The demand side of firm growth: Evidence from Mexico*

Louise Guillouët
Columbia University

Enrique Seira
ITAM

JOB MARKET PAPER

PRELIMINARY AND INCOMPLETE, PLEASE DO NOT CIRCULATE

October 25, 2021
[Click here for the latest version]

Abstract: In order for firms to grow, they must find customers who value their products. Information frictions may prevent customers from knowing the true quality of products, leading them to favor firms that already have a large reputation instead of small firms that could grow. In this paper, we ask how much uncertainty about product quality can account for the differences between large international firms and small domestic firms. We study this question in the context of the consumer goods industry in Mexico, where uncertainty about product quality is prevalent and consumers are willing to pay a higher price for global brands. Leveraging barcode-level consumption data, we show that domestic firms grow more through the growth of surviving goods than through new goods, and that domestic goods have a slower and longer life-cycle than foreign goods. We also show that the new customers of older domestic goods are poorer than the customers of new domestic products, a pattern that does not exist for foreign products. We rationalize these findings using a model of product choice under uncertainty. The possibility of learning through others slows down the most price-sensitive customers from buying a new product, driving down firms' profits. Lastly, we document the mechanisms behind this model by showing evidence of learning, evidence of the importance of the uncertainty margin, and evidence of the relevance of price-sensitivity.

JEL: D22; F23; L25

Keywords: quality uncertainty; international competition; consumer goods

*Corresponding author: Louise Guillouët, Columbia University Department of Economics (email: louise.guillouet@columbia.edu). Louise is very grateful to Eric Verhoogen, Amit Khandelwal and Jack Willis for their continuous encouragement and guidance throughout this project. We wish to thank Vittorio Bassi, Michael Best, Laura Boudreau, Gautam Gowrisankaran, Jonas Hjort, Rocco Macchiavello, David Martimort, Monica Morlacco, Suanna Oh, Golvine de Rochambeau, Matthieu Teachout, Pietro Tebaldi and seminar participants at the development and IO colloquia and workshops at Columbia for helpful feedback. Louise gratefully acknowledges the financial support of the Program for Economic Research and the Center for Economic Development and Policy at Columbia University. We would also like to thank ITAM staff for their key guidance in accessing the data and manipulating it on ITAM's Beethoven server. The data agreement details are registered with the Columbia University Institutional Review Board as the protocol AAAT7865. The views expressed herein are those of the authors and do not necessarily reflect the views of Kantar, Columbia University or ITAM. All errors are our own.

1 Introduction

A prevalent feature in developing countries is the absence or the lack of enforcement of quality regulation, which means customers cannot be certain they are buying a high-quality product. As a result, it may be difficult for firms based in developing countries to find customers who value their products, even if they actually are of high quality (see Verhoogen (2020) for a review of the literature on barriers to firm-level upgrading). One strategy around this issue is to build a reputation for quality over time (Shapiro (1983)).

Globalization increases the size of the market firms potentially have access to [1]. However, as argued by Goldberg and Reed (2020), ultimately firm sales must also come from the middle-class consumers at home. Moreover, globalization also increases the set of competitors firms must contend with at home. In developing countries, multinational corporations (henceforth MNCs) headquartered in high-income countries dominate many product markets. It may then be difficult for small, yet unknown domestic firms to grow a reputation over time when competing with long-established, high-reputation MNCs[2]. In this paper, we ask how important uncertainty about product quality is in explaining the size and growth patterns of domestic firms.

We study this question in the context of the Mexican consumer packaged goods sector. This is an ideal setting to study this question for several reasons. First, quality is salient in the food sector as its absence can lead people to get sick. Yet, the United States' Food and Drug Administrations often recalls product that were exported from Mexico to the United States, a lower bound for the number of (potentially unreported) issues with product that did not leave the domestic market[3]. The Mexican authority for food safety, Cofepris (Comision Federal para la Proteccion contra Riesgos Sanitarios), regularly updates the official norms food manufactures must abide by in an effort to curb quality issues. Second, in Mexico, although 42% of the population lives under the national poverty line, and global brands charge on average a 20% price premium, the aggregate market share of foreign firms is far above 50% for most product categories. This suggests that there are consumers who are willing to pay a higher price for high-quality products, a potential source of growth for high-quality domestic firms if they could harness this demand. Last, conditional on having several products, Mexican firms release less new goods than foreign firms proportionally to their size, a possible consequence of the lack of incentive or ability to invest in product innovation, further reinforcing the problem of lack of quality.

We answer this question by leveraging a rich barcode-level dataset covering the universe of consumer-packaged goods in Mexico from 2010 to 2015. We start by establishing five novel facts based on observing the sales of the 3985 firms selling goods to the households in the dataset. First, most of the growth of domestic firms comes from the growth of the sales of surviving goods, as opposed to the introduction of new goods. At the product level, when foreign firms introduce new goods, sales grow for a short period after the introduction of the good and then

[1] Some recent papers have proposed strategies to give firms access to this additional demand. For example, Atkin et al. (2017a) match rug makers in Egypt to importers in high-income countries.

[2] In the firm-to-firm sector, the presence of large buyers may help small domestic firms as shown by Alfaro-Ureña et al. (2019). Hjort et al. (2020) teach firms how to answer tender calls for large firms or governments.

[3] For a list of recent examples, see this article by digital media platform Sin Embargo and this public health article.

decline for a long period, as demand is cannibalized by newer products of the same firm or "stolen" by other firms. This has been shown in the context of the United States by Argente et al. (2019). When Mexican firms introduce new goods, they start off by selling considerably less than comparable new foreign products. But, conditional on surviving, sales grow and stay higher than in the initial quarter for up to two years. This suggests that Mexican firms are able to retain demand and attract new demand for their products as they age.

For both types of firms the key driver of sales growth is growth in the number of customers, as opposed to growth in the number of units sold to each customer or growth in the sales generated per unit. This fact helps us interpret the rapid demand depletion observed for foreign products as customers who were buying the goods when they are first released, being attracted to newer products by the same firm or by other firms as the products age. By contrast, the sustained demand observed for domestic products reflects either the persistence of the customers who first started buying the goods or the arrival of new customers purchasing the goods as they age. We further show that in order to grow their customer base, domestic firms depend relatively more on the intensive margin of product markets: growing the number of customers they sell each product to, as opposed to the extensive margin: growing the number of products they sell. This suggests that the first-order problem faced by domestic firms is not that their products do not match the taste of domestic customers, but potentially that they must overcome barriers to convince customers that they will appreciate each of their existing products.

Last, we show that the new customers of domestic products that have survived several quarters are poorer than the new customers buying domestic products who have just been released. By contrast, the new customers of foreign products that have survived several quarters are not different than the customers of these same foreign products when they were new. Together with the second fact about the domestic product life-cycle, this finding suggests that customers who face stricter constraints in their consumption decisions do not buy new domestic products immediately, but instead wait until these products survive a certain age before purchasing them.

We next propose a model of consumer decision that generates equilibrium outcomes matching these facts. In this model, consumers face uncertainty about product quality. Consumers who have a tighter budget constraint are less likely to experiment with a product of unknown quality, conditional on price. If it is possible to learn from others, there is a positive option value of waiting until other, less constrained individuals experiment with the product and reveal whether they liked it or not. This generates a delay in product adoption, hurting firms' profits.

We test the implications of the model in three different ways. First, we show evidence of individual learning. We do this by measuring the importance of brand experience in explaining consumers' decisions about purchasing products they haven't tried before. We show that individual exposure to a brand is highly predictive of future purchases of goods from that brand. Moreover, this effect is higher for domestic brands than global brands, despite domestic brands being less successful on average. Second, we show evidence of the uncertainty margin. We do this by exploiting the heterogeneity in product types. For products for which quality is more salient, such as infant formula, prior experience with a brand has a larger effect on the probability to purchase more products. Third, we show that learning is more important for individuals who are more budget constrained: the predictive power of brand exposure is much higher for households in the bottom half of the expenditure distribution.

Section 2 describes the literature this paper aims to contribute to. Section 3 describes the data and the setting. Section 4 discusses the five stylized facts we establish. Section 5 explains the conceptual framework we use to think about our results. Section 6 shows evidence of the mechanisms at work. Section 7 concludes.

2 Literature

The main contribution of this paper is to assess the role of uncertainty about product quality in accounting for the specific demand-side constraints faced by domestic firms in developing countries. This relates to a large literature studying asymmetric information and quality provision issues. Problems arise whenever sellers have more information than buyers, as identified by Akerlof (1970)): both low-quality and high-quality products will be pooled under the same price.

When there is uncertainty on the price that will be obtained, Sandmo (1971) shows that sellers provide less quality, worsening the problem. This paper is related to this theory as one possible consequence of the lack of market valuation for new domestic products is that domestics get discouraged and produce fewer varieties or varieties of lower quality. Shapiro (1983) shows that if consumers learn about quality over time, it is possible to sustain different qualities in the market, sold at different prices which reflect the cost of producing quality. Importantly, the sellers of high quality goods will charge more than their marginal cost, a premium which is proportional to the size of the informational friction and the time it takes for them to establish their reputation as a producer of high quality.

An important factor is therefore whether the sellers believe they can ultimately succeed. If the sellers know the buyers well and regularly contract with each other, "relational" contracts, without the intervention of a third party, can sustain high-value trade (Macchiavello and Morjaria (2015), Macchiavello and Morjaria (2019)). If buyers and sellers are far apart and are likely to never see each other again, it is much more costly for buyers to be certain about the quality of products (Startz (2016)). When relational contracts are not possible and verification is not possible, low-quality equilibria can be sustained until an outside intervention. In rural Uganda for example, misperceptions about anti-malarial drugs are widespread, allowing for the flourishing of low-quality counterfeit drugs. Bjorkman Nyqvist et al. (2012) experimentally introduce a low-price, high-quality anti-malarial drug promoted by an NGO, which ends up replacing the low-quality counterfeit drug prevalent at baseline. In most situations, however, higher-quality products are also more expensive, which can make it difficult for firms to convince customers that they are making the right decision.

Firms may convince customers about the quality of their products through marketing efforts. National or international brands and chains offer an alternative to direct relationship. Bennett and Yin (2019) study how a "high-productivity" pharmacy chain in India, marketing itself as a high-quality firm and consistently delivering high-quality medicine, lead to improved quality and cheaper prices at incumbent pharmacies. Bronnenberg et al. (2015) study the brand premium effect in over-the-counter drugs and grocery staples in the United States. Even in this presumably high-trust and high-transparency setting, the authors find that more informed shoppers (such as pharmacists for drugs, and chefs for groceries) are less likely to buy the branded product, suggesting that a sizable share of the brand premium is due to a lack of information on the demand side.

Using survey data on Pakistani soccer ball manufacturers, Atkin et al. (2017b) show that that the firms who charge the highest markups are not necessarily the most productive ones, but the ones that make the most marketing efforts, for example by participating in international trade fairs. However, marketing technology can sometimes be too costly. Bai (2018) tests experimentally whether a laser could serve as a hard signal to separate "high-quality" watermelons from "low-quality" watermelons in open-air markets in China. She finds that although it functions, once she removes the subsidy for the laser all the firms revert back to the pooling equilibrium because the price premium is not high enough to cover the cost of the technology. In a randomized controlled trial, Hjort et al. (2020) show that teaching firms marketing skills can expand the market they have access to, in particular towards large buyers, and thus enhance growth possibilities.

The trade literature has looked at the implications of the impact of marketing efforts on firms' ability to sell. Arkolakis (2010)' seminal paper shows that the convex cost of reaching additional customers in a given market can explain the puzzle that despite fixed costs to exporting in an additional country, many firms export small volumes in each destination. Hottman et al. (2016) show that it is a firm's "appeal", and not marginal cost, that drives the majority of differences between large and small firms, which is in part driven by high-appeal firms' ability to charge higher markups. Afrouzi et al. (2020) combine the same scanner data with cost data to show that this appeal is affected by spending on advertising and other non-production efforts. In a model of endogenous markup, they show that this spending may increase efficiency in the economy as it directs customers to the most productive firms. The empirical results of Einav et al. (2021), based on the visa credit card dataset, also support the importance of the customer margin. However, they suggest that if we assume constant markup, because firms must find customers, they spend rare resources on marketing, potentially diverting these resources from R&D efforts which could grow the economy more in the long term. This paper suggests that marketing spending may help support an industrial policy aiming to help the domestic sector grow.

We contribute to understanding the life-cycle of products. Argente et al. (2019) study the product life-cycle of consumer goods. They show that product turnover is high and firms must constantly reinvent their product scope in order to avoid business stealing, even though this strategy increases cannibalization. Perla (2019) proposes an alternative model to explain these life-cycles. His central idea is that customers may be "aware" of some products and not others. This means that firms have more market power than what the nominal number of competitors suggests. As products age though, customers learn about their existence through social networks.

The internet may introduce cheaper alternatives for firms to market their products (or put in another way, for consumers to learn about products' existence). Chen and Wu (2020) show that information frictions matter on ecommerce platforms and that online certification tools, while imperfect, may help SMEs sell. For example, they show that a seller displaying an extra "star" in their rating, a rounding effect, increases sales by 32% even when controlling for the true rating. They also show that information frictions increase with geographic and cultural distance between sellers and buyers. While e-commerce is still out of reach for many small firms in developing countries, it is an avenue for growth. In this paper, we study a market yet relatively untouched by the Internet, but which will be increasingly so: in 2020, 25% of Mexican consumers have bought groceries online up from 13% in 2017 and less than 3% in 2014[4], increasing the value of online-specific marketing efforts. Perla (2019) argues that although the targeted advertising that becomes available with the internet may increase the quality of the match between customers' tastes and firms, it would also further increase their market power.

Last, this paper contributes to understanding how trade, through the presence of MNCs, affects the welfare of consumers in developing countries. Fajgelbaum and Khandelwal (2016) show that trade is generally pro-poor because it decreases the price of goods that poor consumers spend a relatively higher share of their budget on. However, using consumer survey data from India, in which consumption is observed at a more detailed level, Atkin (2013) shows that because preferences change slowly, trade-induced decreases in prices may not increase poor consumers' calorie intake as much as one would expect. Using barcode-equivalent data from Mexico, Atkin et al. (2018) further show that the arrival of Walmart in new areas is relatively more beneficial for higher-income consumers, who demand the high-quality, high-price goods that Walmart supplies. In this paper, we have access the precise origin of the firms supplying the goods consumed, which helps us understand how to design pro-poor trade policies.

Atkin and Donaldson (2015) have access to information about the origin of goods but only for a dozen goods across three countries. This exercise allows them to show that intra-national trade costs may explain why consumers in remote areas may have less access to trade-induced price decreases than others. In our study, we have access to the universe of consumer-packaged goods consumed in Mexico and are therefore able to study how the presence of foreign firms influences the demand faced by domestic firms.

3 Data and Setting

In this section, we first introduce the two primary data sources that together allow us to study in detail the consumer goods sector in Mexico from January 2010 to December 2015. After describing each source in detail, we highlight the relevant features of the market.

3.1 Data

The main source of data is a rotating household panel shared by Kantar World Panel[5]. Households are visited twice a week to obtain a complete consumption diary about all of the packaged

[4] OECD, ICT Access and Usage by Households and Individuals

[5] Kantar World Panel is an international company that operates in more than 50 countries. They specialize in the collection of household consumption data for marketing and sales strategy purposes. For more information on the data-sharing agreement, see Aguilar et al. (2021).

goods purchased for at-home consumption. We observe[6] 7182 distinct households per month on average and a total of 15750 unique households from all 32 states of Mexico[7]. The sample is designed to represent metropolitan areas (collection of municipalities) in Mexico with more than 50,000 individuals.

The panel provides information on all packaged goods purchases made over time. The data covers all consumer packaged goods except liquor and tobacco. We mostly don't observe products purchased in bulk or by weight such as fresh fruits and vegetables, meat and fish, etc. Products that are described as being purchased in bulk (such as tortillas, a staple item for Mexican households) or for which the manufacturer is not identified are dropped[8]. For each purchase, we observe the transaction date, several characteristics, the price, the units purchased, the type of the store where the purchase was made[9], whether the product was subject to special promotions, and the payment method. Importantly for the rest of the paper, we also observe the name of the manufacturer of the good. For soft drinks purchases for example, the product characteristics would include whether the drink is "diet", the flavor, the content size and the package type (e.g. can or plastic bottle). It would then indicate "Coca-Cola FEMSA" if the item was produced by the Coca-Cola subsidiary in Mexico.

The panel also contains economic, demographic and geographic information about each household. We observe these variables at the yearly level. They include information about household members' age, gender and occupation. We also observe asset-like characteristics about households: we know whether they have a fridge, a TV, and other appliances. We observe a few dwelling characteristics. Last, we observe a socioeconomic status (SES), which is computed based on households' assets, the dwelling characteristics, the head of household education and purchasing power, as computed by the Mexican Association of Market Intelligence Agencies (AMAI)[10]. We observe five values for the SES. Geographic variables include the neighborhood of residence.

The second source of data is the yearly updated directory of private establishments (Directorio Estadístico Nacional de Unidades Económicas or DENUE) conducted by the Mexican national statistical institute, INEGI. DENUE was first created in 2010 based on the 2009 Economic Census. Since then, DENUE has been used as a sampling frame for business surveys[11]. Although this dataset provides the exact addresses of all 5478689 establishments listed, the KWP dataset only lists one of the names of the firm, which may be multi-establishment. We therefore merge this name with either the given or the official name of an establishment in the administrative dataset, but cannot tell which establishment actually corresponds to each good. This dataset is important because it helps us narrow down our definition of the "firms" we study in this paper.

[6] Given the observations we exclude because the purchases are not identified by their manufacturer.

[7] On average, households stay in the panel for 3.4 years. We observe 1191 households on all 72 months.

[8] For this reason, we are unable to do the analysis on "unbranded goods" as unbranded goods are not identified by their manufacturer.

[9] We observe the name of the retail chain if the purchase was made at a retail chain. If not, we observe whether the purchase was made in a store that does not belong to a chain or in an open-air market.

[10] See AMAI'swebsite for more explanation

[11] See a commentary from the American Statistical Association

The third source of data is the yearly updated register of foreign investment (Registro Nacional de Inversiones Extranjeras or RNIE) which is maintained by the Mexican Economic Secretary (Secretaria de Economia). Firms based in Mexico that receive foreign capital directly (as opposed to through stock) are listed there. There are 65810 firms in total. We match the name of the manufacturers collected by Kantar with the names of firms listed in this list, which defines their status as a foreign firm and in the case of the latter gives us the origin of the foreign investment.

3.2 Descriptive Statistics

Households In Table B.1, we compare the households in our dataset in 2010 to the households surveyed in the expenditure and income survey of households, ENIGH, conducted by the national statistical office, INEGI, in 2010. Households in the Kantar data appear to be slightly larger than households in the ENIGH data and slightly better off in terms of assets, but overall they are comparable.

Goods We observe 66059 different products over 82 categories. For each product, we observe between one and seven characteristics such as brand, flavor, color, size, number of units in the package etc. This description is similarly precise as the one available in better known datasets such as Nielsen which report at the barcode level. In the baseline specification of the empirical analysis, we will refer to these goods defined by Kantar under a product identifier as "barcodes".

Firms These barcodes are manufactured by 3985 different firms, 94% of which receive no foreign direct investment. The few firms who do are much larger as shown in Figure A.1, which plots the distribution of annual total expenditure panelists made on each firm. Among foreign firms, 49% receive FDI from the USA[12]. The rest of the foreign investment is from 30 different countries, which limits our ability to do separate analysis by country of origin.

Foreign firms also charge much higher prices and enjoy larger market shares than domestic firms. To show this, we run the following regressions within each product category:

$$y_{i,g,t} = \alpha + \beta \text{Foreign}_i + \zeta_{gt} + \mu_i + \epsilon_{i,g,t} \tag{1}$$

where we regress barcode i's price y observed in city g in month t on a dummy for whether the manufacturer is foreign or not and control for a set of city and time fixed effects ζ_{gt}. In the baseline specification, we control for the product category that barcode i belongs to μ_i. We show two other specifications, one controlling for the subcategory of product (when mentioned) and one controlling for this and the size of the package sold. Figure 1 shows the Foreign coefficients β obtained for the largest 20 product categories, based on what characteristics are included as controls. In the baseline specification we only control for product category and city and time fixed effects, while in the second regression we control for a narrower product category and in the third regression we use city and time fixed effects, product subcategory, and size of the unit. Almost all the coefficients are positive, which we interpret as the "Foreign price premium". The Foreign price premium can be very large, up to 100% for the milk powder category, which is

[12]This is similar to the share of US investment in the total foreign investment inflows received by Mexico over the last 20 years. Authors' calculation based on the data published by the Secretaria de Economia.

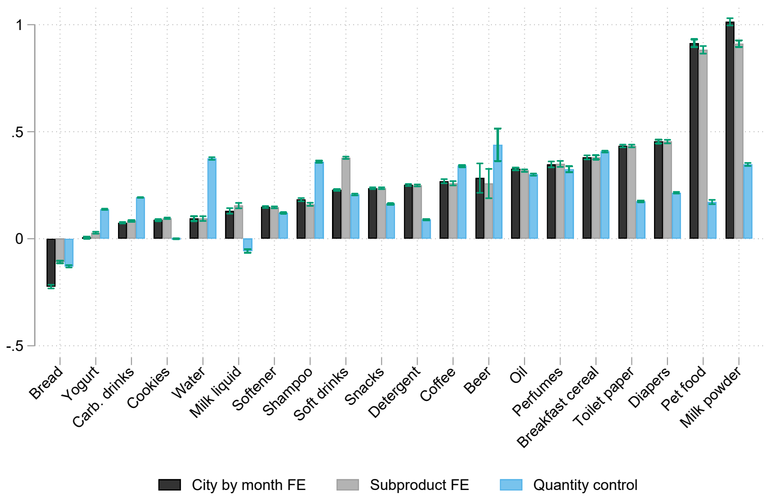

Figure 1: Foreign price premium

Notes: Figure reports the coefficient obtained from the purchase-level regression where a dummy that turns on if the product is sold by a foreign firm and characteristics fixed effects are regressed on the log price of the product. We repeat this regression three times for each of the top 20 categories in the dataset. Each regression uses incrementally additional characteristics as controls, which are described in the legend (grey and blue coefficients are obtained from regressions that also have city-by-month and and subproduct fixed effects, respectively). 95% confidence intervals are reported using the green bars. Products are sorted according to the size of the coefficient obtained in the first regression.

due both to the subsidized price of the domestic products sold by public establishments such as LICONSA and the importance of quality of products such as infant formula.

We run a similar regression for market shares in each product category:

$$y_{k,t} = \alpha + \beta \text{Foreign}_k + \vartheta_t + \epsilon_{k,t} \tag{2}$$

except now i is a firm and t is a year. We look at market shares defined in terms of sales, volume, and quantity (typically leveraging the information given by KWP about the content of each unit, usually in milliliters or grams depending on the nature of the product - since here the analysis is within product categories, we don't think there's a risk of bias). Figure 2 shows the Foreign coefficients β obtained for the largest 20 product categories. All but four coefficients are positive, which we interpret as the "Foreign share premium". We point out the correlation between the negative coefficients observed for the sectors of milk, beer and bread with the existence of three very large Mexican firms in these sectors (Lala, Modelo and Bimbo respectively which are themselves MNCs, based in Mexico). The one for milk is further linked to the importance of a public establishment, LICONSA[13], which sells milk to 18% of urban households. Further, we remark that the negative share premia for bread, milk and beer are coincidental with the smaller price premia for these product categories, emphasizing the idea that in the categories for which there is no domestic powerhouse that is trusted, consumers turn to foreign firms which have proven their ability to deliver quality in other markets.

New goods A large part of the empirical analysis relies on the identification of "new" goods in the dataset. In this paper, the definition of new products is empirical. We define a

[13] See Jiménez-Hernández and Seira (2021) for an assessment of government's role in milk provision.

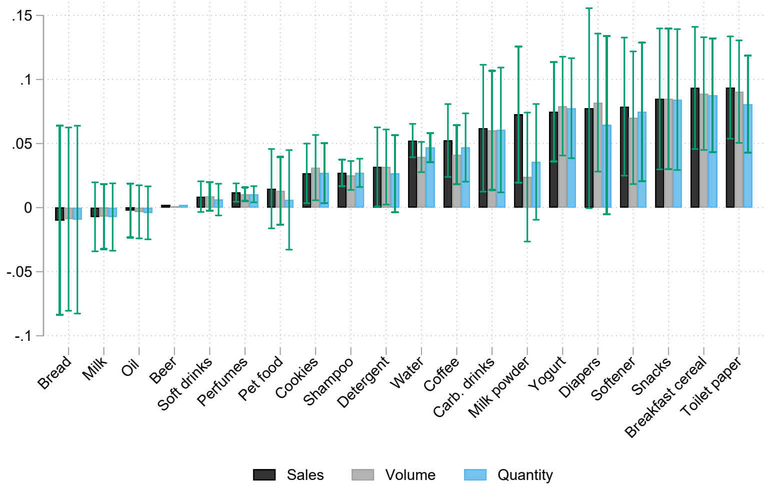

Figure 2: Foreign share premium

Notes: Figure reports the coefficient obtained from the firm-product-category-level regression where a dummy that turns on if the firm is foreign and year fixed effects are regressed on the market share of the firm in that product category. We repeat this regression three times for each of the top 20 categories in the dataset. Each regression uses a different definition of the market share, which are described in the legend. 95% confidence intervals are reported using the green bars. Products are sorted according to the size of the coefficient obtained in the first regression.

product as new if it appears in the dataset more than a year after the start of the dataset (so in or later than January 2011) and if it first appears as being consumed by a household who has been active in the dataset for more than a year. The rationale behind the latter part of the definition is that otherwise we could misclassify products as new when they are just rare, and think they are new to the entire Mexican consumer packaged goods market when they existed before and we are only observing them now thanks to the arrival of a new household in the dataset. We validate this data-driven definition by checking that new products referenced in the marketing firm The Market Think's review[14] were indeed classified as new according to our definition. Although most products do not appear in our dataset, the ones that do were indeed classified as new. Examples include Coca-Cola's Life drink and Nestle's Oikos Greek-style yogurt. On average, 22% of firms' product portfolio in a given year (starting in 2011) are classified as new. When looking at the rate of introduction of new goods, domestic firms appear to introduce slightly more new goods than foreign firms as shown in Columns (1)-(3) of Table 1.

However, this result may be driven by the size heterogeneity between Mexican firms and MNCs: firms that do not have a product that is not new have a 100% new products rate. Therefore, in Columns (4)-(6) of Table 1 we focus on firms that sell at least one good that is not new in th year of observation. We find that the rate of introduction of new goods is smaller, at only 16% for foreign firms on average, and is smaller for Mexican firms, by about 3.7 percentage points. This foreshadows the implications of our model: if certain firms face difficulties finding customers for the same underlying level of quality, they will have less ability to invest in quality and less incentive to do so, reinforcing the prevalence of low quality products in the market.

[14]See https://www.themarkethink.com/lanzamiento-de-productos/

	Share new products					
	(1)	(2)	(3)	(4)	(5)	(6)
Mexican	0.046	0.035	0.041	-0.046	-0.047	-0.037
	(0.012)	(0.013)	(0.013)	(0.007)	(0.007)	(0.007)
Firm sales, mMXN		-0.025	0.017		-0.004	-0.003
		(0.006)	(0.009)		(0.003)	(0.005)
Number of old varieties			-0.000			0.000
			(0.000)			(0.000)
Firm leader in category			-0.062			-0.013
			(0.034)			(0.017)
Category FEs	No	No	Yes	No	No	Yes
Baseline share (foreign)	0.22	0.22	0.22	0.16	0.16	0.16
N	12127	12127	12126	10008	10008	10008
R2	0.06	0.06	0.12	0.01	0.01	0.06

Table 1: Firm-level new goods introduction rate

Notes: Table reports the firm-year level regression of a dummy for whether the firm is Mexican, the firm sales in million MXN, the number of not new varieties, firm fixed effects and other firm-level characteristics described in the table, on the share of new products among the total number of products the firm has in a given year. The first three columns show all firms we have data for. Columns (4), (5) and (6) only show firms that sell at least one not new good in the year of observation. Standard errors are reported in parenthesis.

4 Stylized Facts

In this section, we leverage the rich data about consumer goods and establish a series of five novel facts about the dynamics of demand faced by domestic and multinational companies in Mexico.

4.1 Domestic firms grow relatively more through surviving goods

We follow Argente et al. (2019) who use scanner data to study the life-cycle of products in the United States. We decompose a firm's growth rate into the sum of a "new products" component and a "product life-cycle" component. We use the following approximation:

$$\Delta S_{k,t} = \underbrace{\Delta S_{k,t}^{old,survive} - \overline{S}_{k,t-1}^{old,exit}}_{product\ life-cycle} + \underbrace{n_{k,t}^{new} \times \overline{s}_{k,t}^{new}}_{new\ products} \tag{3}$$

where $\Delta S_{k,t}$ is the firm's annual growth rate, and where the "product life-cycle" component is further decomposed in the annual growth rate of the aggregate sales of products that survived between year $t-1$ and year t, $\Delta S_{k,t}^{old,survive}$, from which we subtract the share of sales that the products that exited between year $t-1$ and t represented in the sales of the firm in year $t-1$, $\overline{S}_{k,t-1}^{old,exit}$. The "new products" component is the product of the rate of introduction of new products in the firm's portfolio $n_{k,t}^{new}$ and the relative sales of new products compared to older products $\overline{s}_{k,t}^{new}$, obtained by taking the ratio of the average sales of a new product in year t to the average sales of a surviving product in year t. We show Figure 3 that domestic firms

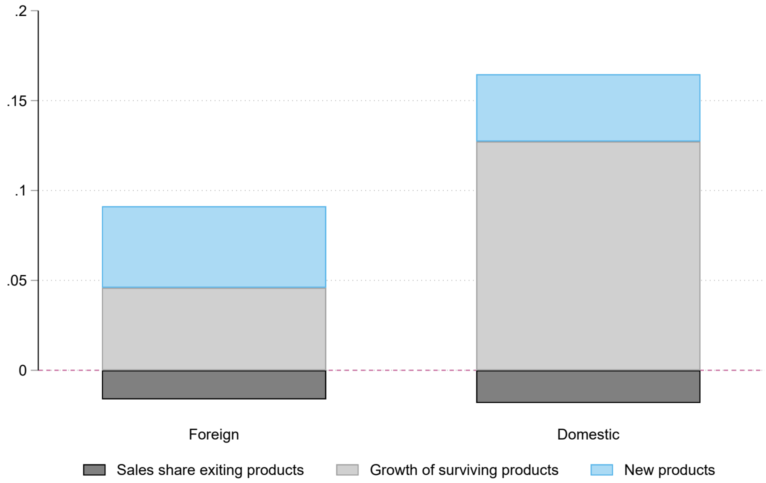

Figure 3: The extensive customer margin is key to firm growth

Notes: Figure represents the average firm-level year-to-year growth components, separated out by the origin of the firms. The sales of exiting products is the average share of sales that the product that exited between the previous year and the current year represented in the previous year. The growth of surviving products is the average growth rate of the total sales of products that survived between the previous year and the current year. The new products components is the product of the average rate of introduction of new products and the average sales of new products compared to older products.

grow more than foreign firms. For foreign firms, growth appears to be driven equally by the introduction of new products and the growth. The "new products" component is comparable across foreign and domestic firms. The difference in average growth between foreign firms and domestic firms is almost entirely driven by the higher growth rate of surviving products for the latter.

This may be surprising given the fact underlined in Table 1 that domestic firms have a higher rate of introduction of new goods, an element of the "new products" components illustrated in Equation (3). However, we also show in that table that this is driven mostly by very small firms who only sell one product and mechanically appear to sell 100% of new products in the year when they change their product.

When making comparisons between domestic firms and foreign firms, one may be concerned that the difference we find is entirely driven by size differences. It resonates in this context since we know that foreign firms are much larger. However, this finding is robust when focusing on the domestic firms that are most comparable in size to foreign firms, as shown in Appendix Figure A.2. This is not surprising given the observation that relatively larger domestic firms have a smaller rate of introduction of new goods than smaller domestic firms, reinforcing.

4.2 Domestic products have a slower life-cycle

We again follow Argente et al. (2019) and analyze the evolution of product-level sales over time. We estimate the following equation:

$$\log y_{i,t} = \alpha + \sum_{a=1}^{14} \beta_a D_a + \lambda_{jt} + \theta_c + \epsilon_{i,t} \tag{4}$$

where i is a barcode observed in a certain quarter t, a is a potential age of the good in quarters (between 0 and 14, the mean age attained by products born between 2011Q1 and 2012Q2 and the maximum age reached by products born in 2012Q2 in the dataset). j is a product category and c is a cohort-quarter. We regress a good u's log sales in a quarter t on dummies for the age of this product, product category interacted with quarter fixed effects, and cohort fixed effects (corrected following Deaton (1997)'s suggestion to avoid collinearity). We restrict the analysis to products born between 2011Q1 and 2012Q2, as the mean duration of a product is 14 quarters, which is the period that can be observed with products born in 2012Q2 until 2015Q4. We only keep products which sold a positive amount in each quarter of their "life". We perform this regression separately for domestic and foreign products.

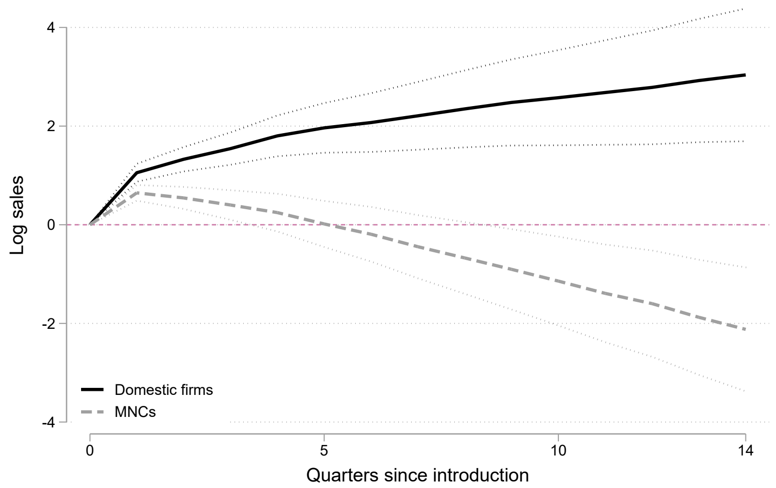

Figure 4: Product life-cycle

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters from a product-quarter level regression where we regress these dummies, quarter by product category fixed effects, and cohort fixed effects on the sales of a product in a given quarter. We run two separate regressions for domestic and foreign products, respectively. 95% confidence intervals are represented by the lighter lines.

In Figure 4, we plot the coefficients obtained on each of the 14 quarter-age dummies. The coefficients obtained for the products introduced by foreign firms are very similar to the ones obtained by Argente et al. (2019) in the United States scanner data: product are born and sell, they exit immediately or sell more, and then their sales decline for a long time until they exit. The authors attribute this to a constant arrival of products in the market, which means that after a while demand for the aging products is captured by newer products released by the same firm (cannibalization) or other firms' products (business stealing). By contrast, when domestic firms introduce new products, sales increase in the initial quarters and then remain high for a long period. While this may seem like an advantage for domestic products, they sell much less at any age, so overall the lifetime total sales of the average domestic product are much lower than the lifetime total sales of the average foreign product. Moreover for a given level of total lifetime sales, displacing sales later in time means that the present discounted values of domestic products is smaller.

A potential explanation for this phenomenon could be the endogenous timing of product of retirement: if foreign firms decide to retire products later than domestic firms conditional on a sales trajectory, we would observe that conditional on surviving a set number of quarters, foreign average product sales decrease over time. Since the product duration is the same between foreign and domestic goods (14 quarters on average, 15 quarters median) it seems unlikely. We can further rule out this hypothesis by showing that the survival rate of foreign products is not different from that one of domestic products as they age, as shown in Appendix Figure A.3.

4.3 The extensive customer margin is key to firm growth

We now decompose firm sales in a different manner. Here, we follow Einav et al. (2021) in studying the exact decomposition of a firm's sales:

$$\text{Sales} \equiv \text{Customers} \times \frac{\text{Quantity}}{\text{Customers}} \times \underbrace{\frac{\text{Sales}}{\text{Quantity}}}_{\text{Unit value}}$$

This equation can be interpreted in the following way: in order to double sales, firms may do one of three things, *ceteris paribus*

(i) double their number of customers - the extensive margin

(ii) double the number of items they sell to each customer - one intensive margin

(iii) double the monetary value obtained from each unit sold - another intensive margin

The firm upgrading literature, and particularly the technology adoption literature, has mostly focused on (iii). The intuitive argument is that if a firm increases the quality of its output, it will be able to sell each unit for a higher price, and probably attract more demand. Here, we remain agnostic and look at which margin seems to generate the most variation in sales. Of course, these three variables are endogenous as highlighted by the sentence above and we are not making any causal claim. We begin by taking logs of each element:

$$\log(\text{Sales}) = \log(\text{Customers}) + \log(\text{Quantity per Customer}) + \log(\text{Unit value})$$

We perform three regressions, regressing each element in turn on the log of Sales.

$$\log(\text{Customers}_{kt}) = \alpha + \beta_C \log(\text{Sales})_{kt} + \eta_k + \vartheta_t + \epsilon_{kt} \tag{5a}$$

$$\log(\text{Quantity per Customer}_{kt}) = \alpha + \beta_Q \log(\text{Sales})_{kt} + \eta_k + \vartheta_t + \epsilon_{kt} \tag{5b}$$

$$\log(\text{Unit value}_{kt}) = \alpha + \beta_U \log(\text{Sales})_{kt} + \eta_k + \vartheta_t + \epsilon_{kt} \tag{5c}$$

where i is a firm and t is a year. We control for firm fixed effects (η_k) and year fixed effects (ϑ_t) which means the coefficients are identified from looking at the years when firms grew faster, or slower, than their own average. By construction,

$$\beta_C + \beta_Q + \beta_U \equiv 1$$

The results are presented in Figure 5. The customer margin accounts for 60% of sales variation

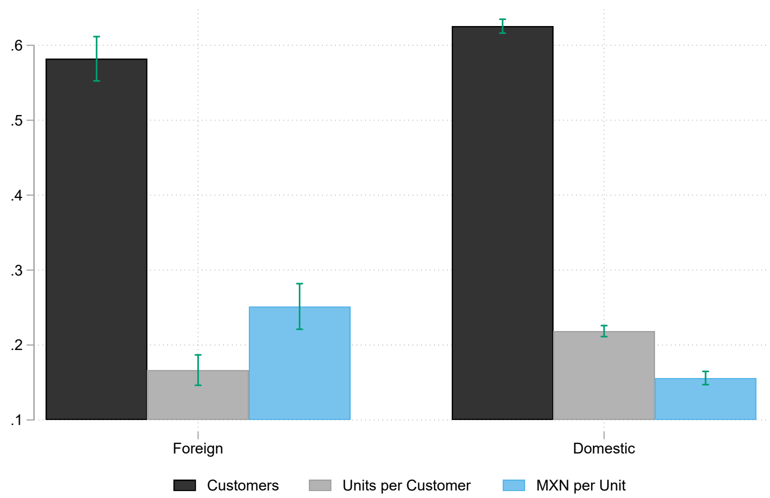

Figure 5: The extensive customer margin is by far the most important

Notes: Figure represents the coefficients obtained on log sales from firm-year-level regressions of log sales, product category and year fixed effects on log customers, log units per customer and log monetary value per unit. We run each of the three regressions separately for foreign and domestic firms. 95% confidence intervals are represented using the bars. Table B.4 shows the numerical values of these coefficients and an alternative specification.

within firms over time. This is true for both domestic and foreign firms, it appears that for domestic firms, the customer margin seems slightly more important, while the value margin seems relatively less important than for foreign firms. Relating this finding to the last fact about the product life-cycle, we interpret the rapid decay of sales of products introduced by foreign firms as they age as customers being attracted to new products in the initial quarters of the life of a product, and customers being increasingly attracted away from these products and towards newer products from the same firm, or other or newer products from other firms. By contrast, the sustained sales of domestic products even after two years could suggest that even if early customers grow tired of these products and stop buying them, more customers arrive to replace them as the products age.

This fact may vary a lot by industry. Because the regressions control for firm fixed effects, effectively comparing years in which firms grow faster than their own average, we are not excessively worried about industry variation. Nevertheless, we study this by taking the analysis to the firm-product category level. Appendix Figure A.4 shows that across the top 20 categories, the customer margin is the most important one in 19 categories, the only exception being perfume for which it makes sense that the price obtained per purchase would matter a lot. This finding confirms that this analysis is capturing an important dimension of how the firms we are studying grow.

4.4 The intensive customer per product margin is key to customer growth

Because the extensive customer margin appears to be so important according to the previous analysis, we turn to study the question of how to acquire more customers. To double their number of customers, firms may firms may do one of three things, *ceteris paribus*

(i) double the number of markets they operate in - the extensive margin

(ii) double the number of customers they reach in each market - the intensive margin

Markets can be understood in several different ways: in terms of geography, distribution channels, or product. The firm upgrading literature has mostly focused on the extensive margin. The intuition behind is that firms face a barrier in accessing new markets, such as exports (Atkin et al. (2017a)) or large, tender-based markets (Hjort et al. (2020)). A few recent papers have looked at the question of product scope expansion. Intuitively, it means that if there are customers who do not like the current products that the firm has in its portfolio, but may like slightly different products, the firm may increase sales by iterating on its current varieties. Product scope expansion might be easier for firms than quality upgrading which often means adopting new technologies or techniques. However, product scope expansion still imposes some fixed cost on the firm. We look at the following exact decomposition:

$$\text{Customers} \equiv \text{Products} \times \frac{\text{Customers}}{\text{Products}}$$

which yields, taking logs

$$\log(\text{Customers}) = \log(\text{Products}) + \log(\text{Customers per product})$$

We perform two regressions, regressing each element in turn on the log of the number of customers:

$$\log(\text{Products})_{kt} = \alpha + \beta_M \log(\text{Customers})_{kt} + \eta_k + \vartheta_t + \epsilon_{kt} \tag{6a}$$

$$\log(\text{Customers per product})_{kt} = \alpha + \beta_C \log(\text{Customers})_{kt} + \eta_k + \vartheta_t + \epsilon_{kt} \tag{6b}$$

where η_k are firm fixed effects and ϑ_t are year fixed effects. By construction

$$\beta_M + \beta_C \equiv 1$$

Figure 6 shows the coefficients obtained for domestic and foreign firms. For the latter, the two margins do not appear to be extremely different. By contrast, for domestic firms the intensive margin accounts for over 70% of the variation in the number of customers, while the extensive margin accounts for less than 30%. This suggests that it is easier for domestic firms to grow by convincing more customers to buy their existing products, than to grow by adding additional products to their portfolio. This is coherent with the first two facts showing that for domestic firms, the growth of the sales of surviving goods is a strong determinant of firm-level

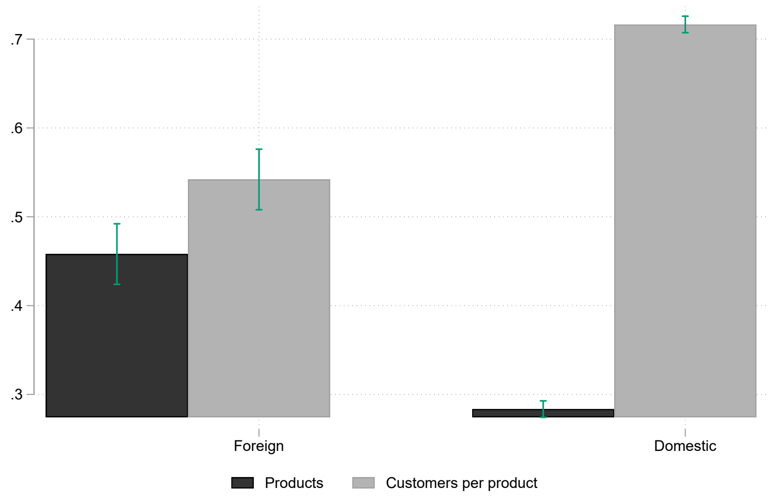

Figure 6: The intensive customer per product margin matters more for domestic firms

Notes: Figure represents the coefficients obtained on log customers from firm-year-level regressions of log customers, product category and year fixed effects on log products and log customers per product. We run each regression separately for foreign and domestic firms. 95% confidence intervals are represented using the bars.

sales growth, and that the products introduced by domestic firms acquire customers over time. Moreover, we showed in Section 3 that when we exclude small domestic firms that only carry one good, domestic firms have a lower rate of introduction of new products than foreign firms on average. This may be a strategic response to the difficulty they face in introducing new products.

4.5 The new customers of older domestic products are poorer

The last fact sheds light on the characteristics of the new customers acquired by products as they age. We estimate the following equation:

$$\log y_{i,u,t} = \alpha + \sum_{a=1} \beta_a D_a + \lambda_{jt} + \theta_c + \delta_g + u_{i,u,t} \qquad (7)$$

where i is a good observed in a certain quarter t, u is the individual who purchased it, a is a potential age of the good in quarters (between 0 and 14, the mean age attained by products born between 2011Q1 and 2012Q2 and the maximum age reached by products born in 2012Q2 in the dataset). j is a product category and c is a cohort-quarter. We regress the annual expenditure of new customers u of good i's in quarter t on dummies for the age of this product, product category interacted with quarter fixed effects, and cohort fixed effects (corrected following Deaton (1997)'s suggestion to avoid collinearity). We restrict the analysis to products born between 2011Q1 and 2012Q2, as the mean duration of a product is 14 quarters, which is the period that can be observed with products born in 2012Q2 (they are effectively 14 quarters old in 2015Q4). We perform this regression separately for domestic and foreign products. We control for city fixed effects δ_g based on where customer u lives.

Figure 7 shows the coefficients obtained on each of the 14 quarter-age dumies. The new

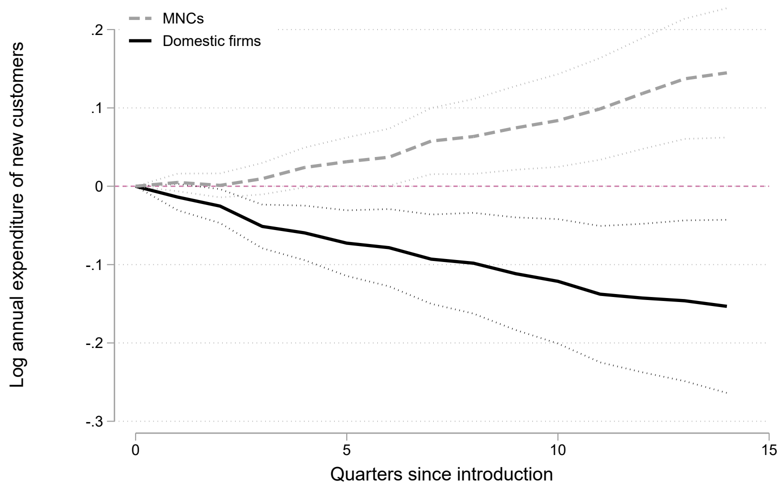

Figure 7: The new customers of older domestic products are poorer

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters from an individual-product-quarter level regression described in Equation (7) where we regress these dummies, quarter by product category fixed effects, cohort fixed effects and city fixed-effects on the average annual expenditure of the new customers of a product in a given quarter. We run two separate regressions for domestic and foreign products, respectively. 95% confidence intervals are represented by the lighter lines. Table B.6 shows the numerical values.

customers who start buying a foreign product as it ages are not different from the customers who started buying the foreign product in the very first quarters of its existence. By contrast, the new customers who start buying a domestic product as it ages are significantly poorer than the customers who started consuming the same product in its initial quarters of existence. This fact suggests that the process by which domestic firms acquire demand is very different from the one foreign firms go through.

One may think that there is a supply-side explanation to this pattern. For example, if domestic products are distributed through certain stores and foreign products are distributed through other stores, it may be that there the distribution dynamics specific to each type of firm account for the divergence the new customers profile over time. Specifically, if domestic products are better able to go to more remote, and poorer areas, then the new customers they are getting are not arriving later because of learning but because the products were not available to them before. We are not too concerned about this because the regression includes city fixed effects, which suggests that this finding holds within cities.

We pursue this concern by asking whether the new cities that are reached by products as they age are different depending on whether the product is sold by a domestic firm (we would expect the cities to be increasingly small and rural) or by a foreign firm (we would expect the products to only diffuse to big cities). To answer this question, we implement a similar estimation equation as shown in Equation (7) but at the barcode-city-quarter level, where we only keep the new cities where the barcode makes sales in each quarter after it is born. The results are shown in Appendix Table B.7. We find that the new cities that products reach when they are older are smaller in population (Columns (1) and (2)) and less dense (Columns (3) and (4)) than the cities reached when the products are new. However, this trend is common to both domestic and foreign goods. We find that the new cities that domestic product reach when they are older are further away from Mexico City, the capital, than the cities by the same products when they were new. This trend does not exist for foreign products: in fact, it appears to be almost the opposite. However, the coefficients for both set of products are very noisy, so it does not appear to be a solid threat to out interpretation of fact # 5.

5 Model

In this section we introduce a conceptual framework in which we highlight the hypothesis that uncertainty about product quality generates these facts.

5.1 Setup

In this model, agents are consumers faced with a new good of unknown quality x. x is a random variable from a known distribution with prior mean μ_0. Agents choose whether to purchase the good or not, maximizing

$$u(\mu) = \max\{\mu - \beta_i p, 0\}$$

where β_i represents price-sensitivity. For simplicity, we suppose there are only two possible types of agents: $\beta_i \in \{\beta_L, \beta_H\}$ where $\beta_H > \beta_L$. There is a higher share of the market who is of the second type $\gamma_H = 1 - \gamma_L$. This game is dynamic. In each period, agents decide whether to purchase the good or not. If they purchase the good, they immediately learn the true quality x. They can then use this information to decide whether they will purchase the good in the next period. We call this learning through individual experimentation "individual learning". Agents buy at most three times in a row (three times if they like the good, one time if they don't). Agents may also learn through social observation: by looking at whether people who have tried the good continue to purchase it or not, they can update their belief about the quality of the good. We call this "social learning".

5.2 Individual learning

Suppose that a period t no agent has purchased the good yet. Everyone has the same prior μ_t. Each agent decides whether

$$\mu_t - \beta_i p \lesseqgtr 0$$

If one agent decides to purchase the good and not the other, it must mean that it is the agent with the lesser price-sensitivity β_L. She immediately learns the true quality x. However, the other agents don't learn anything from observing this initial purchase. $\mu_{t+1} = \mu_t$.

5.3 Social learning

In the following period $t + 1$, the leader assesses whether

$$x - \beta_L p \lesseqgtr 0$$

Upon observing this decision, the other agents learn whether

$$x \lesseqgtr \beta_L p$$

and can update their belief accordingly:

$$\mu_{t+2} = \begin{cases} < \mu_{t+1} & \text{if } x < \beta_L p \\ \geq \mu_{t+1} & \text{if } x \geq \beta_L p \end{cases}$$

5.4 Sales trajectories

This model generates the following sales trajectory for a "successful" product $x > \beta_H p$, when $\mu_0 < x$ ("uncertainty") and for when $\mu = x$ ("no uncertainty"), where each point represents the fixed effects of the product aging one quarter on its sales compared to the initial quarter. Figure 8 strongly resembles Figure 4 which plots the coefficient on the dummies for the age in

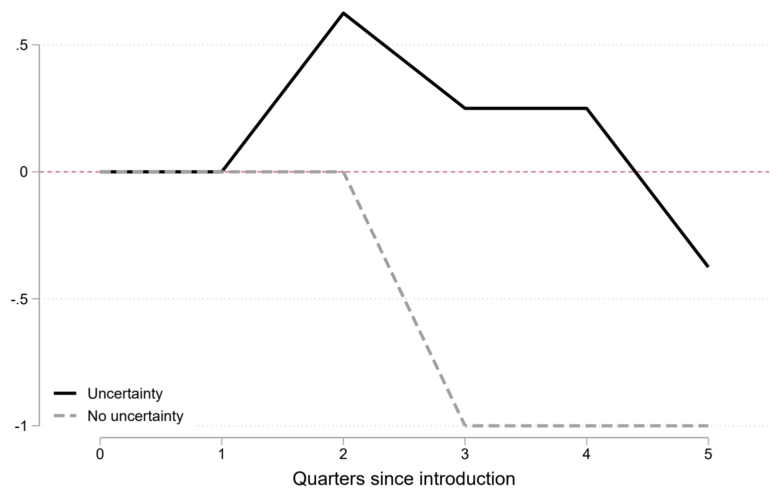

Figure 8: Sales trajectories of a successful product given uncertainty

Notes: Figure represents the simulated coefficient on the dummies for age of the product in quarters one would get by running a product-quarter level regression of these dummies on sales of a successful product in the model.

quarters of the products in the data.

6 Mechanisms

In this section, we document the mechanisms described in the model. First, we show evidence of learning. Second, we show that uncertainty about product quality binds by showing that the learning phenomenon is more important for goods for which quality uncertainty is more salient, or where the lack of quality might be more costly. Third and last, we show that price-sensitivity plays an important role in how learning affects consumption.

6.1 Learning

We propose to shed a light on learning by measuring the influence of brands. Conceptually, this means that we envision brands as a signal about products' qualities. If consumers respond to these signals, it suggests they are learning from them. For example, suppose consumers are worried about the presence of *Salmonella* in food[15]. They might be hesitant to buy new products. However, if they have experience with a brand and have never gotten sick with products from that brand, they will be more willing to try new products from that brand than new products from another brand.

We measure the informational effect of brands by looking at the probability that consumers purchase goods in a given year, depending on whether they have experience with the brand in the previous year. In order to avoid measuring just habit formation, we only look at the probability of purchasing products that the consumer did not purchase in the previous year. The experience with the brand therefore comes from other products. We benchmark this "experience from other goods" effect against the "same-good experience" effect of having consumed a good on the probability to consume it again in the current year. We hypothesize that because domestic products are not seen to be as trustworthy as foreign products *ex ante*, a positive experience with a domestic brand (a domestic good) will have more effect on the probability to buy other goods from that brand (the same good) in the future.

We estimate the following equation:

$$y_{u,ik,t} = \alpha + \beta y_{uk,t-1} + \gamma D_k + \delta D_k \times y_{uk,t-1} + \xi_u + \epsilon_{u,i,t} \tag{8}$$

where u is the consumer, i is the barcode, k is the brand the barcode belongs to, and t is the year. The parameter of interest is β, which measures the effect of the consumer's individual experience with the brand j at $t-1$, $y_{uk,t-1}$ (excluding the particular barcode i we study) on subsequent consumption (at t) of other barcodes from the same brand. The second parameter of interest is the interaction coefficient δ, which measures the differential impact of experience with a brand for a domestic brand compared to a foreign brand.

We cannot claim that experience with a brand in year $t-1$ is exogenous, as individuals' exposure to a brand, both in terms of advertising and choice sets, may be strategically chosen by forward-looking firms in year $t-1$ by firms who anticipate consumers' propensity to consume in year t. However, we can control for individual fixed effects ξ_u, and so we are arguably measuring the average influence of past brand exposure on current consumption choices, controlling for individual preferences, advertising exposure and choice sets.

The results are described in Table 2. In Column (1) we find that previous experience with a brand is predicted to raise the probability of consumption of a barcode non previously consumed by 0.02, *ceteris paribus*, a huge effect compared to the baseline probability of 0.007. This suggests that consumers are inferring from their previous exposure to brands that they might like other products from that brand. As a benchmark, we show in Column (2) the effect of having consumed the exact same good before (Column 2). At 0.326, the "same-good experience" effect is 15 times higher, raising the probability to consume by 0.3 up from 0.013.

[15] https://www.cdc.gov/salmonella/oranienburg-09-21/index.html

This effect could be attributed both to learning about one's preference for the barcode itself, or consumer inertia. The fact that we find an effect when looking at "other-good experience" suggests there is some learning from the brand signal.

The idea that these effects can be attributed to learning and not just habit formation or consumer inertia are confirmed by Columns (3) and (4). In Column (3), we show that previous consumption of other products from the same firm raises the probability to consume a given good not previously consumed by 0.012, a smaller but comparable effect to the brand. Column (4) shows that previous consumption of other products from the same country raises the probability to consume a given good not previously consumed by 0.007, a yet smaller effect. This makes sense because the name of a firm is less salient to consumers than the name of the brand they purchase goods from, so they learn less from the signal. Similarly, while national reputations matter as underlined by Cagé and Rouzet (2015) and Bai et al. (2019), we expect their signal to be less strong than that of firms or brands. Further, the probability to consume

	Current consumption			
	Brand	Barcode	Firm	Country
	(1)	(2)	(3)	(4)
Previous consumption	0.018	0.327	0.012	0.007
	(0.000)	(0.000)	(0.000)	(0.000)
Domestic	-0.002	-0.004	-0.002	-0.004
	(0.000)	(0.000)	(0.000)	(0.000)
Previous consumption X	0.007	0.070	0.009	0.000
Domestic	(0.000)	(0.000)	(0.000)	(.)
Hhd FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Control mean	0.008	0.015	0.006	0.008
N hhd	1640	1640	1640	1640
N	30123860	30686647	30123860	30123860
R2	0.01	0.12	0.01	0.00

Table 2: There is relatively more learning for Mexican products

Notes: Table shows the results of a barcode-household-year level regression of a dummy for consumption in the previous year, a dummy that turns on if the product if Mexican, and the interaction of these two dummies, on a dummy that turns on if the individual has consumed the barcode in the current year. In Column 1, previous-year consumption is defined as previous consumption of the brand that the product belongs to, while we only look at current consumption of products that were not consumed before. In Columns 3 and 4 we do the same but for the firm and country that the product belongs to, respectively. Column 2 looks at previous consumption of the product itself and therefore includes all products, whether consumed or not in the current year. We always control for household and year fixed effects. We work on a random sample of households for computing power reasons. Standard errors are reported in parenthesis.

a domestic product not consumed before is lower by 0.002 (28%), compared to the probability to consume a foreign product not consumed before. However, we find that the coefficient interacting previous brand exposure with the dummy indicating that the brand is Mexican is positive, economically significant at 38% of the value of the brand experience coefficient, and statistically significant. This suggests that households learn more (update more their belief) from consuming domestic products than from consuming foreign products. The fact that the former coefficient is negative suggests that households' prior is lower for domestic products, but we cannot distinguish between a lower or a noisier prior (or both).

6.2 Importance of quality certainty

The previous subsection suggests that there is some consumer learning. However, many forms of learning are possible: learning about how to use products from a certain brand for example, or learning through social exposure that the product is socially valuable. In the model, learning comes from uncertainty about product quality. We therefore turn to showing that quality matters.

The importance of quality may affect how people learn about product quality. For example, if food products are of low quality, it may imply that household members get sick. On the contrary, if paper products are of low quality, it has less dramatic implications. Therefore, our model implies that consumers would take more precaution in buying new food products than in buying new paper products. Further, they would learn more from consuming good brands of food products than from consuming good brands of paper products.

We test this by estimating the following equation:

$$y_{u,ik,t} = \alpha + \beta y_{uk,t-1} + \gamma D_i + \delta D_i \times y_{uk,t-1} + \xi_u + \epsilon_{u,i,t} \tag{9}$$

which is similar to Equation 8: $y_{u,ik,t}$ is current consumption of product i (brand k) by household u and time t, $y_{uk,t-1}$ is previous consumption by household u of brand k excluding barcode i, but importantly the dummy D_i represents whether the quality of that product is salient or not.

We start by comparing food to non-food products. Column (1) of Table 3 shows the results. Previous exposure to a brand is predicted to increase the consumption of any product by 0.017 *ceteris paribus*, similar to what we showed in Table 2. This effect increases by 0.005 or almost a third when the product is a food product as opposed to a non-food product, confirming that when quality is more important, consumers learn more from the brands they consume.

However, food and non-food products are different not only with respect to the importance of quality, but also in terms of the structure of demand, which may influence learning. For example, households typically buy food every week or even several times a week, while they probably buy toilet paper once a month or even less frequently. This would make learning slower for the latter products, regardless of the importance of quality. Moreover, food and non-food products are different in terms of supply chains and marketing strategies, which may influence learning as well. So finding different learning speeds between food and non-food products might be attributable to these factors and not to quality salience.

Therefore, we exploit the heterogeneity of quality salience among more narrow product categories. Most people would agree that the safety of food being fed to babies and young children is relatively more important than the safety of other foods. In Column 2, we therefore compare the effect of brand exposure on the probability of buying an infant formula products to the effect of brand exposure on the probability of buying another milk product. We find that the effect of exposure to the brand is much larger for infant formula products: it increases the probability of buying a product from the same brand by 0.041, an effect that is twice

	Current consumption		
	All	Milk	Pads
	(1)	(2)	(3)
Previous consumption	0.017	0.025	0.014
	(0.000)	(0.001)	(0.001)
Salient category	0.000	-0.001	-0.004
	(0.000)	(0.001)	(0.001)
Previous consumption X Salient	0.005	0.041	0.002
	(0.000)	(0.006)	(0.002)
Hhd FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Control mean	0.008	0.009	0.009
N	6141346	252782	175652
R2	0.01	0.01	0.01

Table 3: Learning by quality salience

Notes: Table shows the household-barcode-year-level regression of a dummy indicating whether the household has consumed products for a similar brand in the previous year, whether the product belongs to a category for which quality is salient, and an interaction of these two dummies, on a dummy indicating whether the household has consumed the barcode in the observation year. In the first column, we look at all products and the salient category is food. In the second column, we keep only milk-products and the salient category is infant formula. In the third category, we keep baby diapers and sanitary pads and the salient category is baby diapers. We always control for household and year fixed effects. Standard errors are reported in parenthesis.

larger than the standalone "other-products" brand effect at 0.025. Similarly, in Column (3) we compare the effect of brand exposure on the probability of buying baby diapers to the effect of brand exposure on the probability of buying adult sanitary pads. We again find that the brand effect is larger for baby products. This analysis therefore suggests that quality matters in the way consumers learn about previous experience when choosing goods they have not consumed before.

6.3 Price sensitivity

In the previous two subsections, we show that consumers seem to be learning, or updating their beliefs about goods they previously did not consume, from the signals generated by consuming from brands. They seem to be learning relatively more from consuming Mexican brands, suggesting that these brands have *ex ante* lower or noisier priors. We further show that they seem to be learning relatively more for products for which quality is more salient, confirming our hypothesis that concerns about quality drive this learning process. We now turn to the last element of our hypothesis: that income is a binding constraint in choosing to experiment with unknown goods. We propose to estimate the following equation:

$$y_{u,ik,t} = \alpha + \beta y_{uk,t-1} + \gamma D_u + \delta D_u \times y_{uk,t-1} + \mu_i + \epsilon_{u,i,t} \tag{10}$$

Which again estimates the impact of $y_{uk,t-1}$, previous consumption of brand k, on $y_{u,ik,t}$, the current consumption of product i (brand k) by household i and time t, but importantly the dummy D_u represents whether the households' expenditure in year $t-1$ puts them in the bottom half of the expenditure distribution in the sample (which we will henceforth call the "low-expenditure" group. We cannot control for household fixed effects in this context, and therefore control for barcode fixed effects μ_i instead.

	Current consumption			
	Brand	Barcode	Firm	Country
	(1)	(2)	(3)	(4)
Previous consumption	0.020	0.290	0.017	0.011
	(0.000)	(0.000)	(0.000)	(0.000)
Low expenditure	-0.004	-0.005	-0.004	-0.007
	(0.001)	(0.001)	(0.001)	(0.003)
Previous consumption X Low expenditure	0.005	0.107	0.002	0.005
	(0.001)	(0.003)	(0.001)	(0.003)
Barcode FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Control mean	0.007	0.013	0.005	0.008
N	6141346	6259248	6141346	6141346
R2	0.05	0.17	0.05	0.05

Table 4: There is relatively more learning among the poor

Notes: Table shows regression of a dummy for consumption in the previous year, a dummy that turns on if the household is in the bottom half of the expenditure distribution, and the interaction of these two dummies, on a dummy that turns on if the individual has consumed the barcode in the current year. In Column 1, previous-year consumption is defined as previous consumption of the brand that the product belongs to, while we only look at current consumption of products that were not consumed before. In Columns 3 and 4 we do the same but for the firm and country that the product belongs to, respectively. Column 2 looks at previous consumption of the product itself and therefore includes all products, whether consumed or not in the current year. We always control for household and year fixed effects. Standard errors are reported in parenthesis.

Table 4 shows the results of the estimation in the same order as in Table 2. We find that having consumed for a brand is predicted to raise the probability to purchase a previously not-purchased product by a large, positive and significant amount, *ceteris paribus*. This effect is considerably larger for low-expenditure households. Focusing on Column 1, we find that the effect is 25% larger for low-expenditure households. As a benchmark, we show in Column 2 the effect of having consumed the exact same good, which is again about 15 times larger than the brand effect, is much larger for low-expenditure households. Last, Columns 3 and 4 show robustness checks looking at larger groups: firm and country, and the results are similar. It therefore appears that sensitivity to price is an important predictor of the important of learning among consumers.

7 Conclusion

Uncertainty about product quality may prevent consumers, particular consumers for whom it is costly to experiment, from purchasing products they don't know. This fundamental stickiness creates a demand barrier for young and small firms to grow. In this paper, we measure the importance of this issue by leveraging barcode-level dataset covering the universe of consumer packaged goods available in Mexico between 2010 and 2015. In this setting, uncertainty about product quality is an important issue and it creates a large advantage to global brands, who dominate the market despite charging higher prices.

We show five new stylized facts: domestic firms grow relatively more through surviving goods as opposed to new goods than foreign firms. Domestic products sales start lower but grow more and remain higher than initially, for a longer period than foreign products. Sales growth in the Mexican consumer goods sector is largely driven by the customer extensive margin. While this appears to be true for both domestic and foreign firms, for domestic firms customer growth is driven more by the intensive customer acquisition margin, within each product, as opposed to the extensive market acquisition margin, by adding products to the portfolio. Last, the new customers acquired by domestic products as they age are poorer than the initial customers who started purchasing the domestic products when it just came out, while the new customers of foreign products as they age are not different from the later customers.

We rationalize these five facts in a stylized model showing that the presence of uncertainty about product quality leads price-sensitive customers to withhold from purchasing a new good. Instead, they prefer to wait and learn from others, hurting firms' profit. Lastly, we provide evidence of the learning mechanisms at play. We show that individual learning matters by showing that a consumers' probability to consumer a given barcode from a brand increases with previous exposure to the brand, excluding the barcode itself, controlling for individual taste. This learning effect is much stronger for goods for which uncertainty is more salient, and it is higher for consumers in the bottom half of the expenditure distribution.

The findings in this paper suggest that uncertainty about product quality result in a lower demand for domestic products than otherwise. In future work, we plan to quantitatively estimate the size of the inefficiency, which will help us think about potential business strategies and policy interventions. Potential avenues for domestic firms to raise demand are informative advertising and dynamic pricing. Policy wise, interventions that raise minimum quality standards or transparency could also help develop the domestic sector.

References

Afrouzi, Hassan, Andres Drenik, and Ryan Kim, "Growing by the Masses: Revisiting the Link between Firm Size and Market Power," *Available at SSRN 3703244*, 2020.

Aguilar, Arturo, Emilio Gutierrez, and Enrique Seira, "The effectiveness of sin food taxes: Evidence from Mexico," *Journal of Health Economics*, 2021, *77*, 102455.

Akerlof, George A., "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Alfaro-Ureña, Alonso, Isabela Manelici, and Jose P Vasquez, "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages," *Available at SSRN 3376129*, 2019.

Argente, David, Munseob Lee, and Sara Moreira, "The Life Cycle of Products: Evidence and Implications," *Available at SSRN 3163195*, 2019.

Arkolakis, Costas, "Market penetration costs and the new consumers margin in international trade," *Journal of political economy*, 2010, *118* (6), 1151–1199.

Atkin, David, "Trade, tastes, and nutrition in India," *American economic review*, 2013, *103* (5), 1629–63.

__, Amit K Khandelwal, and Adam Osman, "Exporting and firm performance: Evidence from a randomized experiment," *The Quarterly Journal of Economics*, 2017, *132* (2), 551–615.

— **and Dave Donaldson**, "Who's getting globalized? The size and implications of intra-national trade costs," Technical Report, National Bureau of Economic Research 2015.

— **, Azam Chaudhry, Shamyla Chaudry, Amit K Khandelwal, and Eric Verhoogen**, "Organizational barriers to technology adoption: Evidence from soccer-ball producers in Pakistan," *The Quarterly Journal of Economics*, 2017, *132* (3), 1101–1164.

— **, Benjamin Faber, and Marco Gonzalez-Navarro**, "Retail globalization and household welfare: Evidence from mexico," *Journal of Political Economy*, 2018, *126* (1), 1–73.

Bai, Jie, "Melons as lemons: Asymmetric information, consumer learning and quality provision," Technical Report, Working paper 2018.

— **, Ludovica Gazze, and Yukun Wang**, "Collective reputation in trade: Evidence from the Chinese dairy industry," Technical Report, National Bureau of Economic Research 2019.

Bennett, Daniel and Wesley Yin, "The market for high-quality medicine: Retail chain entry and drug quality in India," *Review of Economics and Statistics*, 2019, *101* (1), 76–90.

Bronnenberg, Bart J, Jean-Pierre Dubé, Matthew Gentzkow, and Jesse M Shapiro, "Do pharmacists buy Bayer? Informed shoppers and the brand premium," *The Quarterly Journal of Economics*, 2015, *130* (4), 1669–1726.

Cagé, Julia and Dorothée Rouzet, "Improving "national brands": Reputation for quality and export promotion strategies," *Journal of International Economics*, 2015, *95* (2), 274–290.

Chen, Maggie X and Min Wu, "The value of reputation in trade: Evidence from alibaba," *Review of Economics and Statistics*, 2020, pp. 1–45.

Deaton, Angus, *The analysis of household surveys: a microeconometric approach to development policy*, World Bank Publications, 1997.

Einav, Liran, Peter Klenow, Jonathan D Levin, and Raviv Murciano-Goroff, "Customers and retail growth," Technical Report, Working Paper 2021.

Fajgelbaum, Pablo D and Amit K Khandelwal, "Measuring the unequal gains from trade," *The Quarterly Journal of Economics*, 2016, *131* (3), 1113–1180.

Goldberg, Pinelopi K and Tristan Reed, "Income Distribution, International Integration, and Sustained Poverty Reduction," Technical Report, National Bureau of Economic Research 2020.

Hjort, Jonas, Vinayak Iyer, and Golvine De Rochambeau, "Informational Barriers to Market Access: Experimental Evidence from Liberian Firms," Technical Report, National Bureau of Economic Research 2020.

Hottman, Colin J, Stephen J Redding, and David E Weinstein, "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics*, 2016, *131* (3), 1291–1364.

Jiménez-Hernández, Diego and Enrique Seira, "Should the government sell you goods? Evidence from the milk market in Mexico," Technical Report, Stanford University Working Paper 2021.

Macchiavello, Rocco and Ameet Morjaria, "The value of relationships: evidence from a supply shock to Kenyan rose exports," *American Economic Review*, 2015, *105* (9), 2911–45.

— **and** — , "Competition and relational contracts in the Rwanda coffee chain," 2019.

Nyqvist, Martina Bjorkman, Jakob Svensson, and David Yanagizawa-Drott, "Can good products drive out bad? evidence from local markets for (fake?) antimalarial medicine in Uganda," 2012.

Perla, Jesse, "A model of product awareness and industry life cycles," *Unpublished Paper, University of British Columbia*, 2019.

Sandmo, Agnar, "On the theory of the competitive firm under price uncertainty," *The American Economic Review*, 1971, *61* (1), 65–73.

Shapiro, Carl, "Premiums for high quality products as returns to reputations," *The quarterly journal of economics*, 1983, *98* (4), 659–679.

Startz, Meredith, "The value of face-to-face: Search and contracting problems in Nigerian trade," *Available at SSRN 3096685*, 2016.

Verhoogen, Eric, "Firm-Level Upgrading in Developing Countries," *CDEP-CGEG Working Paper*, 2020, (83).

Appendix Figures

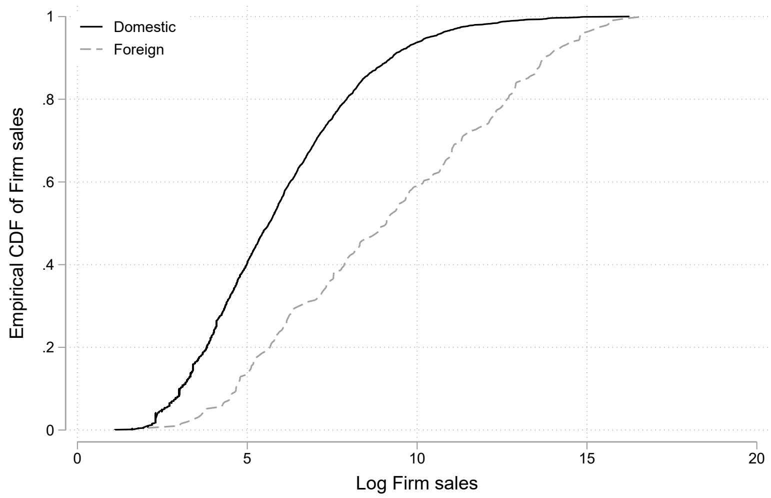

Figure A.1: Firm size distribution

Notes: Figure plots the distribution of the log sales of the firms observed in our dataset, where sales are defined as the total yearly expenditure of the households in that dataset on that firm. We separate out the distribution between foreign and domestic firms.

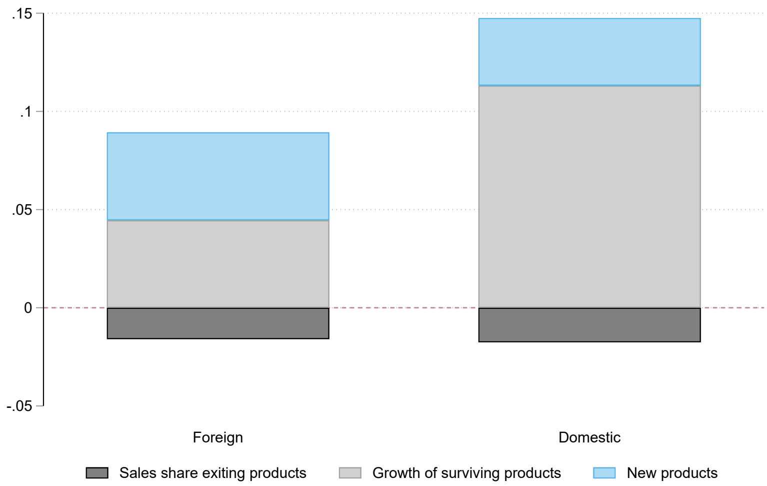

Figure A.2: The extensive customer margin is key to firm growth

Notes: Figure represents the average firm-level year-to-year growth components, separated out by the origin of the firms. We look only at firms belonging to the top quarter of the firm sales distribution of each year (almost all the foreign firms do). The sales of exiting products is the average share of sales that the product that exited between the previous year and the current year represented in the previous year. The growth of surviving products is the average growth rate of the total sales of products that survived between the previous year and the current year. The new products components is the product of the average rate of introduction of new products and the average sales of new products compared to older products.

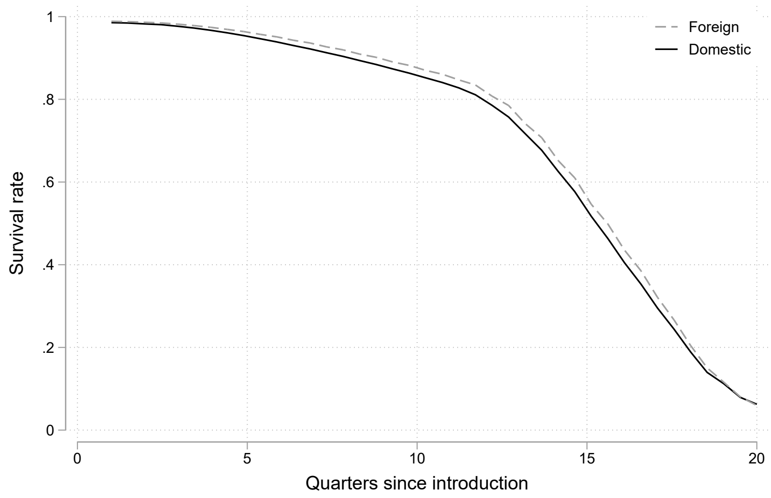

Figure A.3: Survival rate of foreign and domestic products, by quarter

Notes: Figure plots the survival rate of new products in the dataset over time, measured in quarters since the product was introduced. The death of a product is defined at the last quarter it appears in the dataset. We separate the survival rate by whether the product belons to a foreign or a domestic firms.

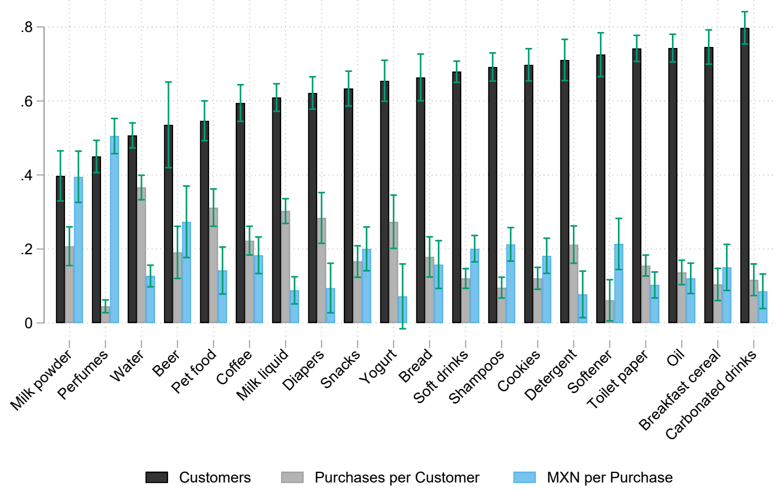

Figure A.4: Customer/Quantity/Value decomposition, by product category

Notes: Figure represents the coefficients obtained on log sales from firm-year-level regressions of log sales and year fixed effects on log customers, log purchases per customer and log monetary value per purchase. We run each of the three regressions separately for each of the top 20 product categories (one observation is the sales of a firm in a given a year in the category). 95% confidence intervals are represented using the bars.

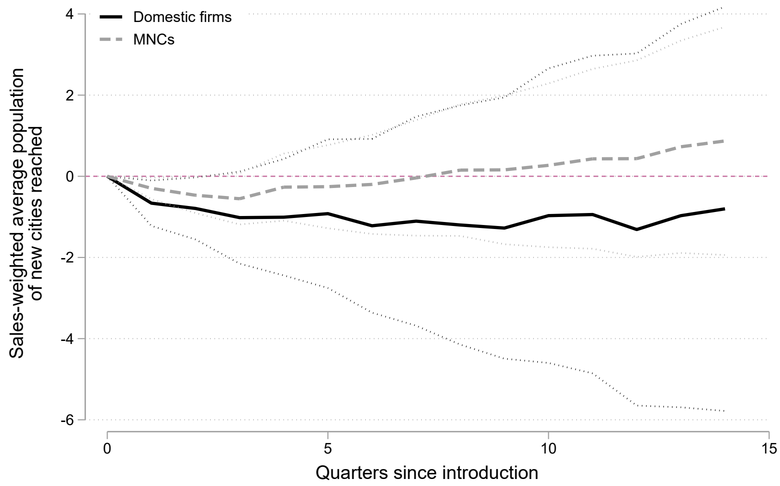

Figure A.5: New cities characteristics, by quarter and product type

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters from a product-quarter level regression where we regress these dummies, quarter by product category fixed effects, and cohort fixed effects on the sales-weighted average of the population of the new cities in which a product sells in a given quarter. We run two separate regressions for domestic and foreign products, respectively. 95% confidence intervals are represented by the lighter lines.

Appendix Tables

	ENIGH			KWP			Difference	
	mean	sd	N	mean	sd	N	diff	p
Number of household members	3.94	1.98	26942	4.37	1.83	8414	0.430	0.00
Number of women in household	2.03	1.27	26942	2.29	1.22	8414	0.267	0.00
Age head of household	48.32	15.62	26942	45.61	14.02	8412	-2.707	0.00
Finished primary	0.84	0.37	26942	0.96	0.20	8414	0.120	0.00
Finished secondary	0.35	0.48	26942	0.65	0.48	8414	0.307	0.00
Finished Post-secondary	0.26	0.44	26942	0.13	0.34	8414	-0.130	0.00
Works full time	0.75	0.44	26942	0.75	0.43	8414	0.006	0.24
Number of cars	0.53	0.80	26942	0.56	0.66	8414	0.030	0.00
Number of PCs	0.31	0.61	26942	0.33	0.47	8414	0.019	0.01
Access to Internet (0/1)	0.19	0.39	26942	0.24	0.42	8414	0.043	0.00
Number of color TVs	1.44	0.92	26942	1.87	0.98	8413	0.426	0.00
Number of fridges	0.83	0.43	26942	0.96	0.19	8412	0.135	0.00
Number of microwaves	0.42	0.51	26942	0.70	0.46	8414	0.287	0.00
Number of bedrooms	2.01	0.97	26385	2.20	0.97	8412	0.188	0.00
Debit or credit card (0/1)	0.21	0.41	26942	0.28	0.45	8414	0.070	0.00
Monthly expenditure (MXN)	1107.30	758.20	26942	1320.09	736.49	8414	212.796	0.00

Table B.1: Household-level summary statistics, KWP vs ENIGH in 2010

Notes: Table compares summary statistics of the main dataset used in the analysis (Kantar World Panel or KWP) in 2010 in Columns (4)-(6) against the official expenditure survey (ENIGH 2010) in Columns (1)-(3). ENIGH provides the national reference values for household characteristics, income and expenditures. When relevant, the variable described is measured for the head of household (adult man if two working-age adults are present). Ownership of a debit or credit card is a variable in the ENIGH survey and it is coded in the KWP to 1 if the household is ever reported to use a card as a mode of payment. Expenditure in the ENIGH survey cannot be compared exactly to expenditure in the KWP but is constructed based on similar categories (spending on personal care, household care, and food for at-home consumption).

	All	Mexican	Foreign
Growth sales	0.11	0.14	0.08
	(0.35)	(0.48)	(0.17)
Product Life Cycle Component	0.06	0.10	0.03
Growth of Surviving	0.08	0.12	0.04
	(0.32)	(0.45)	(0.17)
Sales Share of Exit	-0.02	-0.02	-0.02
	(0.04)	(0.05)	(0.04)
New Products Component	0.04	0.04	0.04
Entry Rate	0.16	0.16	0.16
	(0.12)	(0.15)	(0.09)
Entrants Relative Sales	0.26	0.24	0.27
	(0.28)	(0.31)	(0.26)
Obs	8885	8010	875

Table B.2: Firm growth decomposition

Notes: Table shows the results from the decomposition of annual growth of sales at the firm-year level, as defined in Equation (3). For each firm and year starting in 2011, we compute the contribution of new products the number of new products and their sales in their first year of activity. Table shows the sales-weighted average across all firms and years. The first column groups all firms, while the second and third column separate firms by whether they have received foreign investment ("Foreign") or not ("Mexican"). Standard errors are shown in parenthesis.

	Log sales		Log quantities		Log price	
	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic
	(1)	(2)	(3)	(4)	(5)	(6)
Age=2 quarters	0.659	1.051	0.541	0.969	0.158	0.055
	(0.079)	(0.092)	(0.080)	(0.094)	(0.037)	(0.038)
Age=3 quarters	0.558	1.319	0.322	1.162	0.308	0.085
	(0.111)	(0.125)	(0.111)	(0.127)	(0.052)	(0.051)
Age=4 quarters	0.422	1.543	0.063	1.278	0.446	0.121
	(0.149)	(0.166)	(0.149)	(0.168)	(0.070)	(0.068)
Age=5 quarters	0.276	1.796	-0.194	1.437	0.599	0.157
	(0.190)	(0.209)	(0.190)	(0.212)	(0.089)	(0.086)
Age=6 quarters	0.055	1.967	-0.528	1.498	0.732	0.203
	(0.232)	(0.255)	(0.232)	(0.257)	(0.109)	(0.104)
Age=7 quarters	-0.142	2.088	-0.848	1.514	0.877	0.238
	(0.274)	(0.301)	(0.275)	(0.304)	(0.129)	(0.123)
Age=8 quarters	-0.393	2.236	-1.209	1.546	1.016	0.283
	(0.317)	(0.348)	(0.318)	(0.351)	(0.149)	(0.142)
Age=9 quarters	-0.617	2.396	-1.554	1.605	1.171	0.316
	(0.360)	(0.395)	(0.361)	(0.398)	(0.169)	(0.161)
Age=10 quarters	-0.825	2.542	-1.888	1.625	1.305	0.364
	(0.404)	(0.442)	(0.404)	(0.446)	(0.189)	(0.180)
Age=11 quarters	-1.061	2.630	-2.227	1.630	1.447	0.404
	(0.446)	(0.488)	(0.447)	(0.493)	(0.210)	(0.199)
Age=12 quarters	-1.267	2.745	-2.544	1.621	1.566	0.461
	(0.490)	(0.536)	(0.491)	(0.541)	(0.230)	(0.219)
Age=13 quarters	-1.479	2.876	-2.858	1.613	1.723	0.500
	(0.534)	(0.584)	(0.535)	(0.589)	(0.251)	(0.238)
Age=14 quarters	-1.706	2.999	-3.228	1.646	1.843	0.569
	(0.578)	(0.632)	(0.579)	(0.638)	(0.271)	(0.258)
Age=15 quarters	-1.972	3.135	-3.583	1.677	2.012	0.597
	(0.623)	(0.681)	(0.623)	(0.687)	(0.292)	(0.278)
Product X Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	No	No	Yes	Yes	Yes	Yes
Initial quarter mean	5.47	4.71	5.47	4.71	5.47	4.71
N	16836	12424	16678	12173	16678	12173
R2	0.23	0.28	0.33	0.30	0.56	0.59

Table B.3: Barcode life-cycle: log sales

Notes: Table shows the results from the product-quarter-level regression shown in Equation (4) of dummies for the age of the product in quarters, product interacted with quarter fixed effects (or subproduct interacted with quarter fixed effects), cohort fixed effects on log sales (Columns (1) and (2)), log quantities (Columns (3) and (4)) or log price (Columns (5) and (6)). We only keep products that survived at least 14 quarters and compute their total sales or quantities, or sales-weighted price in the dataset in each quarter. Standard errors are shown in parenthesis.

	Within firms across time			Across firms within time		
	Customers	Items per C	MXN per item	Customers	Items per C	MXN per item
Mexican firms	0.626	0.176	0.198	0.736	0.200	0.064
	(0.005)	(0.003)	(0.005)	(0.003)	(0.002)	(0.003)
N	10934	10934	10934	12081	12081	12081
R2	0.97	0.90	0.86	0.81	0.43	0.03
	Within firms across time			Across firms within time		
	Customers	Items per C	MXN per item	Customers	Items per C	MXN per item
Foreign firms	0.582	0.141	0.277	0.795	0.205	0.000
	(0.015)	(0.008)	(0.015)	(0.007)	(0.005)	(0.008)
N	1126	1126	1126	1157	1157	1157
R2	0.99	0.97	0.90	0.91	0.61	0.00

Table B.4: Sales decomposition

Notes: Table shows the results of the firm-year-level decomposition of sales as explained in Equation (5a). On the left-hand side, we show the coefficients obtained on each firm-level regressions of the log of sales and firm and year fixed effects on the log number of customers, log number of items sold to each customer, and log monetary value of each item sold. The regressions are computed separately for Mexican and Foreign firms, respectively. We show these results graphically in Figure 5. On the right-hand side, we show the coefficients obtained from three similar regressions without firm fixed effects, which amounts to considering each observation as a separate firm and interpreting the coefficient as heterogeneity across firms. Standard errors are shown in parenthesis.

	Within firms across time		Across firms within time	
	Products	Customers per P	Products	Customers per P
Mexican firms	0.284	0.716	0.474	0.526
	(0.005)	(0.005)	(0.003)	(0.003)
N	11273	11273	12599	12599
R2	0.95	0.95	0.67	0.72
	Within firms across time		Across firms within time	
	Products	Customers per P	Products	Customers per P
Foreign firms	0.458	0.542	0.626	0.374
	(0.017)	(0.017)	(0.008)	(0.008)
N	1150	1150	1186	1186
R2	0.98	0.96	0.82	0.63

Table B.5: Customers decomposition

Notes: Table shows the results of the firm-year-level decomposition of the number of customers as explained in Equation (6a). On the left-hand side, we show the coefficients obtained on each firm-level regressions of the log of number of customers and firm and year fixed effects on the log number of products and log customers buying each product. The regressions are computed separately for Mexican and Foreign firms, respectively. We show these results graphically in Figure 6. On the right-hand side, we show the coefficients obtained from two similar regressions without firm fixed effects, which amounts to considering each observation as a separate firm and interpreting the coefficient as heterogeneity across firms. Standard errors are shown in parenthesis.

	Log expenditure				SES (1-6)	
	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic
	(1)	(2)	(3)	(4)	(5)	(6)
Age=2 quarters	0.013	-0.002	0.013	-0.011	0.093	-0.004
	(0.008)	(0.012)	(0.007)	(0.011)	(0.018)	(0.028)
Age=3 quarters	0.014	-0.010	0.013	-0.024	0.143	-0.040
	(0.010)	(0.015)	(0.010)	(0.014)	(0.024)	(0.034)
Age=4 quarters	0.027	-0.037	0.026	-0.061	0.208	-0.058
	(0.013)	(0.019)	(0.012)	(0.018)	(0.031)	(0.043)
Age=5 quarters	0.047	-0.038	0.048	-0.070	0.287	-0.066
	(0.017)	(0.023)	(0.016)	(0.022)	(0.039)	(0.053)
Age=6 quarters	0.064	-0.063	0.062	-0.092	0.336	-0.079
	(0.020)	(0.028)	(0.019)	(0.026)	(0.047)	(0.063)
Age=7 quarters	0.071	-0.072	0.074	-0.105	0.438	-0.113
	(0.024)	(0.033)	(0.022)	(0.031)	(0.055)	(0.074)
Age=8 quarters	0.096	-0.095	0.100	-0.132	0.495	-0.107
	(0.027)	(0.038)	(0.026)	(0.036)	(0.064)	(0.086)
Age=9 quarters	0.107	-0.112	0.111	-0.149	0.570	-0.125
	(0.031)	(0.042)	(0.029)	(0.040)	(0.073)	(0.097)
Age=10 quarters	0.123	-0.135	0.127	-0.171	0.676	-0.124
	(0.034)	(0.047)	(0.032)	(0.045)	(0.081)	(0.108)
Age=11 quarters	0.130	-0.149	0.140	-0.187	0.751	-0.164
	(0.038)	(0.052)	(0.036)	(0.050)	(0.090)	(0.119)
Age=12 quarters	0.149	-0.181	0.160	-0.221	0.828	-0.176
	(0.042)	(0.057)	(0.039)	(0.054)	(0.098)	(0.131)
Age=13 quarters	0.176	-0.189	0.185	-0.228	0.905	-0.188
	(0.046)	(0.062)	(0.043)	(0.059)	(0.107)	(0.142)
Age=14 quarters	0.197	-0.202	0.208	-0.239	0.999	-0.196
	(0.049)	(0.067)	(0.046)	(0.064)	(0.116)	(0.154)
Age=15 quarters	0.209	-0.203	0.220	-0.249	1.081	-0.162
	(0.053)	(0.073)	(0.050)	(0.069)	(0.125)	(0.166)
Product X Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
City FEs	No	No	Yes	Yes	Yes	Yes
Initial quarter mean	9.76	9.72	9.76	9.72	2.43	2.38
N	429795.00	210452.00	429795.00	210452.00	429795.00	210452.00
R2	0	0	0	0	0	0

Table B.6: Barcode life-cycle: characteristics of new customers

Notes: Tables represents the coefficients obtained on the dummies for the age of the product in quarters from a individual-product-quarter level regression described in Equation (7) where we regress these dummies, quarter by product category fixed effects and cohort fixed effects on the annual expenditure of the new customers of a product in a given quarter. In Columns (1) and (2) we don't have city fixed effects. In Columns (3) and (4) we add city fixed effects, and these are the coefficients which are represented in Figure 7. In Columns (5) and (6) we look at consumers' socio-economic status, which is an integer between 1 (lowest) and 6 (highest). For each outcome, we run two separate regressions for foreign and domestic products, respectively. Standard errors are shown in parenthesis.

	Log population		Log density		Log distance to CDMX	
	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic
	(1)	(2)	(3)	(4)	(5)	(6)
Age=2 quarters	-0.610	-0.324	-0.534	-0.467	-0.048	-0.180
	(0.211)	(0.424)	(0.192)	(0.379)	(0.127)	(0.260)
Age=3 quarters	-0.786	-0.369	-0.856	-0.595	-0.068	0.013
	(0.306)	(0.633)	(0.278)	(0.569)	(0.182)	(0.381)
Age=4 quarters	-0.855	-0.562	-0.776	-0.998	-0.153	0.619
	(0.420)	(0.900)	(0.379)	(0.804)	(0.246)	(0.535)
Age=5 quarters	-0.944	-0.470	-0.895	-1.053	-0.165	0.507
	(0.548)	(1.194)	(0.493)	(1.057)	(0.320)	(0.702)
Age=6 quarters	-1.023	-0.444	-0.660	-1.238	-0.249	0.787
	(0.668)	(1.474)	(0.598)	(1.302)	(0.388)	(0.864)
Age=7 quarters	-1.066	-0.406	-0.712	-1.652	-0.370	1.514
	(0.787)	(1.778)	(0.705)	(1.561)	(0.457)	(1.035)
Age=8 quarters	-1.222	-1.299	-0.659	-1.631	-0.448	1.072
	(0.918)	(2.065)	(0.821)	(1.821)	(0.532)	(1.208)
Age=9 quarters	-1.263	-0.754	-0.873	-1.949	-0.292	2.068
	(1.036)	(2.326)	(0.928)	(2.059)	(0.600)	(1.364)
Age=10 quarters	-1.230	-0.384	-0.839	-1.938	-0.434	2.144
	(1.162)	(2.618)	(1.039)	(2.312)	(0.673)	(1.533)
Age=11 quarters	-1.646	-1.442	-0.862	-2.082	-0.467	2.006
	(1.285)	(2.925)	(1.151)	(2.574)	(0.745)	(1.706)
Age=12 quarters	-1.614	-0.828	-0.740	-2.621	-0.458	2.533
	(1.418)	(3.229)	(1.268)	(2.835)	(0.820)	(1.879)
Age=13 quarters	-1.732	-1.199	-0.945	-2.550	-0.351	2.516
	(1.550)	(3.501)	(1.384)	(3.088)	(0.896)	(2.047)
Age=14 quarters	-2.143	-0.575	-1.173	-3.419	-0.453	3.174
	(1.681)	(3.792)	(1.502)	(3.352)	(0.972)	(2.221)
Age=15 quarters	-1.911	0.677	-1.092	-2.128	-0.437	3.016
	(1.812)	(4.087)	(1.616)	(3.607)	(1.046)	(2.391)
Product X Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Initial quarter mean	14.14	14.22	5.26	5.26	6.01	5.97
N	4673	1802	5502	2151	5370	2093
R2	0.16	0.25	0.11	0.18	0.11	0.17

Table B.7: Barcode life-cycle: log sales-weighted average characteristics of new cities

Notes: Tables represents the coefficients obtained on the dummies for the age of the product in quarters from a product-quarter level regression described in Equation (7) where we regress these dummies, quarter by product category fixed effects, and cohort fixed effects on the average population of the new cities reached by a product in a given quarter in Columns (1) and (2). For each outcome, we run two separate regressions for foreign and domestic products, respectively. In Columns (3) and (4), we reproduce the regressions ran in Columns (1) and (2) with the average income per capita of the new cities and in Columns (5) and (6) with the poverty rate of the new cities. Standard errors are shown in parenthesis.