)	
	[.058]	
$\mathbb{1}$(has labels) × Post × Wage Bin	0.010	
	(0.012)	
	[.414]	
$\mathbb{1}$(has labels) × Post × Children	-0.040	
	(0.015)	
	[.01]	
$\mathbb{1}$(has labels) × Post × Sephardic	-0.016	
	(0.012)	
	[.182]	
$\mathbb{1}$(has labels) × Post × Index		0.082
		(0.014)
		[.001]
Observations	344,389	344,389

Notes: The table displays the estimation results for Equation 1, where we interact the coefficient of interest with the demographic characteristics. Column (1) presents the association between the demographic characteristics and the effect of the labels. Column (2) shows how the effect of the labels varies with the store index. In both regressions, we aggregate the data to the year level. Both specifications include product, time, and store fixed effects. Standard errors are clustered at the store level and reported in parentheses. Exact p value reported in the square brackets

To estimate whether producers are able to target the 'right' consumers, we proceed in two steps: First, we relate products to consumer segments by estimating product popularity across stores with different consumer characteristics. In particular, we use only pre-policy data to estimate the following:

$$\log(q_{is}) = \alpha_s + \beta_i + \gamma_i \text{index}_s + \epsilon_{is}, \tag{5}$$

where $index_s$ is the store-level consumer index described above. γ_i thus captures changes in demand for the product with respect to the index, i.e., how does the *same product* differentially sell across stores serving different demographics.

Second, we test whether firms can differentially target products with lower γ_i, which are most popular with low-index, more responsive consumers.[17] We define products as "affected" if their γ_i is below -0.5, and as "unaffected" if their γ_i is above 0.5. Products with $-0.5 < \gamma_i < 0.5$ are omitted from the analysis.[18] We then re-estimate Equation 3 replacing $\mathbb{1}(\text{Aff.}_i)$ with an indicator for affected consumers, rather than category.

The results are presented in Panel B of Table 6. Across specifications, we find that firms are able to focus their reformulation efforts on product sold to consumer segments more susceptible to the new labeling. Similar to the heterogeneity by product categories, we observe weak and statistically insignificant effects on the likelihood of reformulating products popular with less responsive consumers. In contrast, we find positive and statistically sig-

[17]To develop an intuition of the relationship between products and γ, consider the following anecdotal evidence concerning a few popular items: Cream cheese, which is widely sold and highly popular in Israel, has a gamma value of 0.1, indicating that the sales of this item only slightly increase with the index. In contrast, for dairy products derived from goats or sheep, more than a third of them exhibit gamma values surpassing 0.5, while none fall below -0.5. This observation suggests that these items are popular at higher-index stores. In contrast, approximately one-third of the potato chips demonstrate gamma values below -0.5, and none of these products have gamma values above 0.5, implying that the popularity of these products decreases as the index increases.

[18]To provide some model-free evidence, Appendix Figure A3 presents the conditional distribution of γ_i based on whether the product underwent reformulation or remained unchanged. Consistent with firms correctly predicting consumer demand, there are substantially more reformulations of products with low γ_i, i.e., products sold in stores associated with lower index values and stronger responses to the label.

nificant differences in the likelihood of reformulating products popular with consumers more responsive to the label. For example, looking at changes in the reformulations made to avoid the high-sodium label in Column (2), we first see that Israeli firms barely respond by reformulating products mostly purchased by unaffected consumers — the estimated effect is weak and statistically insignificant. In contrast, Israeli firms are 5 percentage points more likely to target products purchased by more responsive consumers, suggesting again that firms accurately anticipate consumer response and target their efforts accordingly.

4.3.2 Which Producers are Better at Targeting?

So far, we have shown that firms strategically respond to the new labeling policy by bunching at minimum standards and focusing their effort on products in more affected categories or those mostly purchased by consumers responsive to the policy. Correctly anticipating consumer response is crucial to firm success, as reformulations are costly, and a misguided change could harm the sales of an already successful product. In this section, we examine whether certain types of firms are better able to accurately target their reformulation efforts. We specifically focus on firms' prior performance, as measured by: (1) the (demeaned) log revenue in 2019, and (2) the (demeaned) log within-category revenue in 2019 .[19] There are several reasons to believe that successful firms would be better able to effectively target the correct set of products, such as access to additional resources or familiarity with the market. While we cannot discern whether firms become more successful due to their ability to predict market responses or whether market success directly leads to better predictions, understanding which firms are more accurate has important implications for firm success and market dynamics, as detailed in Section 5 below.

One potential concern is that larger firms simply tend to reformulate more products across the board. However, as presented in Appendix Table A10, we find that these firms are

[19] The results are robust to alternative measures, including number of items in category and the number of categories in which the firm operates. The results are presented in Appendix Table A12.

only marginally more likely to reformulate their products. Instead, we focus our analysis on whether, conditional on reformulation, these successful firms are differentially better at predicting consumer responses and targeting the appropriate set of products for reformulation. To this end, we employ a quadruple-differences specification, similar to Equation 3, where we interact our main effects with our measures of firm performance.[20]

Strategic manipulation of nutritional values We begin by examining whether successful suppliers are more likely to strategically reformulate products to avoid the new label. The analysis is presented in Figure 2, which echoes Figure 3, except it focuses on the quadruple-interaction $\text{IL}_i \times \text{Post}_t \times Bin_i \times \log(\text{Size}_i)$, i.e., the difference in responsiveness between more and less successful Israeli producers across different nutritional bins. Panels A and B present the results when defining firm size according to its total revenue definition, while Panels C and D focus on the revenue-by-category definition.

Across specifications, we find that the strategic responses documented in Section 4.3.1 are primarily driven by successful firms. Observing, for instance, Panels A and B, we first note that larger firms are more likely to target products just above the threshold for reformulation compared to smaller firms. This effect is evident for both the 2021 and the 2020 thresholds, a captured by the third and fourth bins, respectively. Similarly, we observe that products reformulated by larger firms are substantially more likely to "end up" just below the thresholds, as evident by the effects in the second and third bins, compared to products by smaller firms.

[20]Formally, we estimate regressions of this structure,

$$\begin{aligned} y_{it} = \beta \times \text{IL}_i \times \text{Post}_t \times \mathbb{1}(\text{Aff.}_i) \times \log(\text{Size}_i) \\ + \alpha_1 \text{IL}_i \times \text{Post}_t \times \mathbb{1}(\text{Aff.}_i) + \alpha_2 \text{IL}_i \times \text{Post}_t \times \log(\text{Size}_i) \\ + \alpha_3 \text{Post}_t \times \mathbb{1}(\text{Aff.}_i) \times \log(\text{Size}_i) + \mu_i + \gamma_t + \epsilon_{it}, \end{aligned} \tag{6}$$

where all variables are as defined in Equation 3, with $\mathbb{1}(\text{Aff.}_i)$ is an indicator variable for affected category or consumer. The term $\log(\text{Size}_i)$ represents different measures of firm size.

Heterogeneity across categories and consumer demographics We next examine whether larger firms are better at targeting the "correct" set of product categories and consumers, namely those that were most responsive to the new labeling scheme. The results are presented in Table 8. Columns (1)-(2) present reformulation decisions to avoid the high-sugar label, and Columns (3)-(4) reformulation decisions to avoid the high-sodium label. The odd columns use (demeaned) log revenue in 2019 and the even columns use (demeaned) log revenue within food category. Finally, Panel A presents heterogeneity by category, and Panel B by consumer types.

We find that firms' ability to accurately target affected categories and consumers is primarily driven by the larger, more successful firms. For instance, in Columns (1) and (2) of Panel A, large firms are significantly more likely, compared to the average-sized firm, to reformulate their products to avoid the high-sugar label in affected categories: 2.3 p.p. for each 1% increase in 2019 revenue, and 2.9 p.p. for each 1% increase in revenue within category. Similarly, Panel B indicates that larger firms are significantly more likely to reformulate products sold to more responsive consumers by 2.4% and 3.4% for total and within-category revenue, respectively. For the high-sodium labels, the effects are similar for affected categories but are smaller and statistically indistinguishable from zero when estimating for products mostly purchased by affected consumer segments.

Taken together, these analyses suggest that while all firms respond to the new regulation, larger, more successful firms are better at predicting and correctly adjusting to demand responses. Larger firms are more strategic in their responses, focus their reformulations on products in the most affected categories, and target products that are mostly sold to consumers responsive to the labels.

Table 8: Heterogeneity in the Effect of FoPL on Reformulations by Product Characteristics and Firm Size

Dependent Variable:	Below the thresholds			
Nutrient:	Sugar		Sodium	
	(1)	(2)	(3)	(4)
Panel A: Affected Categories				
Post × IL × Size × Aff. Cat	0.023	0.029	0.028	0.035
	(0.009),[.01]	(0.013),[.024]	(0.011),[.009]	(0.014),[.013]
Producer Size Measure (pre):	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)
Observations	4,396	4,396	5,196	5,196
Panel B: Affected Consumers				
Post × IL × Size × Aff.Con	0.024	0.034	0.004	0.001
	(0.009),[.007]	(0.013),[.011]	(0.010),[.672]	(0.014),[.956]
Producer Size Measure (pre):	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)
Observations	486	486	486	486

Notes: This table displays results from a quadruple-difference estimation, focusing on the association between firm sizes and reformulating the "right" categories or products. Panel A examines reformulations for products within categories that experienced a decrease in demand ('affected categories'). Panel B examines reformulations of products popular among consumers who responded more to the regulation ('affected consumers'). Both panels show the heterogeneity in the effect of the regulation on reformulations of sugar, sodium, and saturated fat content, with the data covering the period from 2019 to 2020. In Columns (1)-(2) we present the results for sugar, in Columns (3)-(4) we present the results for sodium, and In Columns (5)-(6) we present the result for saturated fat. For each nutrient, we repeat this exercise using two measures of size. In the odd columns, the size variable is based on the (demeaned) log of the firm's total revenue in 2019, while in the even columns, it is based on the (demeaned) log of the firm's category revenue in 2019. We use data from 2019-2020, reflecting firms' decisions that were made before the actual demand shock. All regressions include product and year fixed effects, as well as all lower-order interactions. In Panel A the we cluster the standard errors at the category × producer level, and in Panel B we cluster the standard errors at the product level. Standard errors are clustered at the product level and reported in parentheses. Exact p value reported in the square brackets

5 Implications and Discussion

5.1 Changes in Nutrient Consumption

This section evaluates the effectiveness of the policy to promote healthier nutrition by reducing consumption of sugar, sodium, and saturated fat. Our main analysis finds that consumers reduced purchases of labeled products, and suppliers decreased the content of unhealthy nutrients in response to the label. Since the data only includes aggregated purchased quantities, it is challenging to directly estimate the downstream effect of the labels on actual consumption. Nevertheless, we attempt to evaluate changes in average consumption of nutrients per 100 grams of product following the introduction of the nutritional labels. There are several inherent limitations to this analysis: firstly, we cannot observe the entire basket purchased by consumers. For instance, we cannot determine whether consumers substituted prepackaged foods with fresh vegetables and fruits. Secondly, since all Israeli consumers were impacted by the policy, we lack a valid control group for a formal Difference-in-Difference analysis; thus, we regard these findings as suggestive evidence that supports the notion of Israeli citizens improving their dietary intake.[21]

The results are presented in Table 9. In Column (1), our estimate suggests that following the introduction of the new nutritional labels, consumers purchased, on average, 1.4 grams less sugar per 100 grams of product, a decrease of approximately 10% in sugar consumption compared to the baseline. Similarly, sodium was reduced by 9 milligrams per 100 grams, which is equivalent to a decrease of about 3.5%. Finally, consistent with the results regarding the effect of the label on purchases, we find negative but substantially smaller and statistically insignificant effects on the consumption of saturated fat. Taken together, these results suggest that the policy had limited yet notable success in improving overall consumer nutrition.

[21] Formally, we estimate the following event study design:

Table 9: Change in Nutrient Consumption Over Time

Dependent Variables:	Sugar (1)	Sodium (2)	Sat. Fat (3)
Post_{20}	-1.437	-9.029	-0.093
	(0.540)	(3.994)	(0.136)
	[.008]	[.024]	[.494]
Chain-City FE	✓	✓	✓
MoY FE	✓	✓	✓
Observations	1,585,242	1,715,274	1,707,603
Mean DV (weighted)	13.68	256.66	5.22

Notes: The table presents results from regressing the nutrition concentration against post, chain×city, and month of the year. The coefficient "post" captures the change in concentration of the critical nutrient after the regulation began. Observations are weighted by quantity. Standard errors are clustered at the product level and reported in parentheses. Exact p value reported in the square brackets

5.2 Changes in Competitive Environment

In addition to the impact of the labeling policy on nutritional intake, the labeling scheme also had some unintended consequences on the dynamics within the food industry more generally. In particular, we find evidence that larger, more successful firms were differentially able to accurately target their reformulation efforts, potentially improving their market position following the reform. To examine market outcomes more directly, we derive the Herfindahl–Hirschman Index (HHI) to measure market concentration within city-chain-category at each period. If larger firms gain an advantage by accurately reformulating products, then we would expect market concentration to disproportionately rise in categories that experienced a higher share of product reformulations.

The results are presented in Table 10. Column (1) presents the results at the item level and Column (2) at the producer level. To be conservative, the analysis includes reformulations for saturated fat, though the effects grow in magnitude when these reformulations are excluded. Under both specifications, we find that concentration in more responsive categories increased compared to categories that experienced fewer reformulations. For instance, we find that

Table 10: Reformulations and Market Concentration

Dependent Variable:	HHI	
Market Shares based on:	Item level	Producer level
	(1)	(2)
Post_{20} × (# Reform)/(# Above Thresholds)	0.017	0.021
	(0.003)	(0.005)
	[.001]	[.001]
City-Chain-Cat FE	✓	✓
Date FE	✓	✓
Within R^2	0.002	0.001
R^2	0.828	0.881
Observations	50,657	50,657

Notes: All regressions include City × Chain × Category fixed-effects. Standard errors are clustered at the City × Chain × Category level and reported in parentheses. Exact p value reported in the square brackets

producer-level HHI grows by 0.021 in categories more affected by reformulation, an increase of 4.75% compared to the baseline. These results suggest that the policy allowed larger firms to increase their relative market share at the expense of smaller producers.

6 Conclusion

This paper studies a nationwide policy that mandated front-of-package labels on unhealthy products. We find that consumers responded to the policy by reducing their consumption of food products marked as unhealthy. Additionally, local producers reformulated products in order to avoid unhealthy labels. Firms were strategic in their reformulation decisions: they focused their efforts on products close to the labeling threshold and tended to bunch just below it. Moreover, firms were adept at identifying and prioritizing the reformulation of products that would have been most adversely affected by the labels, either by focusing on specific categories or products predominantly purchased by specific consumer segments. Notably, larger, more successful firms were better able to accurately predict consumer responses and focus their efforts on a specific set of products.

Several potential explanations may explain firms' differential ability to predict and respond to consumer demand. For instance, possessing correct beliefs about consumer preferences, as well as the capacity to reformulate products, might be necessary conditions for firms to make such adjustments effectively. Furthermore, it remains unclear whether a firm's past success is the cause of, or a precursor to, its ability to respond to market changes. Separating these channels is beyond the scope of this paper, and we leave such analysis for future research.

Nevertheless, our results emphasize the importance of careful monitoring and consideration of the long-term effects of regulations on market dynamics, including both demand and supply forces. This analysis underscores the need for policymakers to design regulations that encourage competition and innovation without disproportionately benefiting larger players. For managers, our findings highlight that the market environment and consumer demand are dynamic. We provide empirical evidence supporting the concept discussed in Teece et al. (1997), stressing the importance of accurately anticipating and adapting to shifts in consumer demand.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge, "When Should You Adjust Standard Errors for Clustering?," *The Quarterly Journal of Economics*, February 2023, *138* (1), 1–35.

Adalja, Aaron, Jūra Liaukonytė, Emily Wang, and Xinrong Zhu, "GMO and Non-GMO Labeling Effects: Evidence from a Quasi-Natural Experiment," *Marketing Science*, March 2023, *42* (2), 233–250.

Aguirregabiria, Victor and Jihye Jeon, "Firms' Beliefs and Learning: Models, Identification, and Empirical Evidence," *Review of Industrial Organization*, March 2020, *56* (2), 203–235.

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

Alé-Chilet, Jorge and Sarah Moshary, "Beyond Consumer Switching: Supply Responses to Food Packaging and Advertising Regulations," *Marketing Science*, March 2022, *41* (2), 243–270.

Anderson, Michael L, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 2008, *103* (484), 1481–1495.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press, 2009.

Araya, Sebastián, Andrés Elberg, Carlos Noton, and Daniel Schwartz, "Identifying Food Labeling Effects on Consumer Behavior," *Marketing Science*, September 2022, *41* (5), 982–1003.

Ater, Itai and Or Avishay-Rizi, "Price Saliency and Fairness: Evidence from Regulatory Shaming," *Working Paper*, 2022.

Barahona, Nano, Cristóbal Otero, and Sebastián Otero, "Equilibrium Effects of Food Labeling Policies," *Econometrica*, 2023, *91* (3), 839–868.

_ , _ , _ , and Joshua Kim, "Single-Threshold Food Labeling Policies," *Working Paper*, 2025.

Barney, Jay, "Firm Resources and Sustained Competitive Advantage," *Journal of Management*, March 1991, *17* (1), 99–120.

Bernard, Andrew B, Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes, "The origins of firm heterogeneity: A production network approach," *Journal of Political Economy*, 2022, *130* (7), 1765–1804.

Bloom, N. and J. Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *The Quarterly Journal of Economics*, November 2007, *122* (4), 1351–1408.

Bloom, Nicholas and John Van Reenen, "Why Do Management Practices Differ across Firms and Countries?," *Journal of Economic Perspectives*, February 2010, *24* (1), 203–224.

_ , Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, "Does management matter? Evidence from India," *The Quarterly journal of economics*, 2013, *128* (1), 1–51.

Byrne, David P. and Nicolas de Roos, "Learning to Coordinate: A Study in Retail Gasoline," *American Economic Review*, February 2019, *109* (2), 591–619.

Cao, Xinyu and Juanjuan Zhang, "Preference learning and demand forecast," *Marketing Science*, 2021, *40* (1), 62–79.

Davis, Lucas W. and Gilbert E. Metcalf, "Does Better Information Lead to Better Choices? Evidence from Energy-Efficiency Labels," *Journal of the Association of Environmental and Resource Economists*, September 2016, *3* (3), 589–625.

DellaVigna, Stefano and Matthew Gentzkow, "Uniform Pricing in U.S. Retail Chains," *The Quarterly Journal of Economics*, November 2019, *134* (4), 2011–2084.

Doraszelski, Ulrich, Gregory Lewis, and Ariel Pakes, "Just Starting Out: Learning and Equilibrium in a New Market," *American Economic Review*, March 2018, *108* (3), 565–615.

Dubois, Pierre, Rachel Griffith, and Martin O'Connell, "The Effects of Banning Advertising in Junk Food Markets," *The Review of Economic Studies*, January 2018, *85* (1), 396–436.

Ellickson, Paul B, Mitchell J Lovett, and Bhoomija Ranjan, "Product launches with new attributes: a hybrid conjoint–consumer panel technique for estimating demand," *Journal of Marketing Research*, 2019, *56* (5), 709–731.

Goldfarb, Avi and Mo Xiao, "Who Thinks about the Competition? Managerial Ability and Strategic Entry in US Local Telephone Markets," *American Economic Review*, December 2011, *101* (7), 3130–3161.

Hitsch, Günter J., "An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty," *Marketing Science*, January 2006, *25* (1), 25–50.

Hortaçsu, Ali, Olivia R Natan, Hayden Parsley, Timothy Schwieg, and Kevin R Williams, "Organizational structure and pricing: Evidence from a large us airline," *The Quarterly Journal of Economics*, 2024, *139* (2), 1149–1199.

Hortaçsu, Ali and Steven L. Puller, "Understanding Strategic Bidding in Multi-Unit Auctions: A Case Study of the Texas Electricity Spot Market," *The RAND Journal of Economics*, 2008, *39* (1), 86–114.

Houde, Sébastien, "How consumers respond to product certification and the value of energy information," *The RAND Journal of Economics*, June 2018, *49* (2), 453–477.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India *," *Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.

Huang, Yufeng, Paul B. Ellickson, and Mitchell J. Lovett, "Learning to Set Prices," *Journal of Marketing Research*, April 2022, *59* (2), 411–434.

Hui, Xiang, Maryam Saeedi, Zeqian Shen, and Neel Sundaresan, "Reputation and regulations: Evidence from eBay," *Management Science*, 2016, *62* (12), 3604–3616.

Jin, Ginger Zhe and Andrew Kato, "Price, quality, and reputation: Evidence from an online field experiment," *The RAND Journal of Economics*, 2006, *37* (4), 983–1005.

Kim, Youngju, SunAh Kim, and Neeraj Arora, "GMO Labeling Policy and Consumer Choice," *Journal of Marketing*, May 2022, *86* (3), 21–39.

Levitt, Steven D., John A. List, and Chad Syverson, "Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant," *Journal of Political Economy*, 2013, *121* (4), 643–681.

Loecker, Jan De, "Detecting Learning by Exporting," *American Economic Journal: Microeconomics*, August 2013, *5* (3), 1–21.

Luca, Michael, "Reviews, reputation, and revenue: The case of Yelp.com," NOM Unit Working Paper 12-016, Harvard Business School March 2016.

Nevo, Aviv, "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 2001, *69* (2), 307–342.

Pachali, Max J., Marco J.W. Kotschedoff, Arjen van Lin, Bart J. Bronnenberg, and Erica van Herpen, "How Do Nutritional Warning Labels Affect Prices?," *Journal of Marketing Research*, February 2023, *60* (1), 92–109.

Porter, Michael E, *Competitive Advantage: Creating and Sustaining Superior Performance*, The free, New York, 1985.

Rao, Anita and Raluca Ursu, "The Impact of Voluntary Labeling," *Marketing Science*, forthcoming.

Rubinstein, Yona and Dror Brenner, "Pride and Prejudice: Using Ethnic Sounding Names and Inter-ethnic Marriages to Identify Labour Market Discrimination," *Review of Economic Studies*, 2014, *81* (1), 389–425.

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

Strulov-Shlain, Avner, "Firms as Model-Free Decision Makers – Evidence from a Reform," *Working Paper*, August 2021.

Syverson, Chad, "Product Substitutability and Productivity Dispersion," *Review of Economics and Statistics*, May 2004, *86* (2), 534–550.

_ , "What Determines Productivity?," *Journal of Economic Literature*, June 2011, *49* (2), 326–365.

Tadelis, Steven, "Reputation and feedback systems in online platform markets," *Annual review of economics*, 2016, *8* (1), 321–340.

_ , Christopher Hooton, Utsav Manjeer, Daniel Deisenroth, Nils Wernerfelt, Nick Dadson, and Lindsay Greenbaum, "Learning, Sophistication, and the Returns to Advertising: Implications for Differences in Firm Performance," *Working Paper*, 2023.

Teece, David J., Gary Pisano, and Amy Shuen, "Dynamic capabilities and strategic management," *Strategic Management Journal*, August 1997, *18* (7), 509–533.

Wooldridge, Jeffrey, "Econometric Analysis of Cross Section and Panel Data," Technical Report, The MIT Press 2001.

Yang, Nathan, "Learning in retail entry," *International Journal of Research in Marketing*, June 2020, *37* (2), 336–355.

Web Appendix A

Data Cleaning - For online publication

In this section, we add further details on the construction of the hand-collected nutrition data set.

1. **Remove products that appear fewer than three times:** We start by eliminating products from the data that have a low occurrence rate. By setting a threshold of three appearances, we ensure that only products with a sufficient number of instances are considered for further analysis. This helps remove outliers or rare occurrences that may not provide meaningful insights.

2. **Drop nutrients compositions observed only once:** In this step, we remove nutrient compositions that are unique and appear only once in the data. These compositions cannot be validated due to their lack of repetition.

3. **Update nutritional values based on surrounding observations:** If a given observation has different nutritional values compared to its previous and following observations, but the previous and following observations have the same values, we update the values of that observation to match the surrounding ones. This approach helps to detect mistakes in the digitization process. It assumes that it is unlikely for a product to undergo reformulation and then revert to an earlier formula. For instance, if the sequence is "aba," we would change the middle value to match the adjacent values, resulting in "aaa."

4. **Drop observations with inconsistent appearance sequences:** In certain cases, due to stock management issues, we may encounter occurrences of a combination that happened before the previous combination. For instance, in the sequence "aabbaab,"

we identify such inconsistent appearances. To address this, we drop these observations. As a result, we remove the latter two 'a's ("aabb~~aa~~b"), resulting in "aabbb."

5. **Impute missing observations between identical values based on time:** This approach assumes that the number of reformulations is low and that it is unlikely for multiple reformulations to occur, starting and ending with the same nutritional values. Consequently, when an item exhibits identical values on two different dates, we impute all the periods between these dates with the same nutritional value.

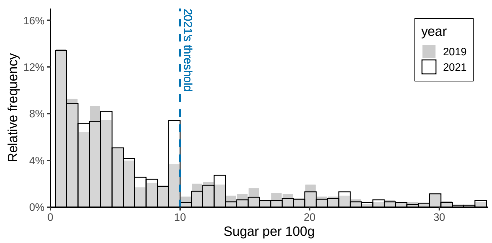
   Code the transition time between nutritional value changes: For products where we detect a change in nutritional values between two observations, we code the transition time as the midpoint between the two dates. Choosing the midpoint aims to minimize the average distance between the actual change point and the assumed change point.

Web Appendix B

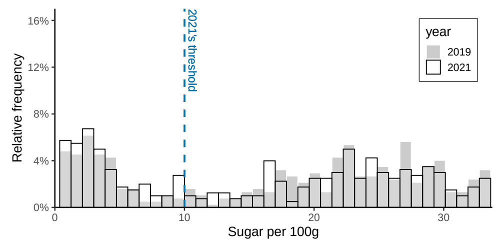
Further Results and Robustness - For online publication

Panel A: Israeli Products

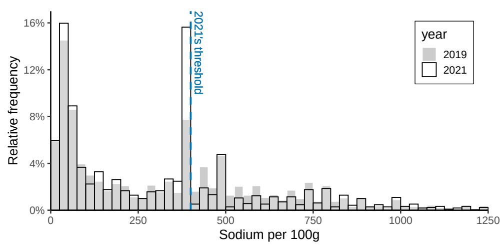
Panel B: International Products



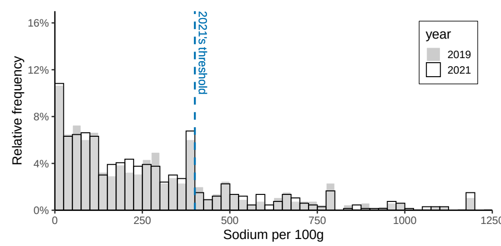
A1: Sugar



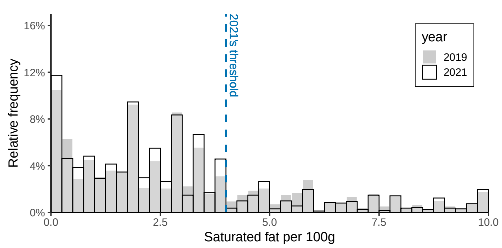
B1: Sugar



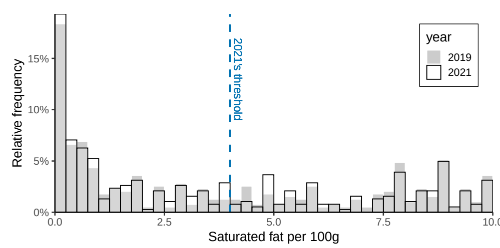
A2: Sodium



B2: Sodium



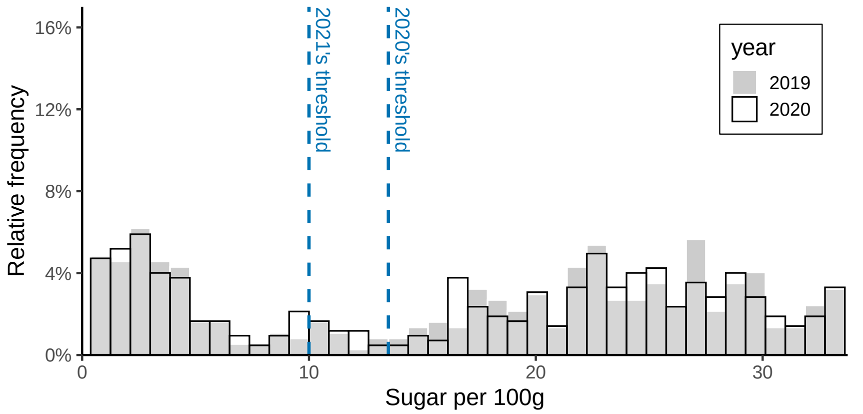
A3: Saturated Fat



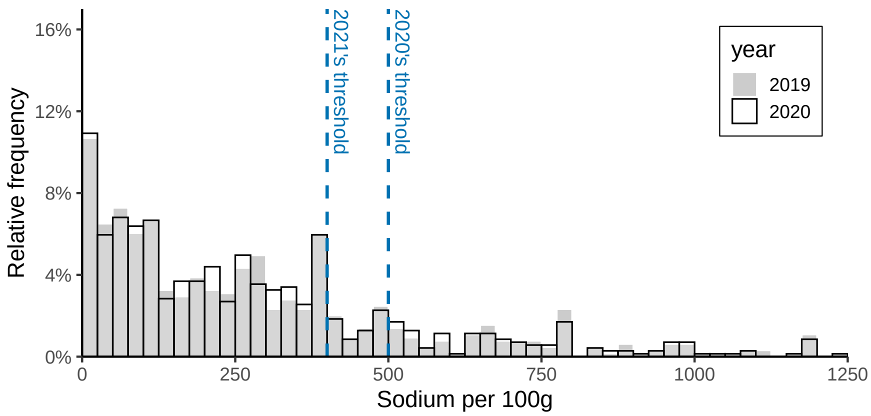
B3: Saturated Fat

Figure A1: Distribution of Products by Nutritional Values

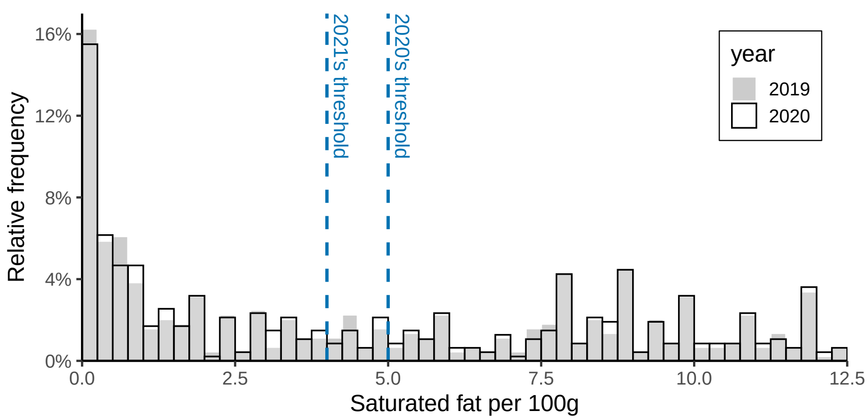
Notes: The figure illustrates the distribution of products based on their nutrient values. Panel A displays the distributions of Israeli products, while Panel B depicts the distributions for international products. The vertical dashed lines represent the thresholds that were implemented at the start of 2020 and 2021. The distribution of products in 2019, indicated by the gray color, is contrasted with the overlayed distribution in 2021, depicted in white.



1: Sugar



B: Sodium



C: Saturated Fat

Figure A2: Distribution of International Products by Nutritional Values

Notes: This figure illustrates the distribution of international products based on their nutrient values. The vertical dashed lines represent the thresholds that were implemented at the start of 2020 and 2021. The distribution of products in 2019, indicated by the gray color, is contrasted with the overlayed distribution in 2020, depicted in white.

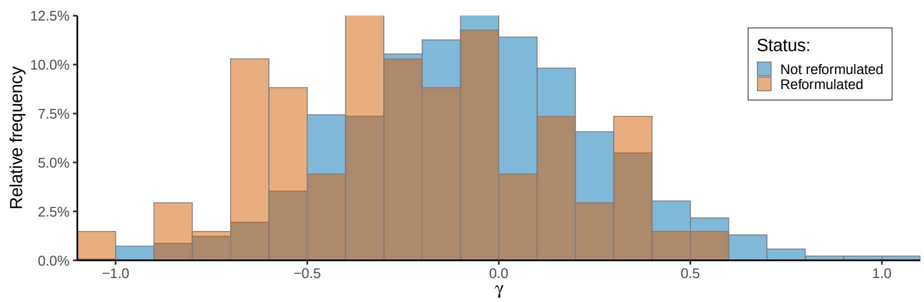


Figure A3: γ Distribution and Reformulations

Notes: The figure illustrates the distribution of γ values obtained from the estimation of Equation 5. The orange bars represent the distribution of products that underwent reformulations, while the blue bars represent the distribution of products that did not undergo reformulation.

Table A1: The Effect of FoPL on Quantities: Allowing for Change in the Demand for Nutrients Over time and Higher Polynomial degree

Dependent Variable:	log(q)					
	(1)	(2)	(3)	(4)	(5)	(6)
Post × $\mathbb{1}$(has labels)	-0.156***			-0.149***		
	(0.036)			(0.049)		
Post × (# of labels)		-0.101***			-0.136***	
		(0.022)			(0.035)	
Post × $\mathbb{1}$(Sugar > 13.5)			-0.250***			-0.231***
			(0.062)			(0.061)
Post × $\mathbb{1}$(Sodium > 500)			-0.105**			-0.171**
			(0.051)			(0.077)
Post × $\mathbb{1}$(Sat. fat > 5)			0.022			-0.023
			(0.052)			(0.062)
log(p)	-2.100***	-2.103***	-2.107***	-2.101***	-2.103***	-2.107***
	(0.094)	(0.095)	(0.093)	(0.094)	(0.095)	(0.094)
Nutrients	✓	✓	✓	✓	✓	✓
Nutrients^2	✓	✓	✓			
Nutrients × Post				✓	✓	✓
Thresholds dummies	✓	✓	✓	✓	✓	✓
Observations	1,402,525	1,402,525	1,402,525	1,402,525	1,402,525	1,402,525

Notes: The table displays the estimation results for Equation 1. Column (1) presents the average effect of labels on quantity, comparing products with and without labels. Column (2), reports the linear relationship between the number of labels and their impact on demand. Column (3), shows the effect of each label on demand. For this robustness check, we interacted the nutrients: sugar, sodium, and saturated fat, with the Post_{20} indicator, to account for any unrelated changes in preferences for these nutrients. The observations are weighted by pre-regulation product × store revenue. Additional covariates include thresholds dummies, nutrient content, product, city × retailer, and month-fixed effects. Standard errors are clustered at the product level and reported in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A2: The Effect of FoPL on Quantities: Different Periods

Dependent Variable:	log(q)					
Sample	2019,2020 (All year)			2019,2021 (First Half)		
	(1)	(2)	(3)	(4)	(5)	(6)
Post $\times$ $\mathbb{1}$(has labels)	-0.109***			-0.074***		
	(0.031)			(0.025)		
Post $\times$ (# of labels)		-0.064***			-0.046***	
		(0.019)			(0.015)	
Post $\times$ $\mathbb{1}$(Sugar > Threshold)			-0.187***			-0.149***
			(0.051)			(0.030)
Post $\times$ $\mathbb{1}$(Sodium > Threshold)			-0.071*			-0.007
			(0.042)			(0.026)
Post $\times$ $\mathbb{1}$(Sat. fat > Threshold)			0.036			0.021
			(0.043)			(0.027)
log(p)	-2.051***	-2.051***	-2.046***	-1.903***	-1.904***	-1.905***
	(0.086)	(0.086)	(0.086)	(0.161)	(0.161)	(0.161)
Sugar + Saturated Fat + Sodium	✓	✓	✓	✓	✓	✓
Thresholds dummies	✓	✓	✓	✓	✓	✓
Observations	1,949,408	1,949,408	1,949,408	1,347,353	1,347,353	1,347,353

Notes: The table displays the estimation results for Equation 1. Column (1) presents the average effect of labels on quantity, comparing products with and without labels. Column (2), reports the linear relationship between the number of labels and their impact on demand. Column (3), shows the effect of each label on demand. The observations are weighted by pre-regulation product $\times$ store revenue. Additional covariates include thresholds dummies, nutrient content, product, city $\times$ retailer, and month-fixed effects. Standard errors are clustered at the product level and reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: The Effect of FoPL on Prices

Dependent Variable:	log(p)		
	(1)	(2)	(3)
Post × $\mathbb{1}$(has labels)	-0.006**		
	(0.003)		
Post × (# of labels)		-0.004**	
		(0.002)	
Post × $\mathbb{1}$(Sugar > 13.5)			-0.010**
			(0.004)
Post × $\mathbb{1}$(Sodium > 500)			-0.011**
			(0.005)
Post × $\mathbb{1}$(Sat. Fat > 5)			0.002
			(0.004)
Observations	1,402,548	1,402,548	1,402,548

Notes: The table displays the effect ot the FoPL on prices. Column (1) presents the average effect of labels on prices, comparing products with and without labels. Column (2), reports the linear relationship between the number of labels and their impact on prices. Column (3), shows the effect of each label on price. The observations are weighted by pre-regulation product × store revenue. Additional covariates include thresholds dummies, nutrient content, product, city × retailer, and month-fixed effects. Standard errors are clustered at the product level and reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: The Effect of FoPL on Reformulations: Different Periods

Panel A: 2020 Thresholds Only			
Dependent Variable:	Below 2020 thresholds		
Nutrient:	Sugar	Sodium	Sat. Fat
	(1)	(2)	(3)
Post × IL	0.012*	0.038***	0.016***
	(0.007)	(0.012)	(0.006)
Observations	4,396	5,196	5,192
Panel B: 2020-2021			
Dependent Variable:	Below 2021 thresholds		
Nutrient:	Sugar	Sodium	Sat. Fat
	(1)	(2)	(3)
Post × IL	0.020**	0.046***	0.006
	(0.010)	(0.011)	(0.005)
Observations	4,496	5,138	5,134
Panel C: 2019-2021			
Dependent Variable:	Below 2021 thresholds		
Nutrient:	Sugar	Sodium	Sat. Fat
	(1)	(2)	(3)
Post × IL	0.025**	0.069***	0.011
	(0.012)	(0.017)	(0.009)
Observations	3,718	4,368	4,364

Notes: This table shows the likelihood of reformulations that firms undertake to avoid warning labels. In Panel A, we use data from 2019-2020 and focus only on changes made with respect to the 2020 thresholds. That is, a product that was between the first and second thresholds in 2019 and below both thresholds in 2020 will be considered as not reformulated since it was always below the 2020 threshold. In Panel B, we use data from all products available in both 2020 and 2021, focusing on the 2021 thresholds. In Panel C, we include all products that were available in 2019 and 2021, using the 2021 thresholds. All regressions include product and year fixed effects. Standard errors are clustered at the category × producer level and reported in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A5: The Effect of FoPL on Reformulations: Placebo Thresholds

Panel A: Sample: Above the thresholds in 2019

Dependent Variables:	I(sug≤20)	I(sod≤800)	I(sat≤8)
Nutrient:	Sugar	Sodium	Sat. Fat
	(1)	(2)	(3)
Post × IL	0.005	-0.022	0.017**
	(0.009)	(0.024)	(0.007)
Observations	1,618	1,360	2,026

Panel B: Sample: Below the thresholds in 2019

Dependent Variables:	I(sug≤5)	I(sod≤200)	I(sat≤2)
Nutrient:	Sugar	Sodium	Sat. Fat
	(1)	(2)	(3)
Post × IL	0.012	0.010	0.002
	(0.014)	(0.007)	(0.015)
Observations	2,778	3,836	3,166

Notes: This table presents a variation of Table 4 where we include only products above the thresholds set for 2020 and use fake thresholds. In Panel A, we include all products above the thresholds, and we set the fictitious thresholds to be at 200% of the 2021 thresholds (i.e., 20g, 800mg, and 8g per 100g for sugar, sodium, and saturated fat, respectively). In the second panel, we keep only products that were below the 2020 thresholds. Here, we set the fictitious thresholds to be at 50% of the 2021 thresholds (i.e., 5g, 200mg, and 2g per 100g for sugar, sodium, and saturated fat, respectively). All the regressions include item and year fixed effect. None of the specifications show a significant negative coefficient.Standard errors are clustered at the category × producer level and reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: The Effect of FoPL on Reformulations

Panel A: Ineffective Reduction			
Dependent Variable:	$\text{Nutrient}_{t,n} < \text{Nutrient}_{\text{pre},n}$		
Nutrient:	Sugar	Sodium	Sat. Fat
	(1)	(2)	(3)
Post × IL	-0.003	-0.005	0.009
	(0.027)	(0.028)	(0.020)
Observations	4,332	4,886	5,076
Panel B: Below the Thresholds Conditional on a Change			
Dependent Variable:	Below the thresholds		
Nutrient:	Sugar	Sodium	Sat. Fat
	(1)	(2)	(3)
Post × IL	0.078**	0.178***	0.069
	(0.035)	(0.052)	(0.047)
Observations	1,040	1,392	880

Notes: This table presents a variation of Table 4. In Panel A, we modify the dependent variable to be 1 if the nutrient value is lower than the pre-regulation nutrient value and 0 otherwise. In the Panel B, we keep only products that have changed their nutrient values, and preserve the regular dependent variable that equals 1 when product nutrient values cross the thresholds. All the regressions include item and year fixed effects. None of the specifications show a significant negative coefficient.Standard errors are clustered at the category × producer level and reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: The Effect of FoPL on Reformulations: Using Different Measures of the Effect on Demand

Dependent Variable:	Below the thresholds			
Nutrient:	Sugar	Sodium	Sugar	Sodium
	(1)	(2)	(3)	(4)
Panel A: Avg. median effect on Cat. and Con.				
Post × IL × Below Median	0.059**	0.062*	0.013**	0.018
	(0.029)	(0.033)	(0.006)	(0.012)
Effect on demand based on::	Categories	Categories	Consumers	Consumers
Observations	4,396	5,196	3,928	3,928
Panel B: Size Median - By categories				
Post × IL × Size × Below Median	0.024***	0.031***	0.032**	0.033**
	(0.009)	(0.011)	(0.013)	(0.015)
Producer Size Measure (pre):	log(Rev)	log(Rev)	log(Cat. Rev)	log(Cat. Rev)
Observations	4,396	5,196	4,396	5,196
Panel C: Size Median - By consumers				
Post × IL × Size × Below Median	0.005*	0.007	0.010**	0.009
	(0.003)	(0.005)	(0.005)	(0.006)
Producer Size Measure (pre):	log(Rev)	log(Rev)	log(Cat. Rev)	log(Cat. Rev)
Observations	3,928	3,928	3,928	3,928
Panel D: Avg. continuous effect on Cat. and Con.				
Post × IL × effect(continuous)	0.032	0.006	0.030**	0.025*
	(0.039)	(0.095)	(0.012)	(0.015)
Effect on demand based on::	Categories	Categories	Consumers	Consumers
Observations	2,598	3,086	486	486
Panel E: Size continuous - By categories				
Post × IL × Size × effect(continuous)	0.024*	0.050	0.018	0.048
	(0.014)	(0.031)	(0.024)	(0.038)
Producer Size Measure (pre):	log(Rev)	log(Rev)	log(Cat. Rev)	log(Cat. Rev)
Observations	2,598	3,086	2,598	3,086
Panel F: Size continuous - By consumers				
Post × IL × Size × effect(continuous)	0.012**	0.000	0.019**	-0.001
	(0.005)	(0.004)	(0.008)	(0.007)
Producer Size Measure (pre):	log(Rev)	log(Rev)	log(Cat. Rev)	log(Cat. Rev)
Observations	486	486	486	486

Notes: This table expands upon the findings of Tables 6 and 8. Rather than categorizing products based on a binary definition of 'affected' versus 'unaffected' categories, determined by a negative within-category effect, this analysis utilizes the coefficient of the within-category effect itself. The aim is to estimate a linear association between the effect size on demand and product reformulations. All regressions incorporate product and year fixed effects, in addition to all lower-order interactions. Standard errors are clustered at the category × producer level and reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Heterogeneity in the Effect of FoPL on Reformulations by Product Characteristics - IV approach

Dependent Variable:	Below the thresholds	
Nutrient	Sugar	Sodium
	(1)	(2)
Panel A: Affected Categories		
Post × IL	0.004*	0.023*
	(0.003)	(0.013)
Post × IL × Aff. Cat	0.054*	0.082***
	(0.028)	(0.030)
Observations	4,396	5,196
Panel B: Affected Consumers		
Post × IL	0.000	0.015
	(0.000)	(0.015)
Post × IL × Aff.Con	0.050***	0.050*
	(0.019)	(0.028)
Observations	486	486

Notes: This table extends the analysis presented in Table 6, applying the same triple-difference estimation methods to examine the relationship between reformulations and the 'right' products. The main difference is that in this analysis we construct the pool of affected categories using pre-regulation nutrition values as an instrument for labeling when we estimate the heterogeneity in demand. All regressions include product and year fixed effects, as well as all lower-order interactions. In Panel A the we cluster the standard errors at the category × producer level, and in Panel B we cluster the standard errors at the product level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Heterogeneity in the Effect of FoPL on Reformulations by Product Characteristics: Different Periods

Dependent Variable:	$\mathbb{1}(\text{Nutrient}_n \leq \text{Threshold}_{tn})$					
Year:	2020-2021			2019-2021		
Nutrient:	Sugar (1)	Sodium (2)	Sat. fat (3)	Sugar (4)	Sodium (5)	Sat. fat (6)
Panel A: Affected Categories						
Post × $\mathbb{1}$(IL)	-0.002 (0.003)	0.019** (0.008)	0.008 (0.007)	-0.001 (0.003)	0.022 (0.017)	0.010 (0.008)
Post × $\mathbb{1}$(IL) × Aff.Cat	0.064*** (0.022)	0.064*** (0.024)	-0.007 (0.011)	0.087*** (0.030)	0.108*** (0.033)	0.001 (0.025)
Observations	5,134	5,862	5,860	4,820	5,603	5,601
Panel B: Affected Consumers						
Post × $\mathbb{1}$(IL)	0.000*** (0.000)	0.018 (0.018)	0.000 (0.000)	0.000*** (0.000)	0.035 (0.025)	0.000 (0.000)
Post × $\mathbb{1}$(IL) × Aff.Con	0.064*** (0.022)	0.054* (0.029)	0.008 (0.008)	0.072*** (0.023)	0.053 (0.037)	0.000 (0.011)
Observations	456	456	456	456	456	456

Notes: This table extends the analysis presented in Table 6, applying the same triple-difference estimation methods to examine the relationship between reformulations and the 'right' products. The key difference here is that this analysis focuses on alternative time periods. Columns (1)-(3) present the changes made between 2020 and 2021, and Columns (4)-(6) present the changes between 2019 and 2021. All regressions include product and year fixed effects, as well as all lower-order interactions. In Panel A the we cluster the standard errors at the category × producer level, and in Panel B we cluster the standard errors at the product level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Heterogeneity in the Effect of FoPL on Reformulations by Firm Size

Panel A: 2019-2020						
Dependent Variable:	Below the thresholds					
Nutrient:	Sugar		Sodium		Sat. Fat	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × IL	0.014**	0.017**	0.046***	0.053***	0.009	0.013
	(0.006)	(0.007)	(0.012)	(0.013)	(0.009)	(0.008)
Post × IL × Size	0.005	0.008	0.010*	0.010	-0.003	-0.001
	(0.004)	(0.006)	(0.005)	(0.007)	(0.003)	(0.003)
Producer Size Measure (pre):	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)
Observations	4,396	4,396	5,196	5,196	5,192	5,192
Panel B: 2019-2021						
Dependent Variable:	Below 2021 thresholds					
Nutrient:	Sugar		Sodium		Sat. Fat	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × IL	0.010	0.019**	0.055***	0.067***	0.005	0.010
	(0.007)	(0.009)	(0.013)	(0.015)	(0.010)	(0.009)
Post × IL × Size	0.009**	0.016**	0.009	0.007	-0.003	-0.006
	(0.005)	(0.007)	(0.006)	(0.008)	(0.004)	(0.004)
Producer Size Measure (pre):	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)
Observations	3,718	3,718	4,368	4,368	4,364	4,364
Panel C: 2020-2021						
Dependent Variable:	Below 2021 thresholds					
Nutrient:	Sugar		Sodium		Sat. Fat	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × IL	0.009*	0.015**	0.034***	0.043***	0.006	0.007
	(0.006)	(0.007)	(0.009)	(0.010)	(0.007)	(0.006)
Post × IL × Size	0.004	0.011**	0.004	0.004	-0.001	-0.003
	(0.004)	(0.006)	(0.004)	(0.006)	(0.003)	(0.003)
Producer Size Measure (pre):	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)
Observations	4,478	4,396	5,118	5,036	5,114	5,032

Notes: The table presents the estimation results of triple-difference estimation, focusing on the associations between firms' size and reformulations. We use two measures of size. In the odd columns, the size variable is based on the (demeaned) log of the firm's total revenue in 2019, while in the even columns, it is based on the (demeaned) log of the firm's category revenue in 2019. We present the results for three nutrients: columns (1) and (2) show sugar, columns (3) and (4) show sodium, and columns (5) and (6) show saturated fat. In Panel A, we use data from 2019-2020 and focus only on changes made with respect to the 2020 thresholds. That is, a product that was between the first and second thresholds in 2019 and below both thresholds in 2020 will be considered as not reformulated since it was always below the 2020 threshold. In Panel B, we use data from all products available in both 2020 and 2021, focusing on the 2021 thresholds. In Panel C, we include all products that were available in 2019 and 2021, using the 2021 thresholds. The table presents weak evidence for the role of size and the likelihood to reformulate. All regressions include product and year fixed effects, as well as all lower-order interactions. Standard errors are clustered at the category × producer level and reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Heterogeneity in the Effect of FoPL on Reformulations by Product Characteristics and Firm Size: Different Periods

Dependent Variable:	Below 2021 thresholds					
Nutrient:	Sugar		Sodium		Sat. Fat	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Reformulation Affected Categories - 2020-2021						
Post × IL	-0.004	-0.002	0.018**	0.019**	0.009	0.008
	(0.003)	(0.002)	(0.009)	(0.008)	(0.008)	(0.007)
Post × IL × Aff. Cat	0.043***	0.051***	0.036*	0.059***	-0.010	-0.005
	(0.014)	(0.014)	(0.019)	(0.022)	(0.012)	(0.011)
Post × IL × Size	-0.001	0.002	-0.002	0.000	-0.005	-0.006
	(0.002)	(0.001)	(0.003)	(0.003)	(0.004)	(0.004)
Post × IL × Size × Aff. Cat	0.013**	0.017**	0.017*	0.018	0.010**	0.008
	(0.006)	(0.008)	(0.010)	(0.013)	(0.005)	(0.006)
Producer Size Measure (pre):	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)
Observations	4,478	4,396	5,118	5,036	5,114	5,032
Panel B: Reformulation Affected Consumers 2020-2021						
Post × IL	0.000***	0.000**	0.019	0.019	0.000	0.000*
	(0.000)	(0.000)	(0.019)	(0.019)	(0.000)	(0.000)
Post × IL × Aff.Con	0.020*	0.049***	0.001	0.043	0.010	0.015
	(0.011)	(0.017)	(0.022)	(0.029)	(0.010)	(0.014)
Post × IL × Size	0.000**	0.000**	0.011	0.001	0.000	0.000***
	(0.000)	(0.000)	(0.011)	(0.002)	(0.000)	(0.000)
Post × IL × Size × Aff.Con	0.025***	0.031**	0.018	0.020	-0.001	-0.014
	(0.009)	(0.013)	(0.015)	(0.019)	(0.001)	(0.013)
Producer Size Measure (pre):	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)	log(Rev)	log(Cat. Rev)
Observations	424	424	424	424	424	424

Notes:
This table extends the analysis presented in Table 8, applying the same quadruple-difference estimation methods to examine the relationship between firm sizes and the reformulation of 'right' products. The key difference here is that this analysis focuses on alternative time periods examining the second phase of the regulation, by comparing 2020 and 2021. All regressions include product and year fixed effects, as well as all lower-order interactions. In Panel A the we cluster the standard errors at the category × producer level, and in Panel B we cluster the standard errors at the product level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Heterogeneity in the Effect of FoPL on Reformulations by Product Characteristics and Different Firm Size Measures

Dependent Variable:	Below the thresholds					
Nutrient:	Sugar		Sodium		Sat. Fat	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Affected Categories						
Post × IL × Aff. Cat	0.020	-0.010	0.023	0.002	-0.001	0.002
	(0.018)	(0.017)	(0.030)	(0.030)	(0.023)	(0.021)
Post × IL × Size × Aff. Cat	0.003***	0.010***	0.005**	0.010***	0.001	-0.002
	(0.001)	(0.003)	(0.002)	(0.004)	(0.002)	(0.006)
Producer Size Measure (pre):	No. of Items in Cat.	No. of Categories	No. of Items in Cat.	No. of Categories	No. of Items in Cat.	No. of Categories
Observations	4,396	4,396	5,196	5,196	5,192	5,192
Panel B: Affected Consumers						
Post × IL × Aff.Con	-0.005	-0.005	0.043	0.026	-0.010	-0.024
	(0.011)	(0.011)	(0.031)	(0.036)	(0.010)	(0.016)
Post × IL × Size × Aff.Con	0.004**	0.006**	0.001	0.003	0.001	0.003*
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.002)
Producer Size Measure (pre):	No. of Items in Cat.	No. of Categories	No. of Items in Cat.	No. of Categories	No. of Items in Cat.	No. of Categories
Observations	486	486	486	486	486	486

Notes: This table displays results from a quadruple-difference estimation, focusing on the association between firm sizes and reformulating the "right" categories or products. Panel A examines reformulations for products within categories that experienced a decrease in demand ('affected categories'). Panel B examines reformulations of products popular among consumers who responded more to the regulation ('affected consumers'). Both panels show the heterogeneity in the effect of the regulation on reformulations of sugar, sodium, and saturated fat content, with the data covering the period from 2019 to 2020. In Columns (1)-(2) we present the results for sugar, in Columns (3)-(4) we present the results for sodium, and In Columns (5)-(6) we present the result for saturated fat. For each nutrient, we repeat this exercise using two measures of size. In the odd columns, the size variable is based on the (demeaned) number of category items offered by the firm in 2019, while in the even columns, it is based on the (demeaned) number of categories the firm operated in during 2019. We use data from 2019-2020, reflecting firms' decisions that were made before the actual demand shock. All regressions include product and year fixed effects, as well as all lower-order interactions. In Panel A the we cluster the standard errors at the category × producer level, and in Panel B we cluster the standard errors at the product level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.