Defensive Specialization:
Theory and Evidence from Mexico's Retail Sector

Tommaso Bondi
Cornell University

Luís Cabral
New York University and *CEPR*

Shreya Kankanhalli
Cornell University

Miguel Angel Talamas Marcos
Inter-American Development Bank

August 28, 2025

Abstract. Large companies are increasingly dominating the retail industry, casting doubt on the future of independent, local retailers. Is this the end of local retail as we know it? What product strategy should local stores follow to attenuate this negative impact? We show that, in response to increasing penetration of large chains, local retailers – especially smaller ones – optimally follow a strategy of *defensive specialization*. Intuitively, the arrival of large chains hurts all independent local stores, but it especially hurts general stores that sell multiple categories. Our empirical evidence, built on two large retail datasets from Mexico, confirms the predictions of the theoretical model: specialty stores are better able to cope with the shock of large-chain entry. Moreover, the entry of a large chain induces a shift toward greater specialization in the traditional retail sector, both through an increase in the absolute number and market share of specialty stores and through changes in the product offerings within local stores towards more specialization.

We are grateful to Song Lin, Sridhar Moorthy, Julien Combe, Henrique Castro-Pires, and seminar participants at NYU, Cornell, IADB, CAF, Toulouse School of Economics, SICS, the ZEW Conference on the Economics of Information and Communication Technologies, the Frank M. Bass FORMS Conference and the Paris Digital Economics Conference for useful comments. We gratefully acknowledge the excellent research assistance from Emilia Bullano and Gonzalo Diaz. The views and conclusions are those of the authors and are not part of the official statistics of the SNIEG or INEGI. The confidential data from INEGI was accessed through the Microdata Laboratory in Mexico City.

1. Introduction

Neighborhood stores around the world are falling victim to the expansion of chains, big-box stores, and online shopping. Some talk about "retail apocalypse" when referring to the dire future of small neighborhood stores.[1] Is this the end of local retail as we know it? What types of local stores are most affected by the rise of retail giants? And what product strategy should local stores follow to attenuate this negative impact?

Our paper addresses these questions theoretically and empirically by studying the optimality of defensive specialization: responding to the entry of a larger competitor by reducing the number of product categories but increasing product variety within remaining categories. On the theory side, we develop a model of competition between a chain (or large) store and a local one. The large chain can offer greater product variety and potentially more product categories than the local store. However, consumers may have a preference for shopping local. We consider the local store's choice of whether to be a general store (carry two product categories) or a specialty store (carry one product category only). On the empirical side, we analyze the response of small local stores in Mexico's traditional sector to the massive entry of convenience and supermarket chains, and we test our model predictions regarding the optimality of defensive specialization.

Before turning to our findings, it is important to clarify how specialization is distinct from differentiation, studied previously in the retailing literature (e.g. Hollenbeck, Hristakeva, and Uetake (2024)). Differentiation occurs when local stores offer products or amenities not carried by the entrant. Specialization, in contrast, means narrowing to a single category and expanding depth within it, even if the entrant also sells every product. To make this as stark as possible, in our baseline model the large store offers all products, so differentiation is impossible. We later extend the model to allow for differentiation in amenities, and defensive specialization remains a robust feature. In practice, local stores may differentiate, specialize, do both, or do neither. Our empirical focus is on explicitly testing for specialization as a defensive response at the market level (number and share of specialized retailers) and the store level (degree of specialization of the product assortment). With this distinction established, we now derive the model's predictions on which local stores respond to entry through specialization.

Our central theoretical result is that the optimal assortment adjustment to the entry of a large chain depends critically on the incumbent's capacity disadvantage. When the capacity disadvantage is not too large (that is, when the number of products stocked by the entrant is not much larger than by the incumbent), the optimal strategy is to remain a generalist—essentially a smaller copy of the entrant. On the contrary, when the capacity disadvantage is substantial, the incumbent optimally switches to a specialist strategy, doubling down on a product category at the expense of the other(s). Thus, rather than a small copy of the larger entrant, the incumbent looks more like the copy of an aisle, or section, of it. This finding helps reconcile the rise of small specialized stores in settings that have experienced substantial entry by large chains (like our empirical one) with the fact that, when independent, similarly small stores compete among themselves, they tend to do so without specializing.

The intuition behind this result is that the incumbent's choice trades off the extensive

1. See, for example, "Retail apocalypse: How e-commerce slowly killed the neighborhood retail store over the past decade".

margin, which favors a generalist approach, and the intensive margin, which favors a specialist approach. A general store attracts more potential customers, but a specialty store is more appealing to consumers in its segment. Entry by a large chain implies that the intensive margins of both general and specialty stores decrease equally, because they both lose customers within each product category. However, the general store's extensive margin decreases at a faster pace than the specialty store's extensive margin, because the general store is losing customers in multiple product categories. Thus, while entry by a large chain is bad news for all local stores, it is particularly bad for local stores carrying multiple product categories. The local store benefits from being a general store when there is no competition, this benefit is decreasing with the entering chain's size and increasing with the local store's size. Therefore, as the size gap between the local store and chain widens, the benefit of being a general store decreases until it becomes optimal to be a specialty store instead.

We then extend our model in several directions, including pricing, consumer eclecticism, the possibility for the local store to invest in offline amenities, and the presence of competition between local stores. By doing this, we show that the optimality of defensive specialization is a robust finding and not an artifact of the simplicity of our baseline model.

We test these theoretical predictions by studying the effect of entry by convenience and supermarket chains on small, local stores in Mexico's traditional retail sector. We leverage comprehensive data and a refined identification strategy to overcome the following key challenges. First, data on small, informal establishments are limited. To address this, we obtain access to confidential Census microdata (collected and sheltered by the Mexican Statistics Institute, INEGI), covering the last two decades and including all retail establishments, both formal and informal, regardless of size. Second, systematically measuring specialization across diverse local stores poses difficulties. To address this, we utilize objective and consistent classifications by the economic census enumerators.

Third, establishing causal effects of entry is complicated by the endogenous nature of chains' entry location and timing. To address this, we pair neighborhood and city-year fixed effects with an instrumental variable strategy introduced in Talamas Marcos (2024). The instrument leverages variation across neighborhoods and across time to predict the entry of chains. The variation across neighborhoods is a measure of the suitability for chain stores based on the prevalence of wide streets, where most chains are located. The time variation builds on the intuition of Jia (2008) and Holmes (2011), is that different stores of the same chain in the same region share various costs (e.g., logistics, marketing, and overhead). Therefore, the expected profitability of a potential new chain store increases with the addition of stores from the same chain in nearby cities, as the average cost declines with the additional store sharing costs. However, the profitability of small, local stores in the traditional sector in neighborhoods suitable for chains remains unaffected by the addition of a chain store in a nearby city, except for the increased likelihood that the chain will open in its neighborhood.

Fourth, even detailed and extensive data collections, such as the one previously described Mexican census, almost never capture within-store product assortment decisions, limiting insights into retailers' strategic responses. We tackle this limitation by conducting two waves of surveys of over 500 retail stores in Mexico City. In this primary dataset, we capture store managers' explicit assortment decisions and their reasons for these changes. Together, these data allow us to rigorously test how local stores respond to the entry of large stores at the market and store level.

Consistent with the results from our theoretical model, we document a series of empirical effects of large chain entry. First, we find that the entry of a large chain store in a market (census tract) leads to an increase in the share and number of specialty stores, as well as their market share. Second, the market-level specialization effect increases with the capacity of the chain entrant, i.e., we find a larger impact on specialization from the entry of a supermarket. The share of specialty stores within the traditional sector increases by 4.7pp (17%) with the entry of a convenience chain store and 29pp (104%) with the entry of a supermarket in the census tract. Third, we find that, in the traditional sector, the entry of a chain leads to an increase in the share of exits by general stores and an increase in the share of entries by specialty stores.

Fourth and finally, using our survey data, we document within-store effects of large chain entry on assortment changes towards specialization. Our central result is that chain entry significantly increases specialization by *smaller* local stores, driving a 0.23 SD increase on an LLM-based specialization index using text data of assortment changes reported by local store managers. Meanwhile, we find negligible and statistically insignificant effects of large chain entry on the same specialization index for *larger* local stores, validating our theoretical predictions on the role of relative capacity disadvantages.

Overall, our theoretical and empirical analysis corroborates the proposition that competition from "giant retail" induces defensive specialization by local stores. The rest of the paper is structured as follows: in the remainder of Section 1 we review the literature and discuss our contributions. In Section 2, we present our main theoretical results and our theory extensions are in Section 3. Section 4 describes the datasets. In Section 5 and 6, respectively, we present our empirical evidence on both market-level specialization and within-store specialization. We conclude in Section 7.

■ **Related literature.** Our paper contributes to three related but separate streams of literature. First, and foremost, the work on the impact of entry by large chain stores, including Igami (2011), Atkin, Faber, and Gonzalez-Navarro (2018), Ellickson, Grieco, and Khvastunov (2020), Arcidiacono et al. (2020), Talamas Marcos (2024), and Caoui, Hollenbeck, and Osborne (2024). We extend this work by studying the impact on local stores based on their degree of specialization. In doing so, we uncover and explain an otherwise puzzling empirical pattern in important retail markets, such as food retail in Mexico, where specialty retailers collectively have *gained* market share in the face of entry by far more productive chains — contrary to the decline observed among generalist local stores (see Figure 4). Consistent with our analysis, Igami (2011) conducts an empirical analysis of Tokyo's grocery market and finds that the rise of large supermarkets does not crowd out small, independent stores, but rather mid-size ones. Building on this literature, we contribute by suggesting that niche specialization — a strategy not available to (or at least not optimal for) mid-size retailers — is an important driver of small stores survival, suggesting that these results might fail to hold in markets in which specialization is not a feasible strategy in the first place.

Our work also relates to a growing literature studying the effects of competition on equilibrium product assortments (e.g., Datta and Sudhir, 2011; Hollenbeck, Hristakeva, and Uetake, 2024; and Allcott et al., 2019). We emphasize that specialization and differentiation, while related, are conceptually distinct. In fact, there is no room for differentiation from the large chain in our analysis, because the large chain stocks every product. We analyze specialization (not differentiation) as a strategic lever where the intensive/extensive margin trade-off — and not price competition — is the key economic force. This deliberate modeling choice allows us to emphasize the separate role played by specialization in niche categories.

Last, and importantly, our findings speak to the literature on small firms in developing economies: the long tail of micro-enterprises (Hsieh and Olken, 2014), the survival of microenterprises (McKenzie and Paffhausen, 2019), and the barriers to small firm growth (McKenzie and Woodruff, 2017; Atkin et al., 2017; de Mel, McKenzie, and Woodruff, 2008). The defensive specialization mechanism that we highlight can contribute to explaining the persistence of small firms despite facing competition from large, productive chains. Moreover, it helps understand why, when only small local stores compete, generalists are prevalent, yet as larger chains enter, there are relatively fewer generalists and more specialty stores. In particular, our findings help explain why local specialty stores have overtaken local general ones as the primary purchase location of food for Mexican consumers (see Figure 4). Finally, defensive specialization, as a strategy for micro-enterprise survival, can be emphasized in business support programs such as those reviewed in McKenzie (2020).

2. Theory

Consider an economy with two stores, a (large chain) and b (local store); and z products divided into two different product categories, x and y. Store a carries all z products, while store b only has capacity for $k < z/2$ products. Although we assume that there exists only one local store, our intent is to model this as a typical local store, assuming that its effective competitor is the chain store. Later we also consider the possibility of competition between local stores.

There is a measure one of consumers. Independently of preferences for specific products or product categories, consumers have a preference for the local store, firm b, with respect to firm a. This may reflect an intrinsic taste for "buying local," the presence (or absence) of additional amenities (which we endogenize in one of our extensions), or an "ideological" aversion to chain stores. We assume that this preference (extra welfare from buying at a local store), denoted by $\widetilde{w}$, is uniformly distributed in $[0, w]$. The assumption that the lower bound of $\widetilde{w}$ is zero simplifies the analysis and is without loss of generality: all our results would be unaffected if we assumed a negative lower bound for $\widetilde{w}$, corresponding to a relative preference for the chain store. The reason for this is that, because the chain store has a size advantage ($z > k$), a positive $\widetilde{w}$ is required to buy from a local store. Put differently, all consumers with $\widetilde{w} \leq 0$ purchase from the chain store, so we can simply assume $\widetilde{w} \geq 0$ and focus our analysis on the competitive segment of consumer demand.

Each consumer only derives utility from one unit of a good. To best illustrate the main focus of our model, we start by assuming that product prices are constant and exogenously given, and with no further loss of generality we assume prices are equal to $1.

Consumers are equally split into two types, x consumers and y consumers.[2] x consumers (resp. y) like goods from product category x (resp. y) but derive no utility from goods from product category y (resp. x).[3] Specifically, in addition to the store preference utility $\widetilde{w}$,

2. Later in the paper, we consider the asymmetric case, that is, the case when one of the product categories has greater demand.

3. Later in the paper, we extend this to the case of eclectic consumers, who have positive valuations for both product categories.

consumers derive utility $m(t)$ from purchasing at a given store, where t is the number of varieties of the consumer's preferred category. For example, if store b only stocks goods from product category x, then x consumers derive utility $m(k)$ from buying from that store, whereas y consumers derive utility $m(0) = 0$. Throughout the paper, we make the following standard assumption regarding $m(t)$:[4]

Assumption 1. *$m(t)$ is increasing and concave.*

One possible justification — but certainly not the only one — for Assumption 1 runs as follows. Consumers of type x (resp. y) have a value $\widetilde{v}$ for one unit of a good from product category x (resp. y) and zero for any good from product category y (resp. x), where the value of $\widetilde{v}$ is generated from a cdf $F(\widetilde{v})$, where $f(\widetilde{v}) > 0$ if and only if $\widetilde{v} \in [0, v]$, where v in turn is possibly infinite. At a given store, consumers can learn both the product category and the value $\widetilde{v}$ of a specific product at no cost. By contrast, when store b chooses what products to carry, it can observe product category but not $\widetilde{v}$. Therefore, the store determines which product category to sell, but otherwise selects a random sample of values $\widetilde{v}$. Each consumer selects the store providing the highest expected value and, within a given store, buys the one product that yields the highest value $\widetilde{v}$. If the store carries t different products of the consumer's preferred product category, then the consumer receives an expected value $m(t)$, where $m(t)$ is the expected value of the highest element of a sample of size t drawn from $F(\widetilde{v})$. It can be shown this set up induces an increasing and concave $m(t)$.

■ **General vs specialty store.** The focus of our analysis is the local store's strategy. We first consider the case when b pays no fixed cost to remain active, so that it is a dominant strategy to do so. What is the store's optimal assortment decision? Should the local store become a specialty store, by stocking k products of one given product category, or a general store, by stocking $k/2$ products of each product category? The first strategy compromises the store appeal for half of its potential consumers; the second, on the other hand, potentially appeals to consumers of both types, but at the expense of variety in each product category.

Another way to think about this tradeoff is as follows: What is the better strategy to compete with a larger competitor, to resemble an aisle of it by focusing on a product category, or to resemble a mini version of it, by offering more product categories with a more limited variety? The answer to this question is not straightforward and depends on both stores' sizes (z and k), as well as on other model parameters.

Profits for a general and a specialty store are respectively given by

$$\begin{aligned} \pi_g(z,k) &= \left(1 - \left(\frac{m(z/2) - m(k/2)}{w}\right)\right), \\ \pi_s(z,k) &= \tfrac{1}{2}\left(1 - \left(\frac{m(z/2) - m(k)}{w}\right)\right). \end{aligned} \tag{1}$$

As the expressions make apparent, the specialty-store strategy gives up half of all potential consumers, but allows the local stores to be more competitive in inventory in the chosen product category ($z/2$ vs k as opposed to $z/2$ vs $k/2$).

Our first two results are based on the following assumption:

4. Considering the large number of different variables used in the paper, Table 1 lists the main notation used in the paper.

Table 1
Main notation used in the paper

Variable	Description
a, b	large chain (e.g. Walmart) and local (small) stores
k, c	store b's capacity and cost per unit of capacity
$\widetilde{d}, d$	horizontal distance from local store b_0; $d = \max \widetilde{d}$
f, F	pdf and cdf of $\widetilde{v}$
g, s	general and specialty store
$m(t)$	maximum $\widetilde{v}$ from t draws out of $F(\widetilde{v})$
p	price
q	local's store market share
t	number of products
$\widetilde{v}, \widetilde{w}$	vertical and horizontal preferences (maximum values: v and w)
x, y	popular and niche product category
z	total number of products (carried by store a)
α, β	popularity of x, fraction of b's capacity devoted to x
π	store b's profit
τ	transportation cost (when b_0 and b_1 compete)

Assumption 2. $w < \min \{m(z/2), v - m(k/2)\}$

Assumption 2 ensures that the solution is interior. When it fails to hold, we are in a corner solution whereby it is a dominant strategy for b to be a general store. If Assumption 2 holds, however, then the choice of general or specialty store depends on the relative values of z and k, as stated in the following result:

Proposition 1. *There exists a threshold $z_{gs} = z_{gs}(k, w)$ such that an active firm b optimally chooses to be a specialty store if and only if $z > z_{gs}$. Moreover, $z_{gs}(k, w)$ is increasing in both k and w.*

Proposition 1 can be equivalently stated in terms of local store heterogeneity, as shown in the following corollary:

Corollary 1. *There exists a threshold $k_{gs} = k_{gs}(z, w)$ such that an active firm b optimally chooses to be a specialty store if and only if $k < k_{gs}$. Moreover, $k_{gs}(z, w)$ is increasing in z and decreasing in w.*

The proof for this and all other results can be found in Appendix B.

Proposition 1 and Corollary 1 correspond to two ways of reading the same result. A local store optimally chooses to be a specialty store if the large store is sufficiently large (Proposition 1) or if the local store is sufficiently small (Corollary 1). In sum, defensive specialization is the optimal strategy if there is a large enough gap between large store and local store size. We stress that our empirical results speak to both heterogeneity in z – e.g., supermarkets vs. smaller chain stores – in Section 5 and heterogeneity in k in Section 6.

To understand the intuition for Proposition 1, note that the choice between a general and a specialty store trades off an "extensive margin" and an "intensive margin" effect. By switching to a specialty strategy, a store forgoes half of its potential customers, those interested in the product category that is no longer stocked (extensive margin). On the other hand, by stocking twice as many items of a given product category, the store increases the expected quality that a patron expects from visiting the store (intensive margin).

As z increases, the expected payoff from visiting chain a, $m(z)$, increases, making chain a relatively more attractive and lowering the demand for store b. This increase in valuation for a hurts the general store b more than the specialty store b. Basically, the general store loses consumers from both product categories, whereas the specialty store only loses consumers from a smaller set. It follows that, starting from a point where a general store strategy is better, there exists a threshold value of z past which a specialty store strategy yields higher profit.

Another way of understanding Proposition 1 is that, as z increases, the profit of both a general and a specialty store decrease. However, the profit of a general store decreases at a faster rate. In other words, specialty stores are better "insured" against chain entry, whereas general stores — who effectively resemble a small version of the chain entrant — suffer bigger profit losses.

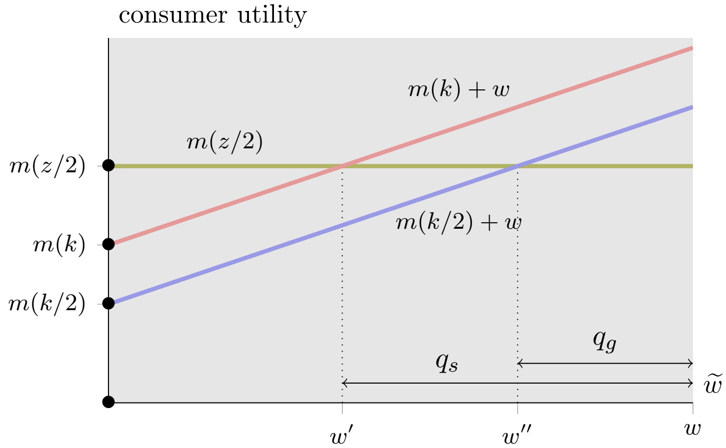
This idea is illustrated by Figure 1. The horizontal axis measures the value of $\widetilde{w}$: the further to the right, the more reluctant a consumer is to purchase from a. The vertical axis measures the value offered by each store (aside from the preference against a). Consider specifically the case when $\widetilde{w} = 0$, so that the consumer has no aversion to buying from a. The four bullet points along the vertical axis indicate the utility of buying from four different types of stores:

- A specialty store b offering k products of the product category the consumer is not interested in: utility zero.
- A general store b offering $k/2$ products of the product category the consumer is interested in (as well as $k/2$ products of the product category the consumer is not interested in): utility $m(k/2)$.
- A specialty store b offering k products of the product category the consumer is interested in: utility $m(k)$.
- Chain a, offering $z/2$ products of the product category the consumer interested in (as well as $z/2$ products of the product category the consumer is not interested in): utility $m(z/2)$.

As the value of $\widetilde{w}$ increases, the utility from shopping at the local store increases, whereas the utility from shopping at chain a remains constant. Thus w' is the threshold value of $\widetilde{w}$ such that a consumer with $\widetilde{w} > w'$ prefers a specialty store (of her preferred product category) with respect to store a, whereas w'' is the threshold value of $\widetilde{w}$ such that a consumer with $\widetilde{w} > w''$ prefers a general store with respect to chain a.

The values of k and z in Figure 1 were chosen so that store b is indifferent between a general and specialty strategy. To see this, notice that the share of customers buying at a specialty store, q_s, is twice as large as the share of customers buying at a general store,

Figure 1
Choice of general vs specialty store

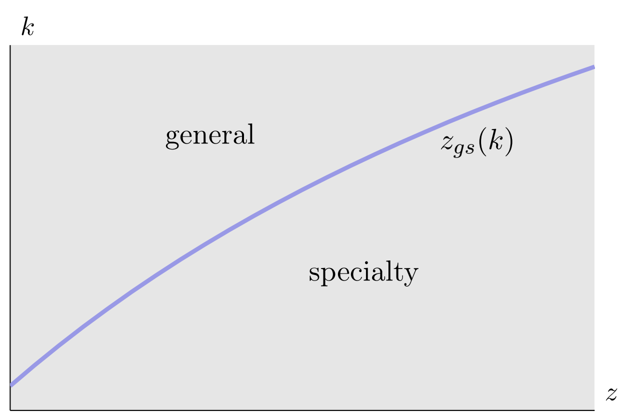


q_g. Since a general store attracts twice as many customers as a specialty store, the two differences exactly balance out.

The point of Proposition 1 is that, as we increase z beyond the value in Figure 1, both q_s and q_g decline at the same rate. However, *proportionately* speaking, q_g drops at a higher rate than q_s. Since the ratio of customers remains fixed at 1:2, it follows that, starting from the indifference point illustrated in Figure 1, becoming a specialty store becomes a dominant strategy for store b.

We consider comparative statics in both k and w. First, for a given value of z, a store with larger capacity k is less likely to specialize, that is, it requires a larger entrant for such a store to abandon a general strategy.

Figure 2
Comparative statics with respect to z and k.



Put differently, store b's decision to specialize is based on its *relative* size with respect to the entrant.[5] Similarly, the threat posed by the entrant is lower the greater w, that is, the greater the buyers' aversion to purchasing from a chain. Accordingly, given z and k, store b is less likely to become a specialty store as a strategy to cope with competition the higher w is.

Proposition 1 highlights the dynamic interpretation of small versus big competition, namely what happens as the chain increases in size. Conversely, Corollary 1 highlights the cross-sectional interpretation, namely what happens to large and to small local stores. The dual interpretation of Proposition 1 and Corollary 1 is illustrated by Figure 2, which plots the critical value $z_{gs}(k)$ derived in Proposition 1 (the plot assumes $\widetilde{v} \sim U[0,1]$). The figure may be read in two ways. One is to consider variations in z, which we think of as variation in the size/capacity of the chain entrant. Alternatively, we can consider variations in k, which we think of as variation in the size/capacity of the incumbent.

■ **Survival.** We now extend our analysis to allow for the possibility of exit. Suppose that the local store must pay a fixed cost $c\,k$ in order to operate, where c is cost per unit of capacity. For simplicity, we return to the assumption that both product categories are equally popular. Store profit is then given by

$$\begin{aligned} \pi_g(z,k) &= \left(1 - \left(\frac{m(z/2) - m(k/2)}{w}\right)\right) - c\,k \\ \pi_s(z,k) &= \tfrac{1}{2}\left(1 - \left(\frac{m(z/2) - m(k)}{w}\right)\right) - c\,k \end{aligned} \tag{2}$$

In this setting, the local store has three choices: remaining open as a general store, remaining open as a specialty store, and exiting. In order for the store's choice not to be trivial, we assume that a switch from general to specialty (or from specialty to general) implies a strictly positive cost (which however can be arbitrarily small). We also assume that the store's fixed cost is randomly selected from a given distribution and independent of store type (in other words, on average, a general and a specialty store have the same fixed cost).

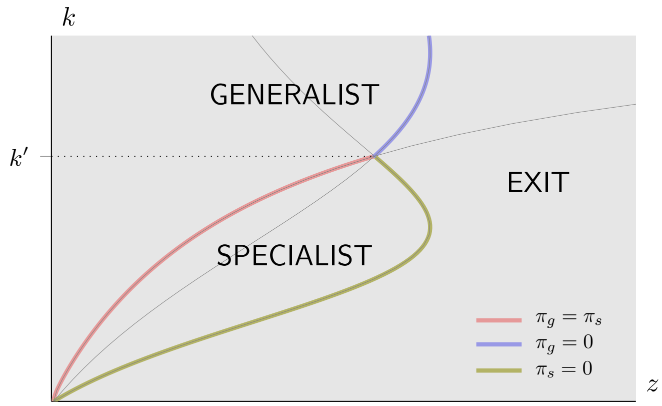
Proposition 2. *If z is high enough, then, upon entry by a large store, general stores are more likely to exit than specialty stores.*

Figure 8 illustrates the new equilibrium, where we assume that $\widetilde{v}$ is uniformly distributed. The red line corresponds to the indifference condition $\pi_g(z,k) = \pi_s(z,k)$. It's the same line as $z_{gs}(k)$ in Figure 2. As per Proposition 1, for points to the NW (higher k or lower z), firm b prefers to be a general store, whereas for points to the SE of the red line firm b prefers to be a specialty store. What Figure 8 adds with respect to Figure 2 is the possibility of exit. Specifically, two additional lines are plotted in Figure 8: the blue line corresponds to the zero profit condition for a general store, whereas the green line corresponds to the zero profit condition for a specialty store.

Note that the three lines must cross at the same point. In fact, when the blue line and the green line cross, both $\pi_g = 0$ and $\pi_s = 0$. Since both stores have the same profit level,

5. Non-linearities in $m(\cdot)$ imply that the ratio k/z is not a sufficient statistic for the specialization decision. Nevertheless, the specialty strategy is more likely when either k is small or z is large.

Figure 3
Comparative statics with respect to z and k when exit is a possibility



firm b is indifferent between being a general store and a specialty store, which in turn implies that the red line must cross at the point.

We thus have three well-defined regions. The general area is located above the red line (g is better than s) and to the left of the blue line (g is better than nothing). The specialty area is located below the red line (s is better than g) and to the left of the green line (s is better than nothing). Finally, the EXIT area is located to the right of the blue and green lines (exit is better than g and s).

For lower values of k, specifically for $k < k'$, the intuition underlying Proposition 1 still applies: as chain a increases in size/capacity, firm b optimally switches its strategy from general to specialty store. However, such strategy can only help to some extent: as firm a continues to increase in size/capacity, eventually firm b optimally exits.

The green and blue lines in Figure 8 also suggest that there is an "optimal" size k for a given store strategy. Not surprisingly, the "optimal" k is higher for a general store than for a specialty store. Relatedly, the figure also suggests that the comparative statics with respect to k are far from trivial. This is particularly the case as we consider variation in the value of k for z slightly higher the point at which the three lines cross: as we increase k from zero, firm b's optimal strategy changes from exit to being a specialty store to exit to being a general store.

Proposition 2 is similar to, but different from, Proposition 1. What the two results have in common is the idea that entry by a large chain affects large local stores proportionately more. In fact, the two results combined allow us to say something about store design.

■ **Store design shift.** Combined, Propositions 1 and 2 imply that, if a large enough chain store ($z > z_{gs}$) enters the market, then some of the surviving general stores optimally become specialty stores. The next Corollary summarizes our previous results:

Corollary 2. *Upon entry by a large chain, small enough general stores either exit or become specialty stores.*

The model developed so far is very stylized. Parsimony is a feature, not a bug: it allows us to identify the main effects at work and to derive testable empirical implications.

Specifically, our theory predicts that entry by a large chain into a local market will have the following effects:

- Reduce the local stores' profitability, particularly that of local general stores
- Lead local stores to become more specialized, especially *smaller* stores or those facing the entry of a *larger chain* competitor
- Increase the exit probability of local stores, especially general stores

In Sections 5 and 6, we test these predictions and derive additional results of interest in the context of entry by large chains in local markets in Mexico. Before then, in the next section, we derive a series of model extensions.

3. Theoretical model extensions

As mentioned earlier, our basic model is based on a series of simplifying assumptions: we consider competition between *one* small store and *one* large entrant; we assume away price competition; we assume individual consumers have preferences for *one* of two (symmetric) product categories; we assume an exogenously given preference for small stores.

Our main results, and their intuition, do not rely on these simplifying assumptions. In this section, we relax each of the elements in our previous paragraph. In particular, in our next four Propositions, we consider what happens with asymmetric product categories; endogenous prices; eclectic consumers, who value both categories; competition between two symmetric stores (as well as with a larger chain entrant); and endogenous investments in amenities.

The goal of this exercise is twofold. First, we show that our *defensive specialization* results are not an artifact of the simplified framework we use to derive Propositions 1 and 2. Second, our model extensions allow us to derive a series of interesting additional results regarding the complementarity of specialization decisions and other strategic choices available to local stores.

■ **Niche product categories.** So far, we have assumed that both product category x and product category y have the same popular appeal. A more realistic case has one of the product categories — say, product category x — be a popular product category, whereas y is a less popular one — a niche product category. Suppose that there is a measure 1 of potential buyers, α of which are only interested in product category x; and suppose that $\alpha > \frac{1}{2}$. (So far, we have implicitly assumed that $\alpha = \frac{1}{2}$.) Consistent with the assumption that product categories x and y have different popular appeal, we assume that a fraction $\alpha\, z$ of the total products are of product category x, and a fraction $(1-\alpha)\, z$ are of product category y.

Proposition 1 states that, as z increases, store b optimally switches from a general to specialty store. The next proposition complements that result by stating that, within the specialty strategy, store b optimally chooses the niche product strategy if z is high enough.

Proposition 3. *There exists an z_{xy} such that an active store b specializes in a niche product category (rather than a popular product category) if $z > z_{xy}$.*

Similar to Proposition 1, Proposition 3 has a dual interpretation where the comparative statics can be done with respect to k rather than z:

Corollary 3. *There exists a k_{xy} such that an active store b specializes in a popular product category if $k > k_{xy}$.*

Figure 4
Store profits from specializing in popular product category (π_x) or niche product category (π_y) as a function of z when $F(\widetilde{v}) = \widetilde{v}/v$.

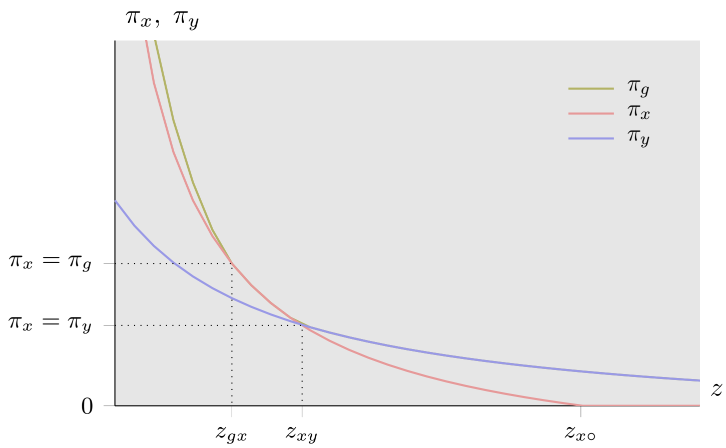


Figure 3 illustrates Proposition 3 and Corollary 3. The key insight is that, *relatively* speaking, a niche-product category store suffers less from an increase in z than a popular-product category store, in a way that is similar to, but different from, the general-specialty trade-off considered in Proposition 1. For low values of z, the advantage of a niche-product category store, in terms of higher intensive margin, is outweighed by the simple fact that a popular product category is more popular, that is, attracts a greater number of potential customers. For high values of z, however, the niche strategy becomes increasingly attractive, as illustrated by Figure 3. Specifically, for $z > z_{xy}$, π_y, the profit from a niche-product category strategy, is greater than π_x, the profit from a popular-product category strategy.

Formally, the proof of Proposition 3 proceeds by deriving the value z_x when $\pi_x = 0$ and establishing that, at that value, $\pi_y > 0$. This proof strategy is similar to that of Proposition 1. There is one difference, however: In Proposition 1, we show that $z > z_{gs}$ is a necessary and sufficient condition for specialization. By contrast, in Proposition 3 $z > z_{xy}$ is only a sufficient condition. The difference stems from the fact that we can prove the monotonicity of $\pi_s - \pi_g$ in general terms but not the monotonicity of $\pi_y - \pi_x$. If we further assume that v is uniformly distributed, then the condition $z > z_{xy}$ becomes a necessary and sufficient condition.[6]

■ **Endogenous prices.** So far, we have assumed that all items are priced \$1. We now explicitly consider pricing choices. Our goal is to verify the robustness of our previous findings as well as to develop additional intuition regarding the comparative statics of the chain's expansion.

6. The proof can be obtained from the authors upon request.

Recall that the actual market structure we have in mind includes one dominant firm and a large number of fringe firms. Although for simplicity we focus on the decisions of one representative fringe firm, it makes sense to treat firms a and b as different types of strategic players. Consistent with this interpretation, we assume that firm a acts a price leader by setting p_a first.

Given p_a, the local store b responds by deciding whether to be a general store or a specialty store as well as by setting its price, which we denote by p_g if the store is a general store and p_s if the store is a specialty store. Our focus is on firm b's decisions. Accordingly, we take p_a as an exogenous variable (and later consider comparative statics with respect to it).[7] Similar to Propositions 1 and 3, we make a parametric assumption so as to eliminate trivial corner solutions (if the assumption below fails to hold, then choosing to be a specialty store is always optimal).

Our next result extends the main intuition of Proposition 1, adding one new dimension of comparative statics.

Proposition 4. *Suppose that* $p_a > \underline{p}$, *where*

$$\underline{p} = w + \frac{m(k) - \sqrt{2}\, m(k/2)}{\sqrt{2} - 1}$$

Then, there exists a threshold z_{gs} *such that store* b *optimally chooses to be a specialty store if* $z > z_{gs}$. *In the right neighborhood of* z_{gs}, *the specialty store sets a higher price, captures a lower market share and earns a higher profit than a general store.*

When discussing Proposition 1, we argued that the trade-off between a general and a specialty store is a trade-off between the extensive margin (which favors a general store) and the intensive margin (which favors a specialty store). The proof of Proposition 4 establishes that, when it comes to price setting, only the intensive margin matters. This explains why a specialty store sets a higher price than a general store. By devoting its space to one product category only, a specialty store elicits a higher willingness to pay from buyers interested in that product category, which in turn allows the store to set higher prices. This in turn increases the store's incentives to specialize.

Similar to Proposition 1, Proposition 4 establishes that, if chain a is big enough (high z), then firm b is better off by becoming a specialty store. The main intuition for the z-threshold part of Proposition 4 is similar to Proposition 1: As total supply z increases, the specialty store option becomes *relatively* more attractive. In sum, the first part of Proposition 4 shows that the intuition from Proposition 1 is robust to the introduction of pricing.

The novel aspect of Proposition 4 is its second part, the statement that, past the disruption level z_{gs}, a specialty store sets a higher price, captures a *lower* market share and earns a higher profit than a general store. We call this the *boutique effect.* The specialty store in the model with fixed prices trades-off extensive margin and intensive margin so as to maximize the number of customers. By switching from general to specialty store, firm b loses potential customers, but its offering becomes so much more attractive to its reduced set of customers that it ends up attracting more customers. By contrast, once we introduce prices we observe that the switch to a specialty-store strategy not only sacrifices *potential* demand

7. Endogenizing the chain's decisions would be a promising direction for future research.

but also sacrifices *actual* demand. Such drop in actual demand is more than compensated by an increase in the intensive margin via higher retail prices.

■ **Eclectic consumers.** So far, we have assumed that consumers are divided into x fans and y fans. Specifically, the value v of an item outside of a consumer's preferred product category is zero. At the opposite extreme, consider the case when consumers are totally eclectic, that is, they value both product categories equally.

Clearly, eclectic consumers are bad news for specialty stores. Before, an x fan valued a specialty store at $m(k)$ and the chain store at $m(z/2)$. By contrast, an eclectic consumer values the chain store at $m(z)$ whereas the specialty store is still valued at $m(k)$ (here we are excluding the preference parameter $\widetilde{w}$).

Regarding a general store, the analysis is not as obvious. Before, the value of a general store was $m(k/2)$ for an x fan or a y fan, whereas the value of the chain was $m(z/2)$. By contrast, an eclectic consumer values the chain at $m(z)$ whereas the general store is valued at $m(k)$ (again, we are excluding the preference parameter $\widetilde{w}$). In which case is the general store better off? The answer depends on which difference is greater, $m(z/2) - m(k/2)$ or $m(z) - m(k)$. Notice that $m(z) - m(k) > m(z/2) - m(k/2)$ if and only if $m(z) - m(z/2) > m(k) - m(k/2)$. Since $z > k$, we have $z - z/2 > k - k/2$, which in turn suggests the inequality holds. However, concavity of $m(t)$ would work against the inequality. Suppose that $F = v$ is linear, so that $m(t) = t/(1+t)$. Then the function $m(t) - m(t/2)$ is non-monotonic, first increasing for $t \in [0, \sqrt{2}]$ and then decreasing. This implies that we can find values of z and k such that the inequality is, in turn, true or false. So, even assuming a specific distribution of v, we cannot guarantee that a general store is better off or worse off when serving eclectic consumers rather than polarized consumers.

It has long been argued that large chains benefit from increased consumer specialization due to their larger size, since local stores cannot, due to their limited size, cater to each consumer's idiosyncrasies (Anderson (2004), Waldfogel (2007)). However, as the above analysis shows, this is not necessarily true when we endogenize local stores' strategies: more specialized consumers allow specialty stores to emerge, which can be detrimental to the chain's profits.

■ **Traditional stores amenities.** We know that, *ceteris paribus*, local stores' survival is crucially dependent on the relative consumer preferences for traditional retail ($\widetilde{w}$). So far, we have treated this distribution as exogenous. We now consider the case when stores make investments that determine consumer preferences. For example, local shops might build relationships with the community. The owner, often a neighbor, can offer tailored customer service and engage in conversations. The store itself may be a gathering point for the neighbors. Moreover, knowing the neighbors so well allows local shops to offer amenities unfeasible for chains, such as informal credit. These features are appealing because chains cannot directly replicate them.

Are amenities a complement or a substitute for specialization? And which stores benefit the most from them? The following result provides an answer:

Proposition 5. *Generalist stores have a higher incentive to invest in amenities:* $\partial \pi_g / \partial w > \partial \pi_s / \partial w > 0$. *This incentive gets stronger the larger the size of the entrant:* ($\partial^2 \pi / \partial w \, \partial z > 0$) *and/or the smaller the size of the store* ($\partial^2 \pi / \partial w \, \partial k < 0$).

The first part of Proposition 5 states that generalist stores get more bang for the buck when it comes to investment in local amenities. The intuition is that, by serving a wider customer base, the investment can be spread on a larger number of consumers.

The second part of Proposition 5 refers to cross-partial derivatives, that is, to moderating factors in evaluating the benefits from investment in amenities. First, the greater the chain is, the more a local store benefits from investment. Intuitively, just as specialization is a strategic reaction to a large chain entry, so is investment in store amenities. Second, the greater the local store, the lower the store's benefit from investment in amenities. Similar to Propositions 1 and 3, these two results (comparative statics with respect to z and to k are connected: ultimately, investment in local store amenities is more profitable the greater the gap between local store size (k) and chain size (z).

Proposition 5 suggests that, for a small store (low k), investing in amenities (i.e., increasing the value of w) may provide an alternative strategy to specialization. This is particularly the case when a significant fraction of consumers are eclectic (so that specialization is not a profitable strategy).

■ **Local store competition.** Up to now, we have considered competition between one large chain and one local store. Implicitly, the idea is that there are a plethora of small (possibly independent) local stores with a catchment area that does not overlap with any other local store. Consider now the case when two local stores, say b_0 and b_1, do compete for the same potential demand. Specifically, we assume a consumer is characterized by a local-store preference $\widetilde{w}$ and a relative preference between stores b_0 and b_1 in the form of a location $\widetilde{d} \in [0,1]$ and transportation cost τ per unit of distance to store b_0 (located at 0) and to store b_1 (located at 1). Moreover, we assume that $\widetilde{d}$ and $\widetilde{w}$ are independently and uniformly distributed: $\widetilde{d} \sim U[0,1]$ and $\widetilde{w} \sim U[0,w]$. Our main result is that, under competition, the choices of product category by stores b_0 and b_1 exhibit strategic complementarities.

Proposition 6. *Let z be such that store b_0 and b_1 are indifferent between a general- and a specialty-store strategy absent competition between local stores. In the neighborhood of z, being a specialty store is a strict best response to the rival choosing to be a specialty store.*

Proposition 6 suggests that competition provides an additional force pushing in the direction of specialization. Suppose that we fix firm b_1's strategy at being a general store. As z crosses a certain threshold, say $z_\circ$, then firm b_0's optimal strategy switches to becoming a specialty firm (of either x or y). However, if firm b_1 has become a specialty firm (choosing, say, product category y), then, *even if z is lower than $z_\circ$* (by a little), firm b_1 also optimally switches to being a specialty (specializing in the niche that firm b_1 did not).

Next we consider data from retail in Mexico as a testing ground of our theory and predictions.

4. Data and background

The empirical analysis relies on two waves of detailed surveys of retail stores in Mexico City collected by the authors, nationwide confidential microdata collected and sheltered by the Mexican Statistics Institute (INEGI), and public data.

■ **Confidential Microdata.** The confidential microdata from INEGI are the Economic Censuses (1999, 2004, 2009, 2014, 2019) and the Income and Expenditure Surveys (2008, 2014, 2018). The Economic Censuses cover all the physical establishments in the country – both formal and informal, of any size. The Economic Censuses classify the establishments according to the North American Industrial Classification System for Mexico (SCIAN).[8] We use this classification at the six-digit level to distinguish between general and specialty stores. We consider establishments that offer a wide variety of product categories as general stores: supermarkets (462111), convenience stores (462112), and neighborhood shops (461110).[9] We consider a firm a large chain if it is a supermarket chain with more than 20 establishments at any point in time, or if it is a convenience chain with over 100 establishments.

The primary specification considers as specialty establishments all the other classifications within the retail sale of groceries, food, beverages, ice, and tobacco category (461), which are retail sale of red meat (461121), poultry meat (461122), fish and seafood (461123), fresh fruits and vegetables (461130), edible seeds and grains, spices, and dried chiles (461140), milk and other dairy products (461150), sweets and pastries (461160), popsicles and ice cream (461170), other food products (461190), wine and spirits (461211), beer (461212), non-alcoholic beverages and ice (461213), cigarettes, cigars, and tobacco (461220).[10] We only exclude establishments in these classifications if they are owned or operated by the government or belong to a chain.[11]

Starting in 2009, INEGI added an establishment identifier to the Economic Censuses. To track establishments before 2009, we use the establishment identifiers created by Busso, Fentanes, and Levy (2018). We use this establishment-level panel from 1999 to 2019 to estimate the effects on exit.

The Income and Expenditure Surveys (ENIGH) contain data on what households buy and in what type of establishment they buy it. While the ENIGH is publicly available, we access the confidential version of it in the INEGI microdata lab to use data on where the household lives down to the census tract (AGEB) level. We identify transactions in specialty stores as those where the purchase location, categorized by INEGI, is a specialty store, defined as "[e]stablishments dedicated to the commercialization of a single line of business—that is, they sell one type of product or service" (ENIGH, 2006). The sample of the ENIGH has grown throughout the years. In 2006, it contained responses from a little more than 20,000 households; by 2018 it included more than 70,000 responses.

INEGI's geo-statistical framework for urban Mexico divides the country into states, municipalities, localities, and urban census tracts (AGEBs). The total number of AGEBs

8. This classification has subtle differences from the one for the United States (NAICS)

9. Code 461110 is for establishments engaged in Retail trade in grocery, variety, and miscellaneous stores. The Mexican Statistical authority describes them as economic units primarily engaged in retail trade, through traditional methods or online, of a wide variety of products such as milk, cheese, cream, cold cuts, candies, cookies, bread, pastries, snacks, fried foods, canned goods, bottled purified water, soft drinks, beer, bottled wines and liquors, cigarettes, eggs, toilet paper, detergent, soap, paper napkins, and disposable kitchen utensils (INEGI, 2025).

10. For robustness, we use two alternative classifications of specialty stores: i) establishments within retail sale of groceries and food (4-digit classification 4611) , and ii) establishments within beverages, ice, and tobacco (4-digit classification 4612). The results are robust to these alternative classifications.

11. The results are robust to including specialized chains. However, our goal is to model the effect of specialization in the traditional sector. Therefore, we exclude specialized chains in the main specification.

ranges from 37,000 to 47,000 (depending on the census year). Each AGEB typically consists of 25 to 50 blocks, which encompass an average of 650 households and around 2,000 residents. AGEBs are precisely defined by streets, avenues, or other easily identifiable characteristics in the field. This design by INEGI is intended to facilitate the data collection process conducted by enumerators.

For the analysis using the Economic Censuses data, the market definition of our main specification is the census tract. However, for the analysis using the ENIGH, we need to construct larger markets, because few census tracts are surveyed in multiple waves of the ENIGH. We use the census tracts to define markets by drawing a buffer (circle) around the centroid of each census tract, and define a market as the union of census tracts that overlap with each buffer. We draw these buffers with a radius of 1 km for our main specification and with radii of 1.5 and 2 km for robustness.

■ **Public Data.** For one of our identification strategies, we use the street width from Open Street Maps. We classify trunk, primary, secondary, and tertiary streets as *wide* and the remaining categories as *narrow*. In the resulting classification, 21% of the total street length is *wide*. The remaining 79% of streets are *narrow*, and the vast majority of the narrow streets, 95%, are residential streets. We construct a measure of the prevalence of wide streets by adding the lengths of all wide streets in the market and dividing it by the market size, specifically dividing by the square root of its area.

■ **Primary Data Collection.** We conducted two waves of in-field business surveys among 554 local retail stores in Mexico City between 2018 and 2020, prior to the onset of the COVID-19 pandemic. Our sampling strategy focused exclusively on local stores (i.e., no chains) with up to 10 employees. Our field team canvassed all major commercial areas of the city to invite any store fitting this profile to participate in our study. The resulting sample is those who consented to respond in this multi-year data collection exercise, and we discuss our evidence on the representativeness of this sample further in this section.

During the baseline survey, we collected data on: (i) the location of the store through GPS coordinates; (ii) the type of store based on the North American Industrial Classification System (SCIAN) code and a open-text description of the main products and categories sold; (iii) the total number of SKUs; and (iv) the top 3 SKUs sold at the store by volume. During the endline survey in early 2020, we asked store managers whether they had added any new products to their assortment, whether or not they had removed any existing products from their assortment, and if they reported either, we collected text data on what products they removed or added. Additionally, we asked managers to report whether the *primary* reason for any assortment change was competition, demand, marginal cost or margin considerations, or any other reason (e.g., a supplier issue).

Figure A.1 maps the surveyed stores in Mexico City, and Table A.1 provides descriptives on store characteristics. The geographical spread in the stores surveyed, pictured in Figure A.1, allows us to obtain useful variation in nearby chain entry and ensures better representativeness of the city's retail landscape. Also, our sample is fairly representative per Table A.1 as means compare closely with citywide means from the 2019 Economic Census for retail establishments in the key sub-sectors represented in our sample (461—groceries, 465—stationery and office supplies, and 722—prepared food).

We merge this panel dataset from the field with the microdata from INEGI's firm registries (DENUE) to identify entries of chain competitors—both supermarket chains and convenience chains—between baseline and endline. A chain is considered a competitor if it enters within pre-defined radii (500 m, 1 km, and 2 km) that correspond to trading areas in urban environments where stores primarily rely on walking traffic.

■ **Background.** The traditional sector, characterized by a large number of small, often family-owned businesses, plays a crucial role in Mexico. It includes more than one million establishments engaged in the retail sale of groceries, food, and beverages, as well as over 500,000 outlets that sell prepared food. Establishments in these two categories alone represent 34% of all establishments, 15% of employment, and 4% of value added in the country (Economic Census, 2019). Within the traditional sector, general and specialty stores coexist. Figure A.2 shows some examples of these two types of establishments.

The general shop is the most common establishment in the country. There are nearly 600,000 of these small establishments, characterized by selling a broad range of food, beverages, and household and kitchen products. They are mostly owner-operated and often located next to the owner's house. Because of their large number, they have almost 30% market share of the food and beverage retail industry. However, this market share has declined by 10% in the last two decades (see Figure 4).

On the other hand, with over 400,000 establishments, specialty stores are quite prevalent in the food retail context. These include, for example, *tortillerias*, butchers, poultry shops, seafood retailers, dairy stores, and fresh produce stores. In the past twenty years, specialty stores have increased their market share in the food sector by more than 30%, establishing themselves as the leading choice for food purchases nationwide (see Figure 4).

Table A.2 compares general and specialty stores in 1999, before most convenience and supermarket chain entries. General and specialty stores were quite similar in size. There was no difference in revenue, and slight differences in profits and age. Specialty stores had 5% lower profits and were 8% younger. They also employed 2% fewer workers but hired 15% more workers (relied more on hired labor). They also had no difference in revenue per worker, and specialty stores had 6% lower profits.

Over the past two decades, Mexico has experienced a considerable expansion of convenience and supermarket chains. The two most prominent participants in the market, Oxxo and Walmart, have established over 20,000 and 3,000 retail outlets, respectively. This proliferation of chains has significantly altered the retail landscape in Mexico (see Figure 5). These chains' establishments are larger than specialty and general stores in the traditional sector, and they are general stores, offering a wide variety of categories.

5. Market-level evidence of specialization

We first provide evidence for Propositions 1 and 2 using nationwide confidential microdata that enables us to estimate the effects of the entry of general chain stores on the degree of specialization of the traditional retail channel. We start by introducing our empirical strategy and estimating equations. Then, we transition to the empirical evidence. We show that the entry of a chain (whether a supermarket or a convenience store) has a negative effect on the average revenue and profits of traditional general and specialty retailers. The effects are larger for the entry of supermarkets and increase with the proximity of the entry. Consistent with our theoretical predictions, the effects are significantly larger for general

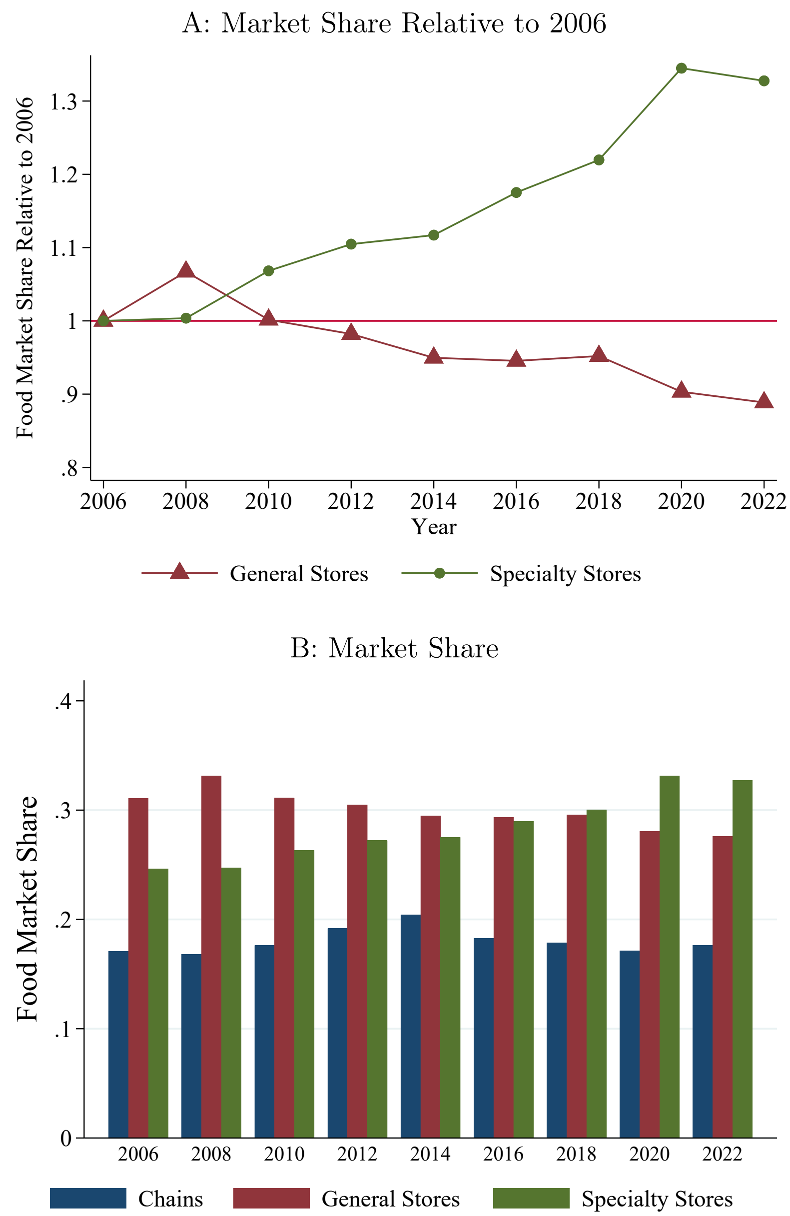


Figure 5
Market Share by Store Format in Mexico's Food Market

Source: National Income and Expenditure (ENIGH) Surveys (2006 - 2022)
Note: Panel A displays food expenditure shares by store format relative to their 2006 share. Panel B presents shares by store type. All food expenditures are included (category "A" in the data). Chains include supermarkets, membership clubs, department stores, and convenience stores. General stores include neighborhood shops, and specialty stores are establishments that focus on a single product category (as defined by ENIGH). The calculations use representative weights provided by ENIGH.

than for specialty stores. Then, we show that as a result of the entry, the traditional sector becomes more specialized. The number of general stores decreases, the number of specialty stores increases, and the share of revenue of specialty stores increases. The estimates based on household consumption data provide consistent evidence: the share of expenditure in specialty establishments and the share of visits to specialty establishments increase.

5.1. Empirical strategy

Chains are typically drawn to markets with high demand and foot traffic. These are also the markets where shops in the traditional sector are more likely to be successful, leading to an endogeneity problem in estimating the relationship between the number of chain stores in a market and outcomes for traditional stores. Including market fixed effects can alleviate this concern because they can control for market-specific time-invariant characteristics. However, our nationwide data spans twenty years, and we only observe firms every five years, potentially leaving the estimation vulnerable to bias arising from market-specific shocks that affect the profitability of chains and traditional sector stores occurring within these five-year windows. To alleviate this concern, we follow an identification strategy based on instrumental variables introduced in Talamas Marcos (2024) that builds on the intuition of Jia (2008) and Holmes (2011), where stores of the same chain in the area share costs (e.g., logistics, marketing, and overhead). Therefore, the expected profitability of a potential new store increases with the addition of stores from the same chain in nearby cities, as the average cost declines with the additional store sharing costs.

The instrument is a cost shifter that reduces chains' costs and increases their profitability, without affecting establishments in the traditional sector. It is based on an interaction of two components. The first one captures the regional economies of scale (cost sharing) of chains, and it is the first term of the product in equation 2. Specifically, it is the square root of the sum of the squared lagged number of stores per chain in nearby municipalities (a Herfindahl-Hirschman Index without normalization), where nearby municipalities are the second-degree neighbors (adjacent municipalities and those adjacent to them).[12] This measure increases with the number of same-chain stores in nearby cities, and it provides variation at the municipality and year level, but does not predict where new chain stores will locate within municipalities.

The second component incorporates variation within cities based on the suitability of the market for chains based on the prevalence of wide streets. While establishments in the traditional sector are often next to the owner's house, chains enter on wide streets and intersections to target traffic customers. We measure the suitability of a market for chain stores using its total length of wide streets divided by the square root of the market area (second term in the product in equation 2). The instrument we use is the product of these two components,

$$Z_{m,c,t} = \underbrace{\left(\sum\nolimits_f (\#StoresNearbyMun_{f,c,t-1})^2\right)^{1/2}}_{\text{Regional Economies of Scale}_{c,t}} \times \underbrace{\frac{Total\ wide\ streets\ length_{m,c}}{Area_{m,c}^{1/2}}}_{\text{Prevalence of Wide Streets}_{m,c}}, \quad (3)$$

where Z is the instrument, m denotes the market, c denotes the municipality (city), t denotes the census year, and f denotes the firm.

12. In the robustness section we show that the results are robust to using third-degree neighbors as well.

Our first and second stages estimating equations are

$$CS_{m,c,t} = \gamma_1 Z_{m,c,t} + \zeta_{m,c} + \eta_{c,t} + \sigma X_{m,c,t} + \mu_{m,c,t}, \text{ and} \tag{4}$$

$$Y_{m,c,t} = \beta_1 \widehat{CS}_{m,c,t} + \zeta_{m,c} + \eta_{c,t} + \sigma X_{m,c,t} + \epsilon_{m,c,t}, \tag{5}$$

where $Y_{m,c,t}$ is the outcome of interest (e.g., number of general establishments, number of specialty establishments, share of specialty establishments, and specialty stores' share of revenue). CS stands for the number of chain stores. The estimation includes market fixed effects, $\zeta_{m,c}$, and municipality-year fixed effects, $\eta_{c,t}$. Standard errors are clustered at the municipality level because the measure of advantages from the regional expansion of chains varies at the municipality level.

We estimate the effect of the entry of a convenience chain and a supermarket chain separately. Therefore, we construct an instrument for the number of convenience chains in the market and another one for the number of supermarket chains. Estimating the effects separately enables us to assess whether the entry of a larger general store has a greater impact on the degree of specialization, as predicted by Proposition 1.

■ **Potential Identification Concerns.** The instruments' exclusion restriction is that traditional general and specialty stores in areas favorable for chains are affected by an increase in the number of chain stores in adjacent municipalities *only* due to the heightened likelihood of a chain establishing itself in their neighborhood.

A first potential concern could be that chains are entering cities that are growing faster, which will also be the cities where the traditional sector sales are increasing. City-year fixed effects control for this and other city-level potential shocks. A second concern is that chains enter affluent neighborhoods, where traditional sector shops are often more successful. Market fixed effects address this issue and others related to cross-sectional time-invariant differences across markets.

A third possible concern is that chains may enter cities that are growing faster in markets where wide streets are more prevalent, which may also be the markets where income and demand are growing faster. This is unlikely to be an issue because while the entry of same-chain stores in nearby cities predicts openings in a city, the openings of competitors in nearby cities do not.[13] Therefore, the entry in a city is likely driven by chain-specific characteristics and not city-wide ones.[14] Additionally, our main outcomes of interest are in relative terms, e.g., the share of specialty stores in the market. Thus, if something affects general and specialty stores in the traditional sector in a similar manner, it will cancel out.

Since Talamas Marcos (2024) has an extensive discussion on the validity of the same instrument in the same context, we defer further discussion to that work.

5.2. Results

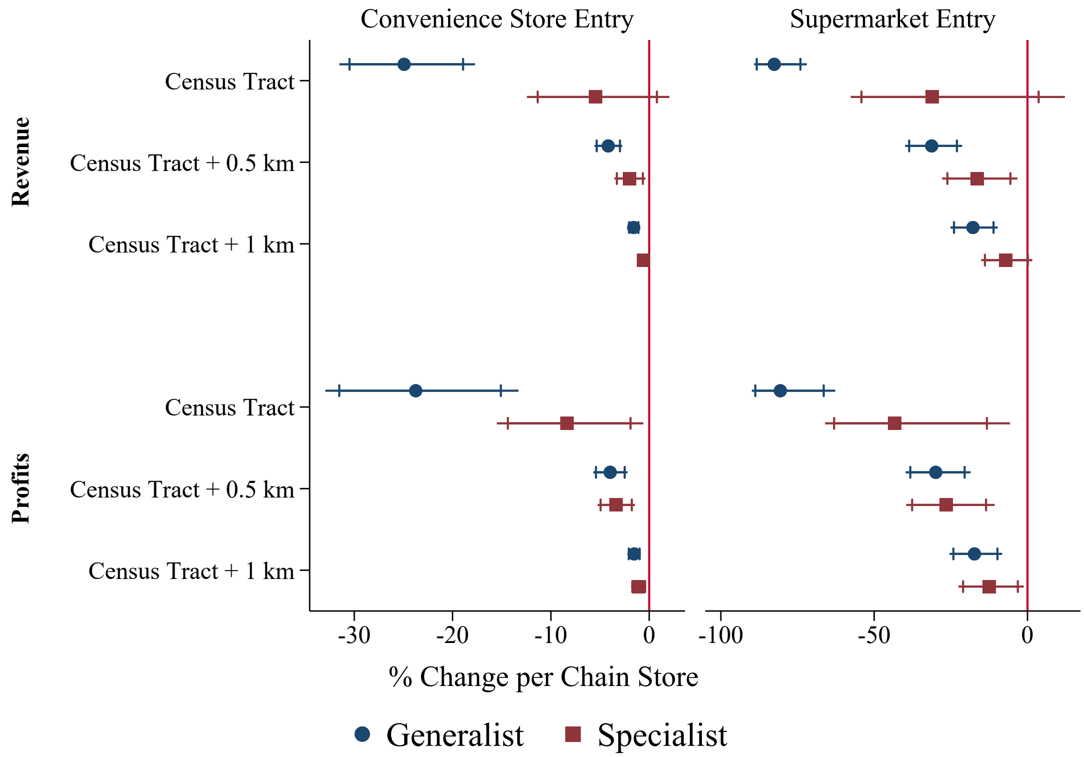
■ **Average profits and revenue.** The entry of convenience chain stores and supermarkets lowers the average profits and revenues of food retailers in the traditional sector (Figure 6).

13. See Online Appendix C of Talamas Marcos (2024) for details on this analysis.

14. If city-wide phenomena determined chains' entry, chains would expand in the same cities at the same time. They do not.

The decline is significantly larger for general stores. An additional convenience chain store in the census tract entails a reduction in average profits and revenue of general stores of almost 30%, while specialty stores suffer a decline of close to 10%. The supermarket entry has a similar pattern, but the magnitude of the effect is much more prominent. General stores average profits and revenue decline by 80%, while specialty stores experience a decline between 40% and 50%. As expected, these effects shrink as the market size increases because the average distance to the entrant increases.

Figure 6
Effect of the Entry of a Large Chain



Note: The figure displays the effect of the entry of a convenience chain store or a supermarket chain on the revenues and profits of traditional general and specialty stores, estimated using equation 4 using 2SLS. The estimates are for three market sizes: census tracts, census tracts within 0.5 km, and census tracts within 1 km. Standard errors are clustered at the municipality level. All coefficients are displayed with their 90% and 95% confidence interval.

■ **Market specialization.** The entry of a convenience chain store or a supermarket increases the degree of specialization in the traditional sector (Table 2). In particular, the percentage of specialty stores increases by 4.7 percentage points (pp; 17%) with the entry of a convenience chain store and 29 pp (104%) with the entry of a supermarket in the census tract. The main driver of the effect is a steep decline in the number of general stores, which does not occur for the number of specialty stores. The number of general establishments declines by 2.5 with the entry of a convenience chain store and by 15 with that of a supermarket. On the other hand, the number of specialty stores does not decline; it even increases. We discuss this increase in more detail later when analyzing the effects on entry and exit.

Table 2
Entry of Large Chains and Specialization in the Traditional Channel

	Convenience Chain Entry				Supermarket Chain Entry			
Dependant Variable:	% Specialized (N)	Generalist Stores (N)	Specialized Stores (N)	% Spec. Revenue	% Specialized (N)	Generalist Stores (N)	Specialized Stores (N)	% Spec. Revenue
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Chain Stores	4.73***	-2.49***	0.88*	5.72***	29.11***	-15.34***	5.43**	35.21***
	(0.73)	(0.38)	(0.45)	(0.86)	(5.79)	(2.12)	(2.67)	(6.28)
Observations	160,103	160,103	160,103	160,038	160,103	160,103	160,103	160,038
Year x Mun. FE	Y	Y	Y	Y	Y	Y	Y	Y
Market FE	Y	Y	Y	Y	Y	Y	Y	Y
Mean Dep. Var	27.9	12.8	8.0	32.1	27.9	12.8	8.0	32.1
Mean Ch. Stores	0.3	0.3	0.3	0.3	0.1	0.1	0.1	0.1
KP-F Statistic	99.8	99.8	99.8	99.8	55.0	55.0	55.0	55.1

Note: The table displays the effect of the entry of a convenience chain store (columns 1-4) or a supermarket (columns 5-8) at the census tract level, estimated based on equation 4 using 2SLS. The % Specialty (N) is the number of specialty stores divided by the total general and specialty stores in the traditional sector in the census tract. The % Spec. Revenue is the total revenue of specialty stores divided by the total revenue of general and specialty stores in the traditional sector in the census tract. Standard errors in parentheses are clustered at the municipality level. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

As shown in Figure 6, the entry of a supermarket has a larger effect on the traditional sector than the entry of a convenience chain. This is consistent with the idea that a larger entrant would steal more business from the incumbents. This does not necessarily imply that the supermarket would also have a larger effect on market-level specialization (i.e., the ratio of specialty to the sum of specialty and general). For example, if the percentage effect for both specialty and general stores were the same, the ratio of specialty over the total of specialty plus general would remain unchanged, implying a null impact on specialization. Therefore, our specialization effects capture the asymmetry in the effects on specialty and general stores. In our model, the entrant's size, z, matters—Proposition 1 and Figure 2 in Section 2 show that bigger entrants lead to more specialization.

In particular, the model predicts that the asymmetry of the effects on specialty and general stores widens the larger the entrant is. This is indeed what we observe empirically when comparing the effect on specialization of the entry of a convenience chain store and a supermarket (Columns 1 and 5 of Table 7). The entry of a supermarket increases the share of specialty retail in the traditional sector six times more than the entry of a convenience chain store. Overall, the market becomes more specialized in a broad range of measures. Table 2 shows that the share of specialty establishments and the share of revenue in specialty establishments within the traditional sector increases. Similarly, the share of profits, value-added, employment, and hours worked of specialty establishments in the traditional sector increases (Table A.4).

Consistent with our estimation based on the Economic Censuses, estimates based on household income and expenditure surveys indicate that the entry of a convenience chain store leads to an increase in the specialization of consumption within the traditional retail sector. Table 3 shows that each entry of a convenience chain store in the market (census tracts within 1 km) increases the share of consumption in specialty stores by 0.4 pp and the share of trips to specialty stores by 0.4 pp. These estimates are robust to alternative market definitions, such as all census tracts within 1.5 and 2 km, and as expected, the magnitude of the estimates decline with larger market sizes (Table A.7).

Table 3
Entry of Large Chains and Specialty's Share of Entry and Exits

Dependant Variable: Specialty Share of Entries and Exits				
	Convenience Chain Entry		Supermarket Chain Entry	
	Entry	Exit	Entry	Exit
	(1)	(2)	(3)	(4)
Number of Chain Stores	4.25***	-2.62*	26.17***	-13.63*
	(1.12)	(1.47)	(6.78)	(7.56)
Observations	148,582	101,470	148,582	101,470
Year x Mun. FE	Y	Y	Y	Y
Market FE	Y	Y	Y	Y
Mean Dep. Var	35.4	32.7	35.4	32.7
Mean Ch. Stores	0.3	0.3	0.1	0.1
KP-F Statistic	101.9	88.9	49.3	60.2

Note: The table displays the effect of the entry of a convenience chain store (columns 1 and 2) or a supermarket (columns 3 and 4) at the census tract level, on the share of entry and exit of specialty stores in the traditional sector, estimated based on equation 4 using 2SLS. The specialty share of entries is the number of entries from specialty stores divided by the total number of entries from specialty stores and general stores in the census tract. The specialty share of exits is the equivalent measure for the exits. Standard errors in parentheses are clustered at the municipality level. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

We now analyze the effect of a chain's entry on the degree of specialization through the lens of firm dynamics. We find that the entry of a chain results in an increased likelihood of traditional sector exits being by general stores and entries being by specialty stores. Table 4 shows that the entry of a chain leads to an increase in the probability of an entrant being a specialty store by 4.2 pp with an additional convenience chain store and by 26 pp with an additional supermarket. Exit also contributes to the market's specialization. The entry of a chain leads to a decrease in the probability of a store closure being a specialty store by 2.6 pp with an additional convenience chain store and by 13 pp with an additional supermarket.

These estimates aim to illustrate the effects through business dynamics. Still, they could also include the effects through within-firm specialization, because when a generalist pivots into a specialty store, it may be recorded as an exit and an entry in the data. Section 6 investigates and isolates the effect on within-firm specialization by leveraging our two waves of data collection of retail establishments in Mexico City.

■ **Robustness.** The finding of the entry of a large chain leading to an increase in specialization in the traditional sector is robust to alternative market specifications, instrument constructions, and classifications of specialized establishments. Table A.5 displays the estimates of the effect of the entry of a convenience chain store on the share of specialized establishments within the retail sector with different market sizes: census tract (baseline specification), census tracts within 0.5 km, 1 km, 1.5 km, and 2 km. All of these specifications lead to the same conclusion; however, as expected, the effect decreases as the definition of the market size increases because the average distance to the entrants increases. Table A.3 replicates Table 2 but using third-degree neighboring cities instead of second-degree neighboring cities to construct the instrument. The estimates are similar and consistent. Table A.6 uses two alternative classifications of specialty stores: establishments within retail sale of groceries and food (4-digit classification 4611) and establishments within beverages, ice, and tobacco (4-digit classification 4612). The results are robust to these alternative classifications, with estimates slightly larger for the first and slightly smaller for the latter.

6. Within-firm evidence of specialization

We leverage our two waves of data collection on retail establishments in Mexico City to provide within-firm evidence that upon entry by a large chain, surviving local stores optimally shift towards specialization. We particularly focus on testing Corollary 1 of our theory, which implies that the shift toward specialization should be more pronounced for local stores that are more capacity-constrained (small k, per the model). We first describe our empirical strategy. We then present empirical evidence demonstrating that local stores respond to chain entry with assortment changes explicitly driven by competition or customer demand, in line with our theory. Conversely, we document null effects for assortment changes reported for placebo reasons unrelated to our theoretical mechanisms (e.g., supplier issues or input cost changes). We further show that these assortment adjustments predominantly occur in small, capacity-constrained stores (below-median number of SKUs), consistent with the theoretical emphasis on relative capacity advantages. Finally, we introduce an LLM-based specialization index to explicitly demonstrate that assortment changes in response to large chain entry are in the direction of greater specialization.

6.1. Empirical strategy

Our primary dataset, collected within Mexico City, precludes employing the instrumental variables strategy described in Section 5, which relies on variation across municipalities. Instead, we adopt a first-differences approach, exploiting within-store variation in product assortments between baseline ($t = 1$) and endline ($t = 2$). This strategy controls for unobserved, time-invariant store characteristics that could be correlated with chain entry decisions. Given the brief 18-month interval between baseline and endline, we are less concerned about market-specific time-varying shocks, as discussed in Section 5. Moreover, our main empirical objective is to test Corollary 1 by comparing the differential effects of large chain entry between small k (capacity-constrained) and large k (less capacity-constrained) local stores. This focus on *relative* effects further mitigates concerns about market-specific time-varying shocks that affect all local stores, which would be virtually canceled out when we compare the impact on capacity-constrained and less capacity-constrained local stores.

We estimate the following regressions at the store level:

$$\Delta Assortment_{i,t=2} = \alpha + \beta\, ChainEntry_{i,(t=1,t=2)} + X'_{i,t=1}\gamma + \varepsilon_i \tag{6}$$

where $\Delta Assortment_{i,t=2}$ measures changes in store i's assortment (e.g., product additions and removals explicitly for competition or demand reasons) between baseline ($t = 1$) and endline ($t = 2$). $ChainEntry_{i,(t=1,t=2)}$ is an indicator equal to 1 if a large chain store (supermarket or convenience chain) entered within a defined radius (1 km in our standard specification) of store i between baseline and endline; and 0 otherwise. Stores with no entry within 1 km serve as the control group. $X_{i,t=1}$ is a vector of baseline store-level covariates, including store revenue, managerial practices, and formal registration status, used to improve precision. Our key parameter is β, reflecting the impact of chain entry on assortment adjustments.

We consider various measures of assortment change as dependent variables. First, we measure whether stores report adding new products specifically due to competition or customers as a binary dependent variable. Second, we measure whether stores report removing existing products, again explicitly for customer or competition reasons. As placebo tests, we separately examine product additions and removals reported for reasons unrelated to competition or customers (e.g., supplier issues, margin adjustments), expecting null results.

Third, we construct a specialization index leveraging our text data on what products were added and removed, if any. Specifically, we use a Large Language Model (LLM) to assign each product reported by store managers into hierarchical product clusters corresponding to categories. Then, assortment changes can be compared against the primary categories stocked by the store at baseline (t=1) per their self-description and the categories of their top products by volume. Using this intuition, we compute a specialization measure as the log ratio of *assortment depth*—number of products divided by number of categories stocked—at endline (t=2) versus baseline (t=1). This continuous measure increases, for example, when the products added by store managers belong to the same categories as those that were reported as the primary categories at baseline, as well as when products removed belong to different categories than those reported as the primary categories at baseline. When computing the assortment depth, there are two intermediate steps: (i) we need to specify the level of category granularity and thus, to obtain a robust measure, we repeat the assortment depth calculation under five granularity levels (prompting the LLM to impose between 10 to 100 product categories in the data), and (ii) we finetune using a standard few-shot learning approach where we manually assign products to categories for a small subset of observations. We normalize the final index such that coefficients can be interpreted in terms of standard deviations.

6.2. Results

Table 5 reports estimates of equation 5 on binary measures of assortment changes made due to competition or customers (Columns 1 to 3) versus placebo reasons (Columns 4 to 6). We find large, positive, and statistically significant effects of chain entry on assortment changes for reasons related to competition or customers. For instance, the probability that such a product removal was made increases by 9 percentage points, corresponding to a 42.8% relative increase over the likelihood that a control store made such a removal (base probability of 0.21). The probability that a product was added for reasons of competition or customers increases by 11 pp, again for a large relative effect of 31.4%. In contrast, we do not observe these effects of large chain entry on the placebo assortment changes. This pattern of results is consistent with store managers being responsive to large chain entry when it comes to assortment adjustments. In the appendix, we show that these results are not sensitive to the definition of the trade area—they hold when we consider large chain entries within 2 km (Table A.8) and 0.5 km (Table A.9).

Table 6 supports Corollary 1 in showing that chain entry drives assortment changes in

Table 4
Entry of Large Chains and Within-Firm Assortment Changes

	Assortment Changes: Competition/Customers			Assortment Changes: Placebo Reasons		
Dependant Variable:	Removed	Added	Changed	Removed	Added	Changed
	(1)	(2)	(3)	(4)	(5)	(6)
Chain Store Entered	0.09**	0.11***	0.12***	-0.01	-0.07**	-0.05
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Observations	554	554	554	554	554	554
Store Controls	Y	Y	Y	Y	Y	Y
Mean Dep. Var	0.21	0.35	0.48	0.24	0.27	0.43

Note: The table displays the assortment changes within stores resulting from the entry of large chains (within a 1 km radius), estimating equation 5 using a linear probability model with observations at the store level. Variation across columns denote diverse outcomes, respectively: assortment addition due to competition or customers, removal due to competition or customers, change due to competition or customers, addition due to other (placebo) reasons, removal due to other (placebo) reasons, and change due to other (placebo) reasons. The indicated regressions include store characteristic controls for precision (monthly profits, total employees, number of weekly customers, tax registration status, and an index of business practices). Robust standard errors are in parentheses. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

stores that are more capacity-constrained. The median store in our data carried 60 to 80 SKUs at baseline, and thus we define a capacity-constrained store (small k) as one with up to 60 SKUs, in line with our model's conceptualization of k. Table 6 reports estimates of equation 5 for small k versus large k stores. We find null effects, with point estimates close to zero, of chain entry on the likelihood of assortment changes for large k stores. The positive effects of chain entry on specialization are concentrated in the small k stores. Again, this pattern of results is not sensitive to the definition of trade area, holding firm when we consider large chain entries within 2 km (Table A.10) and 0.5 km (Table A.11).

Finally, using our LLM-based specialization index, we explicitly document that the stores facing nearby chain entry significantly increase their specialization relative to baseline. These results are presented in Table 7. The estimated effect of chain entry on our normalized specialization index is 0.16 SD. This effect is particularly pronounced among capacity-constrained local stores, precisely as Corollary 1 outlines. Column (2) shows the effect is statistically significant and positive for small k stores (0.23 SD), while Column (3) shows null effects with point estimates close to zero for large k stores. These results are also robust to our definition of trade area, applying when we consider large chain entries within 2 km (Table A.12) and 0.5 km (Table A.13).

In sum, our within-store findings sharply support our theoretical prediction that local stores strategically respond to increased competition from large chains by shifting assortments towards greater specialization, particularly among capacity-constrained stores.

7. Conclusion

How do small neighborhood stores survive in a retail world increasingly dominated by giants such as Walmart and Amazon? Leveraging confidential nationwide microdata from the Mexican Statistics Institute (INEGI) and original primary surveys from hundreds of

Table 5
Entry of Large Chains and Within-Firm Assortment Changes by Store Capacity Constraints

	Assortment Changes: Competition/Customers					
	More Capacity-Constrained (small k)			Less Capacity-Constrained (large k)		
Dependant Variable:	Removed	Added	Changed	Removed	Added	Changed
	(1)	(2)	(3)	(4)	(5)	(6)
Chain Store Entered	0.15***	0.20***	0.22***	-0.03	-0.001	-0.03
	(0.04)	(0.05)	(0.05)	(0.07)	(0.07)	(0.07)
Observations	345	345	345	209	209	209
Store Controls	Y	Y	Y	Y	Y	Y
Mean Dep. Var	0.15	0.38	0.46	0.19	0.2	0.32

Note: The table displays the assortment changes within stores resulting from the entry of large chains (within a 1 km radius), estimating equation 5 using a linear probability model with observations at the store level. Variation across columns denote diverse outcomes and subgroups, respectively: assortment addition due to competition or customers for more capacity-constrained stores, removal due to competition or customers for more capacity-constrained stores, change due to competition or customers for more capacity-constrained stores, addition due to competition or customers for less capacity-constrained stores, removal due to competition or customers for less capacity-constrained stores, and change due to competition or customers for less capacity-constrained stores. Robust standard errors are in parentheses. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table 6
Entry of Large Chains and Within-Firm Specialization Changes by Store Capacity Constraints

Dependant Variable: Fine-tuned Specialization Index			
	Full Sample	More Capacity-Constrained (small k)	Less Capacity-Constrained (large k)
	(1)	(2)	(3)
Chain Store Entered	0.16*	0.23**	0.01
	(0.09)	(0.11)	(0.14)
Observations	553	343	210
Store Controls	Y	Y	Y
Mean Dep. Var	0.66	0.63	0.74

Note: The table displays the specialization within stores resulting from the entry of large chains (within a 1 km radius), estimated using OLS regressions with observations at the store level. The dependent variable is the specialization index described in Section 6.1. Variation across columns denote different subgroups, respectively: the full sample, more capacity-constrained stores, and less capacity-constrained stores. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

small, offline retail stores, we document a key strategy adopted by these stores: defensive specialization. Our main finding is consistent with descriptive evidence in other contexts, such as highly specialized offline book retailers in developed countries where they face online competition (Alter and Harris, 2022).

Specialization in our analysis comes at a cost: by focusing on a product category that only appeals to a few consumers, a specialty store loses a majority of its potential customers.

However, as large stores become more dominant, for small local stores, specialization is a price worth paying: it is better to strongly appeal to some consumers and be ignored by others than to leave all consumers lukewarm. This conclusion is robust to (and, in fact, strengthened by) a variety of theoretical extensions, including endogenous prices and local store competition. We identify a "boutique effect," whereby specialty stores give up market shares but price high enough to more than compensate for their reduced sales.

While our empirical work focuses on the specific context of Mexican retail, the results are related to some important ideas on the impact of e-commerce on local stores and consumers alike. Waldfogel's (2007) "tyranny of the majority" central claim is that, in the presence of substantial fixed costs, local businesses will disproportionately serve the majority of consumers. Taste minorities, in other words, are the main benefactors of large firms. Along these lines, Choi and Bell (2011) show that geographic variation in preference minority status of target customers explains geographic variation in online sales. Contrary to this view, we argue that, in a world in which chains and large firms are dominant, more and more local stores will find it optimal to specialize in narrow niches, forgoing a majority of potential consumers but capturing higher market shares in their domain of specialization. Thus, in equilibrium, at least some taste minorities will be well served by small retailers. Nevertheless, we share Waldfogel (2007)'s intuition that specialization is insufficient: consumer surplus would increase if more local stores specialized.

Regarding online retail, our findings also allow us to rethink, and qualify, the celebrated long tail theory of Anderson (2004), and to add two novel elements to it: first, while the online long tail has been shown to grow longer over time (Anderson (2006), Brynjolfsson, Hu, and Smith (2010)), we argue that it is unclear whether it is growing *relatively* longer than its offline equivalent, contrary to Anderson's central claim. That is, unlike Anderson, we do not believe simply stocking bestselling items is a viable strategy for smaller local retailers, unlike before the entry of the large competitor. In this sense, while Anderson's large firm's long tail is predicated on the lack of capacity constraints, our small retailer one builds exactly on their presence.

We conclude by noting that our paper is about how local stores "limit the damage", not about how they thrive. Specialization allows stores to partly insure themselves from competition with larger alternatives, but it need not prove a viable long-term strategy, especially as optimal niches will become narrower and narrower.

References

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell (2019), "Food deserts and the causes of nutritional inequality," *The Quarterly Journal of Economics*, 134, 1793–1844.

Alter, Alexandra and Elizabeth A. Harris (2022), "Some Surprising Good News: Bookstores Are Booming and Becoming More Diverse," The New York Times, July 10.

Anderson, Chris (2004), "The Long Tail," *Wired Magazine*, October 1.

Anderson, Chris (2006), *The Long Tail: Why the Future of Business Is Selling Less of More*, Hachette UK.

Arcidiacono, Peter, Paul B Ellickson, Carl F Mela, and John D Singleton (2020), "The competitive effects of entry: Evidence from supercenter expansion," *American Economic Journal: Applied Economics*, 12, 175–206.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen (2017), "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan*," *The Quarterly Journal of Economics*, 132, 1101–1164.

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

Brynjolfsson, Erik, Yu Jeffrey Hu, and Michael D Smith (2010), "The Longer Tail: The Changing Shape of Amazon's Sales Distribution Curve," SSRN 1679991.

Busso, Matías, O. Fentanes, and S. Levy (2018), "The Longitudinal Linkage of Mexico's Economic Census 1999-2014," *IDB Technical Note ; 1477*.

Caoui, El Hadi, Brett Hollenbeck, and Matthew Osborne (2024), "Dynamic Entry & Spatial Competition: An Application to Dollar Store Expansion," *Available at SSRN*.

Choi, Jeonghye and David R Bell (2011), "Preference Minorities and the Internet," *Journal of Marketing Research*, 48, 670–682.

Datta, Sumon and K Sudhir (2011), "The agglomeration-differentiation tradeoff in spatial location choice," *manuscript (Yale School of Management, Yale University, New Haven, Connecticut, USA)*.

David, H. A. (1997), "Augmented Order Statistics and the Biasing Effect of Outliers," *Statistics & Probability Letters*, 36, 199–204.

de Mel, Suresh, David McKenzie, and Christopher Woodruff (2008), "Returns to Capital in Microenterprises: Evidence from a Field Experiment*," *The Quarterly Journal of Economics*, 123, 1329–1372.

Economic Census (2019), "Censos Económicos 2019," `https://www.inegi.org.mx/programas/ce/2019/`, accessed: 2025-07-02.

Ellickson, Paul B, Paul LE Grieco, and Oleksii Khvastunov (2020), "Measuring competition in spatial retail," *The RAND Journal of Economics*, 51, 189–232.

ENIGH (2006), "Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022," Accessed: 2025-04-17.

Hollenbeck, Brett, Sylvia Hristakeva, and Kosuke Uetake (2024), "Retailer Price Competition and Assortment Differentiation: Evidence from Entry Lotteries," .

Holmes, Thomas J. (2011), "The Diffusion of Wal-Mart and Economies of Density," *Econometrica*, 79, 253–302.

Hsieh, Chang-Tai and Benjamin A. Olken (2014), "The Missing "Missing Middle"," *Journal of Economic Perspectives*, 28, 89–108.

Igami, Mitsuru (2011), "Does Big Drive out Small?" *Review of Industrial Organization*, 38, 1–21.

INEGI (2025), "Sistema de Clasificación Industrial de América del Norte (SCIAN)," Accessed: 2025-07-02.

Jia, Panle (2008), "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 76, 1263–1316.

Mankiw, N. Gregory and Michael D. Whinston (1986), "Free Entry and Social Inefficiency," *The RAND Journal of Economics*, 17, 48–58.

McKenzie, David (2020), "Small Business Training to Improve Management Practices in Developing Countries," *Development Research*.

McKenzie, David and Anna Luisa Paffhausen (2019), "Small Firm Death in Developing Countries," *The Review of Economics and Statistics*, 101, 645–657.

McKenzie, David and Christopher Woodruff (2017), "Business Practices in Small Firms in Developing Countries," *Management Science*, 63, 2967–2981.

Raffaelli, Ryan (2020), "Reinventing Retail: The Novel Resurgence of Independent Bookstores," HBS Working Paper 20-068, Harvard University.

Saxena, Jaya (2022), "The Coolest Place to Drink Is Your Local Bookstore," *Eater NY*, January 13.

Talamas Marcos, Miguel Ángel (2024), "Surviving Competition: Neighbourhood Shops versus Convenience Chains," *The Review of Economic Studies*, 92, 553–585.

Waldfogel, Joel (2007), *The Tyranny of the Market: Why You Can't Always Get What You Want*, Harvard University Press, Cambridge, Mass.

APPENDICES

A. Proofs

Proof of Proposition 1: Part (a): Consider the case of a general store. For a x (or y) consumer, visiting b yields expected value

$$\widetilde{w} + m(k/2)$$

By contrast, buying at a yields expected value

$$m(z/2)$$

given that half of the total products correspond to product category x (or y). The indifferent buyer is characterized by

$$\widetilde{w} = m(z/2) - m(k/2)$$

whenever $m(z/2) - m(k/2) < w$. (Otherwise, every consumer strictly prefers seller a and b makes zero profits.) Finally, b's expected profit (when strictly positive) is given by

$$\pi_g = 1 - \big(m(z/2) - m(k/2)\big)/w \tag{7}$$

Consider now the case of a store specializing in product category x. For an x consumer, visiting b yields expected value

$$\widetilde{w} + m(k)$$

For a y consumer, the value of the x specialty store is zero. As before, buying at a yields expected value

$$m(z/2)$$

both for x and for y consumers. The indifferent x buyer is now characterized by

$$\widetilde{w} = m(z/2) - m(k)$$

whenever $m(z/2) - m(k) < w$. (Otherwise, every consumer strictly prefers seller a and b makes zero profits.) Finally, b's expected profit (when strictly positive) is given by

$$\pi_s = \tfrac{1}{2}\Big(1 - \big(m(z/2) - m(k)\big)/w\Big) \tag{8}$$

(Note that, by specializing, b expects to make, at most, $\frac{1}{2}$ in sales. This is because it will have lost all potential consumers from the product category it did not specialize in.)

If $z = 0$, that is, if the chain is out of the picture, then being a general store is trivially a dominant strategy: the store sells to a measure 1 of consumers, whereas the specialty store sells to a measure $\frac{1}{2}$ only (at the same price). Specifically, a general store's profits are equal to 1, the highest value possible, while a specialty store would only achieve its upper bound, $\frac{1}{2}$.

At the opposite end, let z_g is such that $\left(m(z_g/2) - m(k/2)\right)/w = 1$. For $z = z_g$, we have $\pi_g = 0$, whereas

$$\pi_s = \tfrac{1}{2}\left(1 - \big(m(z_g/2) - m(k)\big)/w\right) > \tfrac{1}{2}\left(1 - \big(m(z_g/2) - m(k/2)\big)/w\right) = 0$$

Such an z will exist whenever $\lim_{z\to\infty}\big(m(z/2) - m(k/2)\big)\,/w > 1$, which is implied by the condition in the Proposition. (As mentioned in the text, if this condition does not hold — for instance because w or k are very large, or $m(n)$ is very flat —, then it may always be optimal for the store to be general.)

Given continuity of π_g and π_s, it follows from the intermediate value theorem that there exists an $z_{gs} \in (0, z_g)$ such that $\pi_g(z_{gs}) = \pi_s(z_{gs})$, where for notational simplicity we have suppressed the store profit's dependence on k and w. To show that z_{gs} is unique we note that

$$\frac{d\,(\pi_s - \pi_g)}{d\,z} = \big(-m'(z/2) + 2\,m'(z/2)\big)\,/(4\,w) = m'(z/2)/(4\,w) > 0 \tag{9}$$

where the inequality follows from the fact that $m(z)$ is strictly increasing for every z. This concludes the first part of the proof.

To show that $z_{gs}(k, w)$ increases in k and w, we compute the derivative of the profit difference $(\pi_s - \pi_g)$ with respect to k and w:

$$\frac{\partial\,(\pi_s - \pi_g)}{\partial\,k} = \frac{m'(k)}{2\,w} - \frac{m'(k/2)}{2\,w} = \frac{1}{2\,w}\,\big(m'(k) - m'(k/2)\big) < 0 \tag{10}$$

where the inequality follows from concavity of m (David, 1997). Similarly,

$$\frac{\partial\,(\pi_s - \pi_g)}{\partial w} = \frac{m(z/2) - m(k)}{2\,w^2} - \frac{m(z/2) - m(k/2)}{w^2} = (\tfrac{1}{2} - \pi_s)/w - (1 - \pi_g)/w$$

where the second equality follows from (12) and (13). By definition, $\pi_s = \pi_g = \overline{\pi}$ at $z = z_{gs}$. It follows that

$$\left.\frac{\partial\,(\pi_s - \pi_g)}{\partial w}\right|_{z\,=\,z_{gs}} = (\tfrac{1}{2} - \overline{\pi})/w - (1 - \overline{\pi})/w = -1/(2\,w) < 0 \tag{11}$$

By the implicit function theorem,

$$\frac{\partial\,z_{gs}(k, w)}{\partial k} = -\frac{\partial\,(\pi_s - \pi_g)\,/\,\partial\,k}{\partial\,(\pi_s - \pi_g)\,/\,\partial\,z} > 0$$

where the inequality follows from (14) and (10). Also by the implicit function theorem,

$$\left.\frac{\partial\,z_{gs}(k, w)}{\partial w}\right|_{z\,=\,z_{gs}} = -\frac{\partial\,(\pi_s - \pi_g)\,/\,\partial\,w\,|_{s\,=\,z_{gs}}}{\partial\,(\pi_s - \pi_g)\,/\,\partial\,z} > 0$$

where the inequality follows from (14) and (11).

Part (b): We have that

$$\frac{\partial(\pi_g - \pi_s)}{\partial k} = \frac{1}{2}m'(k/2) - \frac{1}{2}m'(k) > 0$$

by concavity of k. Moreover, we know that, as $k \to z$, $\pi_g \to 1$, $\pi_s \to 1/2$, and thus $k_g > k_s$ whenever k is large enough.

Conversely, we know that $\pi_g = 0$ whenever $m(z/2) - m(k/2) \geq w$, while $\pi_s = 0$ whenever $m(z/2) - m(k) \geq w$. Denote by k_g^* and k_s^* the two values of k that satisfy these two with equality. Because both expressions are decreasing in k, these exist and are non-negative if and only if $m(z/2) \geq w$, which is implied by the condition in the proposition.

Now, notice that $k_g^* = 2k_s^*$. Thus, whenever k_g^* and k_s^* are positive, we have that $k_g^* > k_s^*$ or, in other words,

$$\pi_s > \pi_g = 0, \quad \forall k \in [k_s^*,\ k_g^*].$$

Combining our observations, we have that the difference $\pi_g - \pi_s$ is negative for $k \in [k_s^*,\ k_s^*]$ and monotonically increases, becoming strictly positive for $k \to s$. Thus, there exists a unique k_{gs} such that $\pi_s(k_{gs}, s) = \pi_g(k_{gs}, s)$.

Now, we want to show that $k_g s(z, w)$ is decreasing in z and increasing in w. To do so, we appeal to the Implicit Function Theorem again, in a similar fashion as in part (a).

We have that

$$\frac{\partial k_{gs}(z, w)}{\partial s} = -\frac{\partial(\pi_g - \pi_s)/\partial s}{\partial(\pi_g - \pi_s)/\partial k} > 0$$

and

$$\frac{\partial k_{gs}(z, w)}{\partial w} = -\frac{\partial(\pi_g - \pi_s)/\partial s}{\partial(\pi_g - \pi_s)/\partial w} < 0,$$

which concludes part (b) of the proof.■

Proof of Proposition 2: Consider the case of a general store. For a x (or y) consumer, shopping at b yields expected value

$$\widetilde{w} + m(k/2)$$

By contrast, shopping at a yields expected value

$$m(z/2)$$

given that half of the total products correspond to category x (or y). The indifferent shopper is characterized by

$$\widetilde{w} = m(z/2) - m(k/2)$$

whenever $m(z/2) - m(k/2) < w$. (Otherwise, every consumer strictly prefers seller a and b makes zero profits.) Finally, b's expected profit (when strictly positive) is given by

$$\pi_g = 1 - \big(m(z/2) - m(k/2)\big)\,/w \tag{12}$$

Consider now the case of a local store specializing in category x. For an x consumer, shopping at b yields expected value

$$\widetilde{w} + m(k)$$

For a y consumer, the value of the x specialty store is zero. As before, buying at a yields expected value

$$m(z/2)$$

both for x and for y consumers. The indifferent x consumer is now characterized by

$$\widetilde{w} = m(z/2) - m(k)$$

whenever $m(z/2) - m(k) < w$. (Otherwise, every consumer strictly prefers seller a and b makes zero profits.) Finally, b's expected profit (when strictly positive) is given by

$$\pi_s = \tfrac{1}{2}\Big(1 - \big(m(z/2) - m(k)\big)\,/w\Big) \tag{13}$$

(Note that, by specializing, b expects to make, at most, $\frac{1}{2}$ in sales. This is because it will have lost all potential shoppers from the category it did not specialize in.)

If $z = 0$, that is, if large store is not active, then being a general store is trivially a dominant strategy: the store sells to a measure 1 of consumers, whereas the specialty store sells to a measure $\frac{1}{2}$ only (at the same price). Specifically, a general store's profits are equal to 1, the highest value possible, while a specialty store would only achieve its upper bound, $\frac{1}{2}$.

At the opposite end, let z_g is such that $\left(m(z_g/2) - m(k/2)\right)/w = 1$. For $z = z_g$, we have $\pi_g = 0$, whereas

$$\pi_s = \tfrac{1}{2}\,\big(1 - \big(m(z_g/2) - m(k)\big)\,/w\big) > \tfrac{1}{2}\,\big(1 - \big(m(z_g/2) - m(k/2)\big)\,/w\big) = 0$$

Such a z exists whenever $\lim_{z\to\infty}\big(m(z/2) - m(k/2)\big)\,/w > 1$, which is equivalent to Assumption 2. (As mentioned in the text, if this condition does not hold — for instance because w or k are very large, or $m(n)$ is very flat —, then it may always be optimal for the store to be a general store.)

Given continuity of π_g and π_s, it follows from the intermediate value theorem that there exists a $z_{gs} \in (0, z_g)$ such that $\pi_g(z_{gs}) = \pi_s(z_{gs})$, where for notational simplicity we have suppressed the store profit's dependence on k and w. To show that z_{gs} is unique we note that

$$\frac{d\,(\pi_s - \pi_g)}{d\,z} = \big(-\,m'(z/2) + 2\,m'(z/2)\big)\,/(4\,w) = m'(z/2)/(4\,w) > 0 \tag{14}$$

where the inequality follows from the fact that $m(z)$ is strictly increasing for every z. This concludes the first part of the proof.

It follow that, if $z > z_{gs}$, then, conditional on entry by the chain store, a local store's variable profit is higher as a specialty store. Finally, assuming that fixed cost is randomly distributed with a distribution that depends on size but not the nature of the store, the result follows. ■

Proof of Proposition 2: Assumption 1 implies that

$$\frac{\partial\,(\pi_g - \pi_s)}{\partial\,k} = \tfrac{1}{2}\,m'(k/2) - \tfrac{1}{2}\,m'(k) > 0$$

If $k = 0$, then $\pi_g = 1 - m(z/2)/w > \pi_s = \frac{1}{2}\big(1 - m(z/2)/w\big)$. Let k_s be such that

$$\pi_g = 1 - \big(m(z/2) - m(k_g/2)\big)\,/w = 0$$

When $\pi_g = 0$, that is, when $k = k_s$, we have

$$\pi_s = \tfrac{1}{2}\Big(1 - \big(m(z/2) - m(k_g)\big)\,/w\Big) > 0$$

Together, the above facts imply that there exists a $k' \in [0, k_g]$ such that, if $k < k'$, then $\pi_s > \pi_g$. which in turn implies the result. ■

Proof of Proposition 3: Suppose store b specializes in product category x, the popular product category ($\alpha > \frac{1}{2}$). Then store b reaches at most α of its potential customers. The indifferent customer (indifferent between store a and store b) has z such that

$$m(\alpha\, z) = m(k) + \widetilde{w}$$

where $\alpha\, z$ is total supply of products of product category x, all of which are available at store a; and k is the supply of products of product category x at store b (in other words, all of store b's capacity, k, is devoted to carrying product category x products). It follows that, of the k store-b potential customers, a fraction $\alpha\, k$ is interested in the product category offered by store b, and a fraction $\big(m(\alpha\, z) - m(k)\big)\,/\widetilde{w}$ of this fraction prefers store b to store a. This implies that store b's profit from specializing in product category x is given by

$$\pi_x = \alpha\Big(1 - \big(m(\alpha\, z) - m(k)\big)\,/\widetilde{w}\Big)$$

Similarly, the profit from specializing in product category y is given by

$$\pi_y = (1-\alpha)\Bigg(1 - \Big(m\big((1-\alpha)\, z\big) - m(k)\Big)\,/\widetilde{w}\Bigg)$$

If $z = 0$, that is, if the chain is out of the picture, then the popular product category x is trivially a dominant strategy: the store sells to a measure α of consumers, whereas the niche-product category store sells to a measure $1 - \alpha < \alpha$ only (and at the same price). At the opposite end, let z_x be the value of z such that $\pi_x = 0$. Such an z will exist whenever $\lim_{z\to\infty}\big(m(\alpha\, z) - m(k)\big)\,/w > 1$, which is equivalent to the condition in the Proposition. We then have

$$\pi_y = (1-\alpha)\Bigg(1 - \Big(m\big((1-\alpha)\, z_x\big) - m(k)\Big)\,/\widetilde{w}\Bigg) > \alpha\Big(1 - \big(m(\alpha\, z_x) - m(k)\big)\,/\widetilde{w}\Big) = 0$$

(If this condition does not hold — for instance because w or k are very large, or $m(n)$ is very flat —, then it may always be optimal for the store to choose the popular product category.)

Given continuity of π_x and π_y, the intermediate value theorem implies that there exists at least one value $\widehat{z}_{xy} \in (0, z_x)$ such that $\pi_g(\widehat{z}_{xy}) = \pi_s(\widehat{z}_{xy})$, where for notational simplicity we have suppressed the store profit's dependence on k and $\widetilde{w}$. Let z_{xy} be the highest of these values. Then $\pi_y \geq \pi_x$ for $z > z_{xy}$.

Consider now the comparative statics with respect to k. First notice that there are only $(1-\alpha)\, z$ products of product category y. Therefore, $m\big((1-\alpha)\, z_x\big)$ is an upper bound of the benefit from stocking only y product category products. Therefore, for $k > (1-\alpha)\, z_x$, $\pi_y = (1-\alpha)\, k$. As to π_x, we can see that it is increasing in k and, as k reaches $k = m(\alpha\, z)$, $\pi_x = \alpha\, k > \pi_y$. It follows that there exist a k_{xy} such that $\pi_x > \pi_y$ if $k > k_{xy}$. ■

Proof of Proposition 4: We first solve for the optimal prices of a general store given that store a sets p_a. Store g's profit is given by $\pi_g = p_g\, q_g$, where q_g, the store's sales, are given by

$$q_g = 1 - \big(m(z/2) - m(k/2) - p_a + p_g\big)\, /w$$

The profit-maximizing price, quantity and profit levels are given by

$$\widehat{p}_g \;=\; \tfrac{1}{2}\big(w - m(z/2) + m(k/2) + p_a\big) \tag{15}$$
$$\widehat{q}_g \;=\; \tfrac{1}{2}\big(w - m(z/2) + m(k/2) + p_a\big)\, /w \;=\; \widehat{p}_g/w \tag{16}$$
$$\widehat{\pi}_g \;=\; \widehat{p}_g\, \widehat{q}_g \;=\; (\widehat{p}_g)^2/w \tag{17}$$

In the case of a specialty store, profit is given by $\pi_s = p_s\, q_s$, where q_s, the store's sales, are given by

$$q_s = \tfrac{1}{2}\Big(1 - \big(m(z/2) - m(k) - p_a + p_s\big)\, /w\Big)$$

The profit-maximizing price, quantity and profit levels are given by

$$\widehat{p}_s \;=\; \tfrac{1}{2}\big(w - m(z/2) + m(k) + p_a\big) \tag{18}$$
$$\widehat{q}_s \;=\; \tfrac{1}{4}\big(w - m(z/2) + m(k) + p_a\big)\, /w \;=\; \widehat{p}_s/(2\, w) \tag{19}$$
$$\widehat{\pi}_s \;=\; \widehat{p}_s\, \widehat{q}_s \;=\; (\widehat{p}_s)^2/(2\, w) \tag{20}$$

Direct inspection of (15) and (18) reveals that

$$\widehat{p}_s \;>\; \widehat{p}_g$$

that is, in equilibrium specialty stores set a higher price. Moreover, from (15)–(16) and (18)–(19) we conclude that

$$\widehat{p}_s/\widehat{q}_s \;=\; 2\, w \;>\; \widehat{p}_g/\widehat{q}_g \;=\; w \tag{21}$$

Consider the extreme case when $z = 0$. Straightforward computation shows that $\widehat{\pi}_g > \widehat{\pi}_s$ if and only if the condition in the Proposition holds. At the opposite end, let z_g be such that $\widehat{p}_g = 0$. Comparing (15) and (18), we see that, at $z = z_g$, $\widehat{p}_s > \widehat{p}_g = 0$. From (17) and (20) we conclude that, at $z = z_g$, $\widehat{\pi}_s > \widehat{\pi}_g = 0$. Since both $\widehat{\pi}_s$ and $\widehat{\pi}_g$ are continuous we conclude by the intermediate-value theorem that there exists at least one $\widetilde{z}_{gs}$ such that $\widehat{\pi}_s = \widehat{\pi}_g$. Let z_{gs} be the highest of these values. Then $\widehat{\pi}_s > \widehat{\pi}_g$ when $z_{gs} < s < z_g$.

Finally, notice that, at $z = z_{gs}$, $\widehat{\pi}_g = \widehat{\pi}_s$, that is, $\widehat{p}_g\, \widehat{q}_g = \widehat{p}_s\, \widehat{q}_s$. Since, from (21), $\widehat{p}_s/\widehat{q}_s > \widehat{p}_g/\widehat{q}_g$, it must be that, at $z = z_{gs}$, $\widehat{p}_s > \widehat{p}_g$ and $\widehat{q}_s < \widehat{q}_g$. Since these are strict inequalities, they also hold in the neighborhood of $z = z_{gs}$. It follows that, in the right neighborhood of $z = z_{gs}$, a specialty store earns a higher profit, sets a higher price, and captures a lower market share. ■

Proof of Proposition 5: The proof follows straightforwardly from the definitions of π_g and π_s. Specifically, from (12) and (13) we derive

$$\begin{aligned}\frac{\partial \pi_g}{\partial w} &= \frac{m(z/2) - m(k/2)}{w^2} \\ \frac{\partial \pi_s}{\partial w} &= \frac{m(z/2) - m(k)}{2\,w^2}\end{aligned} \tag{22}$$

which implies that

$$\frac{\partial \pi_s}{\partial w} = \tfrac{1}{2}\,\frac{\partial \pi_s}{\partial w} > 0$$

Taking derivatives of (22) with respect to z, we get

$$0 < \frac{\partial^2 \pi_s}{\partial w\, \partial z} = \tfrac{1}{2}\,\frac{\partial^2 \pi_s}{\partial w\, \partial z}$$

Taking derivatives of (22) with respect to k, we get

$$0 < \frac{\partial^2 \pi_s}{\partial w\, \partial k} = \tfrac{1}{2}\,\frac{\partial^2 \pi_s}{\partial w\, \partial k}$$

Finally, taking derivatives of (22) with respect to w we get

$$\frac{d^2 \pi_g}{d w^2} < \frac{d^2 \pi_s}{d w^2} < 0$$

which concludes the proof. ■

Proof of Proposition 6: Figure 10 illustrates the competition case. On the horizontal axis we measure the consumer location d, where $d = 0$ corresponds to local store b_0 and $d = 1$ corresponds to local store b_1. On the vertical axis we measure z, the relative preference for a local store. We assume that d and z are independently and uniformly distributed: $\widetilde{d} \sim U[0,1]$ and $\widetilde{w} \sim U[0,w]$. Since there are two different product categories, we need to plot one graph per product category, product category x on the top panel and product category y on the bottom panel.

Figure 10 illustrates the case when both b_0 and b_1 are general stores. Store b_0's demand of product category x is given by the area in blue in the top panel, whereas store b_0's demand of product category y is given by the area in red in the top panel. To understand that, notice that store b_0 must beat both store a and store b_1. Beating store a requires

$$m(k/2) + z - \tau\,\widetilde{d} > m(z/2)$$

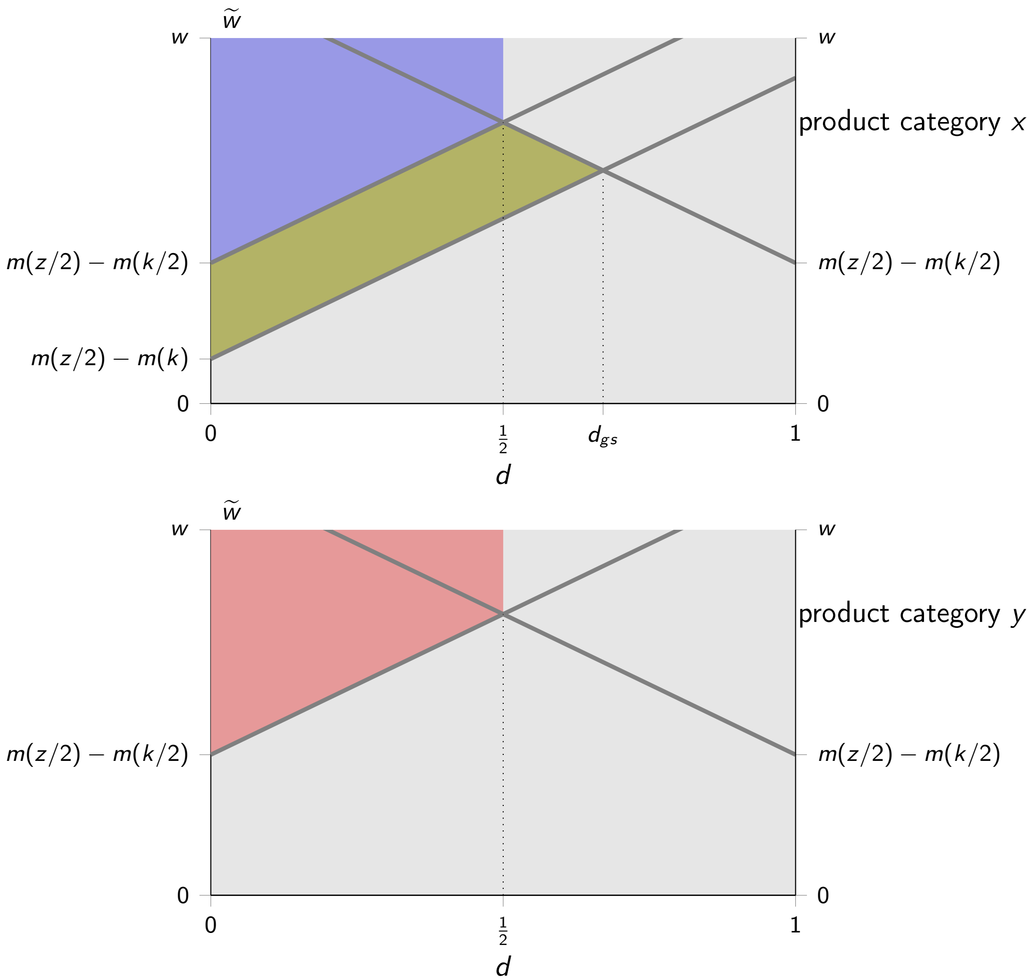
whereas beating store b_1 requires

$$m(k/2) + z - \tau\,\widetilde{d} > m(k/2) + z - \tau\,(1 - \widetilde{d})$$

This results in the following set of inequalities

$$\begin{aligned}\widetilde{w} &> m(z/2) - m(k/2) + \tau\,\widetilde{d} \\ \widetilde{d} &< \tfrac{1}{2}\end{aligned}$$

Figure 7
Store strategy under local competition



which in turn correspond to the areas in blue (top panel) and red (bottom panel).

Given that b_1 chooses to be a general store, how does b_0 change its profits by specializing in product category x? Store b_1's demand from x consumers is now determined by

$$m(k) + \widetilde{w} - \tau\,\widetilde{d} > m(z/2)$$

(beat firm a) and

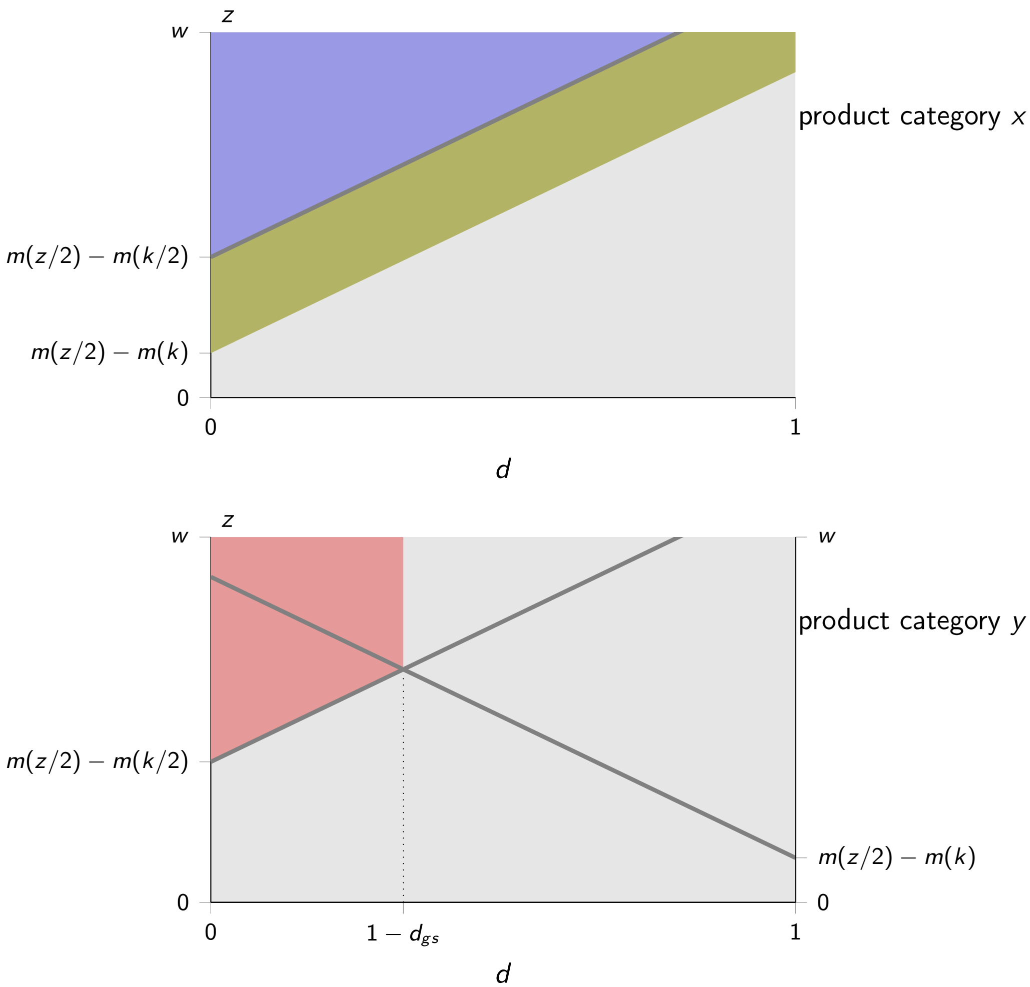
$$m(k) + \widetilde{w} - \tau\,\widetilde{d} > m(k/2) + \widetilde{w} - \tau\,(1 - \widetilde{d})$$

(beat firm b_1). This simplifies to

$$\begin{aligned} \widetilde{w} &> m(z/2) - m(k) + \tau\,\widetilde{d} \\ \widetilde{d} &< d_{gs} \equiv \tfrac{1}{2} + \big(m(k) - m(k/2)\big)\,/\tau \end{aligned}$$

This corresponds to an increase in demand for product category x given by the area in green on the top panel and a loss in demand for product category y given by the area in red

Figure 8
Store strategy under local competition



on the bottom panel. The green area on the top panel corresponds entirely to consumers who purchased from a when both b_0 and b_1 were general stores and now prefer to buy from b_0, the product category x specialty store. The red area on the bottom panel corresponds to consumers who were interested in store b_0 when it was a general store but are now not interested since it no longer carries any product category y products.

The values of z and k in Figure 10 were chosen so that the areas in green and red are equal. This implies that, given that store b_1 follows a general-store strategy, store b_0 is indifferent between being a general store and being a specialty store. Suppose now that b_1 chooses to be a y-specialty store. What is the gain for store b_0 from specializing in x? This alternative scenario is described in Figure 11. In terms of x consumers, the battle is now limited to firms b_0 and a, since firm b_1 is absent from this product category. Demand for firm b_0 is determined by

$$m(k/2) + \widetilde{w} - \tau\,\widetilde{d} > m(z/2)$$

which corresponds to the area in blue. Regarding product category y (bottom panel), we still need to consider both competition by a and competition by b_1. Since b_1 is a product category y specialty store, we now have

$$\begin{aligned} \widetilde{w} &> m(z/2) - m(k) + \tau\,\widetilde{d} \\ \widetilde{d} &< 1 - d_{gs} \equiv \tfrac{1}{2} + \big(m(k/2) - m(k)\big)\,/\tau \end{aligned}$$

which corresponds to the area in red. What happens to firm b_0's profit as it switches from a general store to a product category x specialty store? On the top panel (that is, in terms of x sales), it experiences a profit increase given by the green area. On the top panel (that is, in terms of y sales), it experiences a profit loss given by the red area.

Immediate inspection reveals that the green area in the top panel of Figure 11 is greater than the green area in the top panel of Figure 10, whereas the red area in the bottom panel of Figure 11 is lower than the red area in the bottom panel of Figure 10. This implies that, if firm b_0 is indifferent between being a general store and being a specialty store when its rival is a general store, then it strictly prefers to be a specialty store when its rival is a specialty store. ■

B. Theory extensions

The base model developed in Section 2 helped us make the point that, as the chain increases in size, local stores have a tendency to specialize in a limited number of product categories and, within this specialization strategy, have a tendency to select niche product categories. In the previous section we provided some supporting empirical evidence from traditional retail. In this section, we present a number of extensions of the basic framework introduced in Section 2.

■ **General, popular-product category, and niche-product category stores.** A natural extension of the analysis so far is to integrate the choice of general vs specialty (Proposition 1) with the analysis of product category of specialization (Proposition 3). In our initial model, we assumed two equal-sized product categories, x and y. In this context, a general store is one that stocks x and y in equal amounts, whereas a specialty store is one that stocks either only x or only y. When there are two product categories of different sizes, as in the model underlying Proposition 3, the decision of how to stock is not trivial. Suppose that a fraction α of the products (and a fraction α of the potential demand) correspond to product category x. Let β be the fraction of a general store that carries product category x. Should β be greater than, equal to, or lower than α?

Figure 9
Optimal stocking policy for general store (assuming v is uniformly distributed). α is the fraction of product category x buyers, whereas β is the fraction of product category x optimally stocked by a general store.

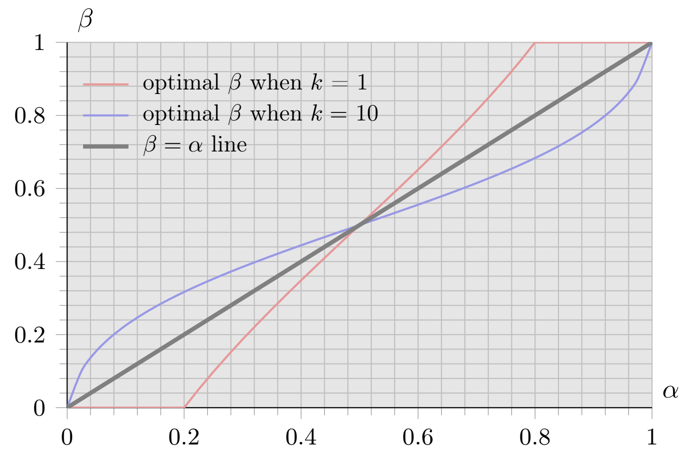


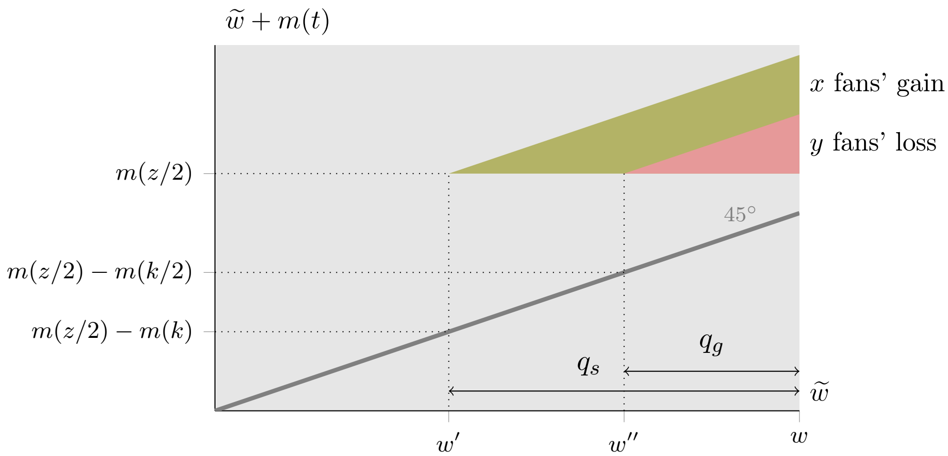
Figure 7 illustrates this decision in the case when $F = v$, and so $m(t) = t/(1+t)$. If the value of k is small ($k = 1$ in the present example), then the optimal stocking policy is to over-stock the most popular product category. This is shown by $\beta > \alpha$ for $\alpha > \frac{1}{2}$ (red line). By contrast, if the value of k is large ($k = 10$ in this example), then the optimal stocking policy is to over-stock the least popular product category. This is shown by $\beta > \alpha$ for $\alpha < \frac{1}{2}$ (blue line). Intuitively, when k is large, then the marginal value of an extra product is lower, due to concavity of $m(k)$. This is particularly true for a popular product category. Therefore, in relative terms and at the margin, the seller is better off by stocking a product of a niche product category. By contrast, if k is small, then the extensive margin effect dominates and the seller is better off by overstocking (relatively speaking) the popular product category.

Taking into account the optimal stocking strategy, Figure 3 plots the profit of a general store (as well as the profit function of a specialty store focused on a popular product category, x, or on a niche product category, xy). As can be seen, as z increases, firm b's optimal choice shifts from being a general store to being a specialty store focused on the popular product category to finally being a specialty store focused on the niche product category. In this way, Figure 3 illustrates both Proposition 1 and Proposition 3.

■ **Welfare analysis.** All of our analysis so far has focused on firm b's profits and optimal choices. A natural follow-up question is the relation between firm b's decisions and consumer welfare. Let us go back to the model with fixed prices and one local store, firm b. Let us consider, as in the initial model, the choice between being a general and being a specialty store. Suppose social welfare is given by consumer surplus plus firm profits. Since all sellers set $p = 1$ and the market is covered (all consumers make a purchase), consumer surplus is a sufficient statistic of social welfare.

Figure 9 illustrates the contrast between a general and a specialty store when competing against firm a. On the horizontal axis we measure each consumer's value of z, that is, their disutility from buying from firm a. On the vertical axis we measure the advantage, in terms

Figure 10
Firm profit and consumer welfare. Effects of switching from general to specialty x store.



of vertical quality, of the chain store with respect to the local store. The 45° line measures the points at which the "horizontal" differentiation advantage of firm b exactly compensates the "vertical" differentiation advantage of firm a.

Consider first the case of a general store b. Its disadvantage with respect to store a is given by $m(z/2) - m(k/2)$. It follows that only consumers with a value of $\widetilde{w}$ greater than w'' purchase at the local store. Since $\widetilde{w}$ is uniformly distributed, we conclude that firm b's market share is given by $q_g = w - w''$.

Consider now the case of a specialty store b. Its disadvantage with respect to store a is given by $m(z/2) - m(k)$. It follows that only consumers with a value of $\widetilde{w}$ greater than w' purchase at the local store. Since $\widetilde{w}$ is uniformly distributed, we conclude that firm b's market share (among its product category followers) is given by $q_s = w - w'$. However, we must keep in mind that if firm b focuses on product category x, for example, then it loses potential buyers who are only interested in y. In other words, by becoming a specialty store firm b halves its potential demand. Therefore, its market share is $(w - w')/2$.

The values of z and k were selected so that $\pi_g = w - w'' = (w - w')/2 = \pi_s$. In other words, for the particular values of z and k underlying Figure 9, firm b is indifferent between being a general store or being a specialty store. Consumers, however, are not indifferent between the two types of store. Consumer surplus is given by the area below

$$\max\{m(z/2), \widetilde{w} + m(\widetilde{k})\}$$

where $\widetilde{k} = k/2$ or $\widetilde{k} = k$ for a general and a specialty store, respectively. It follows that, for product category x consumers, the switch from a general to a product category x specialty store implies an increase in consumer surplus given by the green trapezoid in Figure 9. By contrast, for product category y consumers the switch implies a decrease in consumer surplus given by the red area in Figure 9. By construction, the green area is greater than the red area. More generally, we have just established the following result:

Proposition 7. *When store b is indifferent between being a general or a specialty store, the average consumer strictly prefers the latter.*

Intuitively, consumer surplus is "convex" in the vertical utility provided by the local store. This implies that consumers prefer the "bet" of having a specialty store of their preferred product category with probability 50% than a general store with probability 100%.

This intuition is related to a number of results in the IO literature. Mankiw and Whinston (1986) provide conditions such that, in equilibrium, there is excess entry into a market. Intuitively, the entrant does not correctly take into account the positive externality it creates for consumers nor the negative externality it creates for its competitors. Similarly, our firm b does not take into account the positive surplus effect it has on the consumers who like the product category in which they specialize.

ONLINE APPENDIX

A. Additional Figures and Tables: Empirical Evidence

Figure A.1
Map of N=554 Stores in Primary Dataset

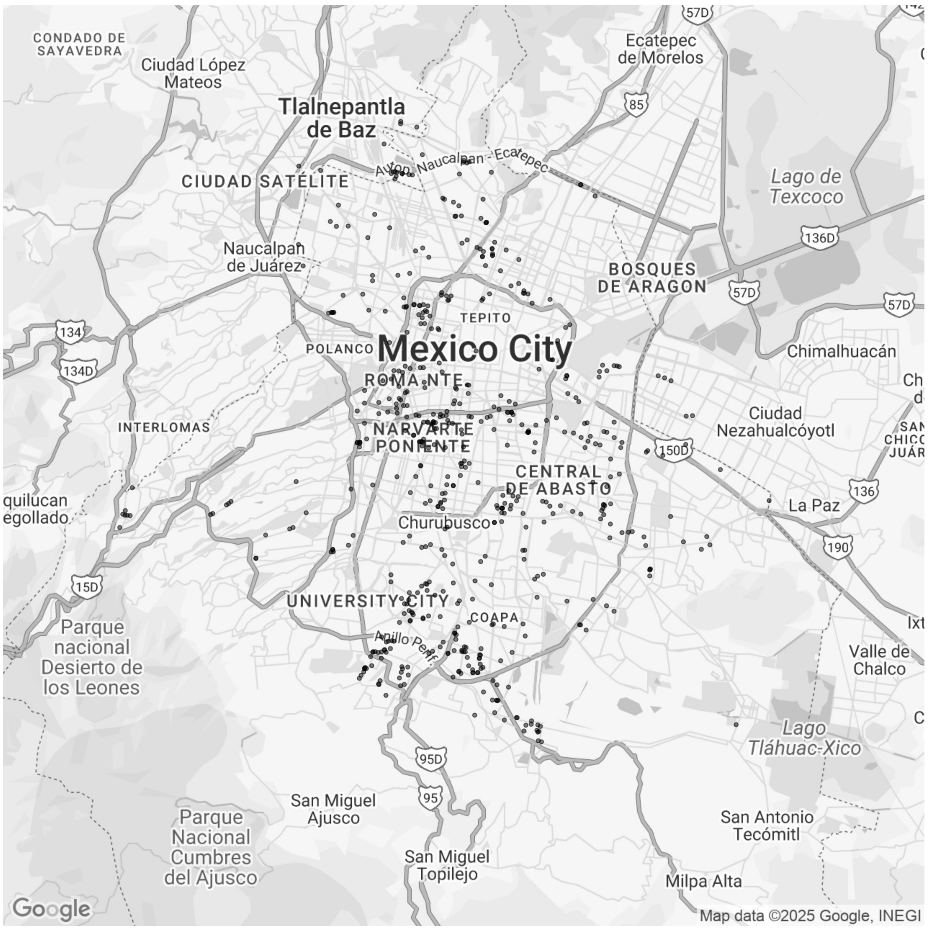


Figure A.2
General and Specialty Stores

General Stores

Miscellaneous Grocery Retailer

Specialty Stores

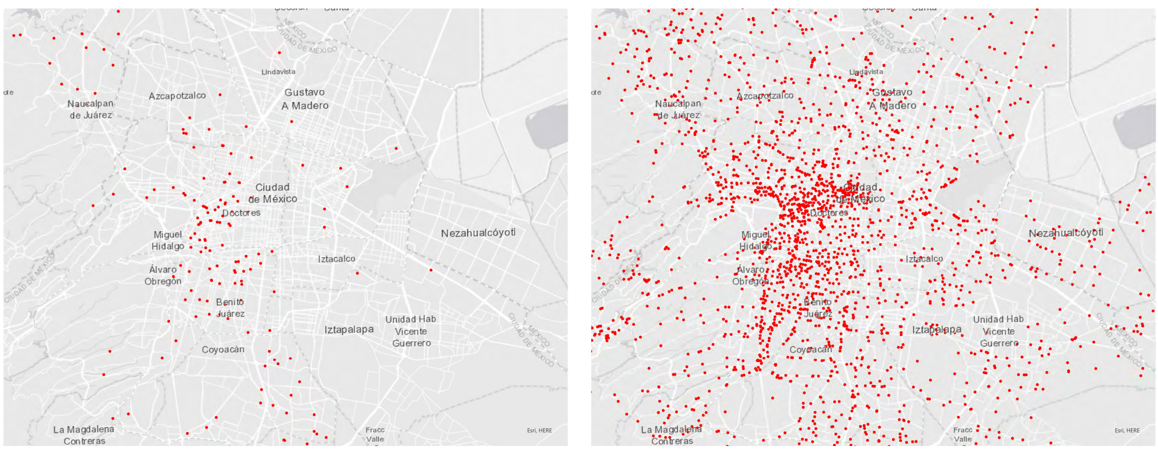
Bread and Baked Goods

Dairy and Deli Meat

Candy

Figure A.3

Convenience Chains in Mexico City



1999 2019

Source: Esri, HERE, DENUE (INEGI), Marco Geostadistico (INEGI)

Note: The maps display the location of convenience chain stores for 1999 and 2019. Locations for 1999 are approximated using the 1999 Economic Census Data. Locations for 2019 are from INEGI's business registry (DENUE).

Table A.1

Summary Statistics of Stores at Baseline (t=1)

	Mean	SD	Census 2019 SCIAN: 461	Census 2019 SCIAN: 465	Census 2019 SCIAN: 722
Monthly Revenue	59,765.78	10,485.70	60,966.01	35,700.70	67,045.12
Monthly Profits	10,449.66	16,571.73	15,029.0	5,326.05	19,130.75
Hired Employees	2.21	2.25	0.52	0.62	2.08
Total Employed	2.51	2.32	1.88	1.84	3.58

Note: The table displays summary statistics on N=554 retail stores in the sample collected at baseline at the business location. All monetary values are expressed in 2018 MXN Pesos. The statistics for the 2019 Census include small and medium-sized establishments (those that do not have a large-firm identifier in the micro data) within the stated 3-digit SCIAN classification and located in Mexico City. Profits are calculated as revenue minus reported expenses (rent, wages, cost of goods, and other costs).

Table A.2

General and Specialty Stores in 1999

Dependant Variable (log):	Revenue	Profits	Age	Total Employed	Hired Employees	Revenue x Wrkr	Profits x Wrkr
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
I[Specialist]	-0.002	-0.051***	-0.085***	-0.018***	0.148***	0.007	-0.059***
	(0.013)	(0.011)	(0.008)	(0.004)	(0.006)	(0.012)	(0.011)
Observations	719,866	719,866	719,866	719,866	719,866	717,007	708,800
Census Tract FE	Y	Y	Y	Y	Y	Y	Y

Note: The table displays the differences between general and specialty stores estimated using the following equation: $Y_f = \beta_0 + \beta_1 1[specialty]_f + \delta_{m(f)} + \varepsilon_i$, where Y_f denotes the outcome of interest in logs for firm f, β_1 is the estimated difference between specialty and general stores, and $\delta_{m(f)}$ is the census tract fixed effect. The sample does not include establishments owned by a chain or the government. Standard errors in parentheses are clustered at the municipality level. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.3
Robustness: Market Specialization (3rd Degree Neighboring Cities in Instrument)

	Convenience Chain Entry				Supermarket Chain Entry			
Dependant Variable:	% Specialized (N)	Generalist Stores (N)	Specialized Stores (N)	% Spec. Revenue	% Specialized (N)	Generalist Stores (N)	Specialized Stores (N)	% Spec. Revenue
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Chain Stores	5.05***	-2.66***	1.17***	6.05***	29.73***	-15.65***	6.89***	35.58***
	(0.70)	(0.38)	(0.44)	(0.79)	(5.31)	(2.10)	(2.46)	(5.49)
Observations	160,103	160,103	160,103	160,038	160,103	160,103	160,103	160,038
Year x Mun. FE	Y	Y	Y	Y	Y	Y	Y	Y
Market FE	Y	Y	Y	Y	Y	Y	Y	Y
Mean Dep. Var	27.9	12.8	8.0	32.1	27.9	12.8	8.0	32.1
Mean Ch. Stores	0.3	0.3	0.3	0.3	0.1	0.1	0.1	0.1
KP-F Statistic	140.5	140.5	140.5	140.6	72.7	72.7	72.7	72.7

Note: The table displays the effect of the entry of a convenience chain store (columns 1-4) or a supermarket (columns 5-8) at the census tract level, estimated based on equation 4 but using third-degree neighboring cities instead of second-degree neighboring cities to construct the instrument. The % outcomes are the share of the variable for specialized stores over the traditional sector. Standard errors in parentheses are clustered at the municipality level. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.4
Specialization: Additional Outcomes

	Convenience Chain Entry				Supermarket Chain Entry			
Dependant Variable:	% Spec. Value Added	% Spec. Jobs	% Spec. Hours	% Spec. Profits	% Spec. Value Added	% Spec. Jobs	% Spec. Hours	% Spec. Profits
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Chain Stores	6.04***	5.41***	6.16***	6.32***	37.25***	33.27***	37.93***	38.98***
	(0.94)	(0.84)	(0.86)	(0.93)	(6.00)	(6.60)	(6.73)	(6.03)
Observations	158,375	160,101	160,066	157,878	158,375	160,101	160,066	157,878
Year x Mun. FE	Y	Y	Y	Y	Y	Y	Y	Y
Market FE	Y	Y	Y	Y	Y	Y	Y	Y
Mean Dep. Var	31.3	28.9	26.0	31.4	31.3	28.9	26.0	31.4
Mean Ch. Stores	0.3	0.3	0.3	0.3	0.1	0.1	0.1	0.1
KP-F Statistic	99.7	99.8	99.8	99.7	56.8	55.0	55.1	56.5

Note: The table displays the effect of the entry of a convenience chain store (columns 1-4) or a supermarket (columns 5-8) at the census tract level, estimated based on equation using 2SLS. The % outcomes are the share of the variable for specialized stores over the traditional sector. Standard errors in parentheses are clustered at the municipality level. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.5

Entry of Large Convenience Chains, Food Expenditure and Trips in the Traditional Channel

Dependant Variable: Share of Consumption and Trips in/to Specialty Stores				
	Consumption ($)		Trips	
	(1)	(2)	(3)	(4)
Number of Conv. Chain Stores	0.42***	0.43***	0.38***	0.40***
	(0.12)	(0.12)	(0.14)	(0.14)
Observations	999,657	999,657	999,657	999,657
Year x Mun. FE	Y	Y	Y	Y
Market FE	Y	Y	Y	Y
HH Controls		Y		Y
Mean Dep. Var	44.4	44.4	47.8	47.8
Mean Ch. Stores	8.6	8.6	8.6	8.6
KP-F Statistic	117.4	117.3	117.4	117.3

Note: The table displays the effect of the entry of a convenience chain store in the market (census tracts within 1 km) estimated through 2SLS based on equation 4 at the household level. The monetary share of consumption is the expenses in specialty stores divided by the expenses in specialty stores and neighborhood shops. The share of trips is the number of days the household visited a specialty store within the week divided by the number of days it visited a specialty store plus the number of days it visited a neighborhood shop. Columns 2 and 4 include household-level controls: income per capita, income, household size, number of adults, age of household head, and monetary expenditures. Standard errors in parentheses are clustered at the municipality level. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.6

Robustness: Specialization (Alternative Market Size Definitions)

Dependant Variable: % Specialty (N)										
	Convenience Chain Entry					Supermarket Chain Entry				
	AGEB	0.5 km	1 km	1.5 km	2 km	AGEB	0.5 km	1 km	1.5 km	2 km
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Number of Chain Stores	4.73***	1.11***	0.53***	0.31***	0.20***	29.11***	9.80***	6.51***	3.95***	2.59***
	(0.73)	(0.21)	(0.10)	(0.07)	(0.05)	(5.79)	(2.77)	(2.17)	(1.45)	(0.95)
Observations	160,103	160,242	160,300	160,321	160,324	160,103	160,242	160,300	160,321	160,324
Year x Mun. FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Market FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean Dep. Var	27.9	22.3	21.6	21.3	21.2	27.9	22.3	21.6	21.3	21.2
Mean Ch. Stores	0.3	1.8	3.9	6.8	10.3	0.1	0.4	1.0	1.7	2.5
KP-F Statistic	99.8	135.8	149.4	154.6	152.7	55.0	29.6	14.5	12.3	15.9

Note: The table displays the effect of the entry of a convenience chain store (columns 1-5) or a supermarket (columns 6-10) in a given market on the specialization of the traditional channel, estimated based on equation 4 using 2SLS. The outcome is the share of specialized stores over the traditional sector. Variations across columns denote different definitions of markets (census tract, census tract within 0.5 km, 1 km, 1.5 km, and 2 km). Standard errors in parentheses are clustered at the municipality level. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.7
Robustness: Specialization (Alternative Definitions of Specialized Stores)

Dependant Variable: % Specialty (N)

	Convenience Chain Entry			Supermarket Chain Entry		
	461	4611	4612	461	4611	4612
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Chain Stores	4.73***	5.32***	3.45***	29.11***	32.70***	21.46***
	(0.73)	(0.79)	(0.49)	(5.79)	(6.24)	(3.72)
Observations	160,103	158,480	157,976	160,103	158,480	157,976
Year x Mun. FE	Y	Y	Y	Y	Y	Y
Market FE	Y	Y	Y	Y	Y	Y
Mean Dep. Var	27.9	22.3	9.7	27.9	22.3	9.7
Mean Ch. Stores	0.3	0.3	0.3	0.1	0.1	0.1
KP-F Statistic	99.8	98.8	96.0	55.0	54.8	50.7

Note: The table displays the effect of the entry of a convenience chain store (columns 1-3) or a supermarket (columns 4-6) in the census tract on the specialization of the traditional channel, estimated based on equation 4 using 2SLS. The outcome is the share of specialized stores over the traditional sector. Variations across columns denote different classifications of specialized stores: 3-digit classification 461 (establishments within retail sale of groceries, food, beverages, ice, and tobacco), 4-digit classification 4611 (establishments within retail sale of groceries and food), and 4-digit classification 4612 (establishments within beverages, ice, and tobacco). Standard errors in parentheses are clustered at the municipality level. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.8
Robustness: Specialization in Consumption

Dependant Variable: Share of Consumption and Trips in/to Specialty Stores

	Consumption ($)				Trips			
	1.5 km		2 km		1.5 km		2 km	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Conv. Chain Stores	0.25***	0.25***	0.15***	0.16***	0.25***	0.25***	0.14**	0.15**
	(0.07)	(0.07)	(0.05)	(0.05)	(0.09)	(0.09)	(0.07)	(0.06)
Observations	1,667,381	1,667,381	2,439,339	2,439,339	1,667,381	1,667,381	2,439,339	2,439,339
Year x Mun. FE	Y	Y	Y	Y	Y	Y	Y	Y
Market FE	Y	Y	Y	Y	Y	Y	Y	Y
HH Controls		Y		Y		Y		Y
Mean Dep. Var	44.6	44.6	44.7	44.7	48.1	48.1	48.3	48.3
Mean Ch. Stores	15.4	15.4	23.9	23.9	15.4	15.4	23.9	23.9
KP-F Statistic	98.3	98.2	108.7	108.7	98.3	98.2	108.7	108.7

Note: The table displays the effect of the entry of a convenience chain store in a given market estimated through 2SLS based on equation 4 at the household level. The monetary share of consumption is the expenses in specialized stores divided by the expenses in specialized stores and neighborhood shops. The share of trips is the number of days the household visited a specialized store within the week divided by the number of days it visited a specialized store plus the number of days it visited a neighborhood shop. Variations across columns denote different definitions of markets (census tracts within 1.5 km, and 2 km) and the inclusion of household-level controls: income per capita, income, household size, number of adults, age of household head, and monetary expenditures. Standard errors in parentheses are clustered at the municipality level. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.9
Robustness: Within-Firm Assortment Changes (Alternative Trade Area Definitions: 2 km)

	Assortment Changes: Competition/Customers			Assortment Changes: Placebo Reasons		
Dependant Variable:	Removed	Added	Changed	Removed	Added	Changed
	(1)	(2)	(3)	(4)	(5)	(6)
Chain Store Entered	0.11***	0.11**	0.15***	0.02	-0.04	-0.01
	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)
Observations	554	554	554	554	554	554
Store Controls	Y	Y	Y	Y	Y	Y
Mean Dep. Var	0.17	0.33	0.42	0.22	0.26	0.41

Note: The table displays the assortment changes within stores resulting from the entry of large chains within a 2 km radius, estimating equation 5 using a linear probability model with observations at the store level. Variation across columns denote diverse outcomes, respectively: assortment addition due to competition or customers, removal due to competition or customers, change due to competition or customers, addition due to other (placebo) reasons, removal due to other (placebo) reasons, and change due to other (placebo) reasons. The indicated regressions include store characteristic controls for precision (monthly profits, total employees, number of weekly customers, tax registration status, and an index of business practices). Robust standard errors are in parentheses. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.10
Robustness: Within-Firm Assortment Changes (Alternative Trade Area Definition: 0.5 km)

	Assortment Changes: Competition/Customers			Assortment Changes: Placebo Reasons		
Dependant Variable:	Removed	Added	Changed	Removed	Added	Changed
	(1)	(2)	(3)	(4)	(5)	(6)
Chain Store Entered	0.16***	0.12*	0.18***	0.004	-0.05	-0.03
	(0.06)	(0.07)	(0.07)	(0.06)	(0.06)	(0.07)
Observations	225	225	225	225	225	225
Store Controls	Y	Y	Y	Y	Y	Y
Mean Dep. Var	0.17	0.33	0.42	0.22	0.26	0.41

Note: The table displays the assortment changes within stores resulting from the entry of large chains within a 0.5 km radius, estimating equation 5 using a linear probability model with observations at the store level. Our control group consists of stores who did not experience the entry of large chains within 1 km, for consistency with our main specification. Variation across columns denote diverse outcomes, respectively: assortment addition due to competition or customers, removal due to competition or customers, change due to competition or customers, addition due to other (placebo) reasons, removal due to other (placebo) reasons, and change due to other (placebo) reasons. The indicated regressions include store characteristic controls for precision (monthly profits, total employees, number of weekly customers, tax registration status, and an index of business practices). Robust standard errors are in parentheses. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.11

Robustness: Within-Firm Assortment Changes by Capacity Constraints (Alternative Trade Area Definitions: 2 km)

	Assortment Changes: Competition/Customers					
	More Capacity-Constrained (small k)			Less Capacity-Constrained (large k)		
Dependant Variable:	Removed	Added	Changed	Removed	Added	Changed
	(1)	(2)	(3)	(4)	(5)	(6)
Chain Store Entered	0.15***	0.20***	0.22***	-0.03	-0.001	-0.03
	(0.04)	(0.05)	(0.05)	(0.07)	(0.07)	(0.07)
Observations	345	345	345	209	209	209
Store Controls	Y	Y	Y	Y	Y	Y
Mean Dep. Var	0.15	0.38	0.46	0.19	0.2	0.32

Note: This table displays the assortment changes within stores resulting from the entry of large chains within a 2 km radius, estimating equation 5 using a linear probability model with observations at the store level. Variation across columns denote diverse outcomes and subgroups, respectively: assortment addition due to competition or customers for more capacity-constrained stores, removal due to competition or customers for more capacity-constrained stores, change due to competition or customers for more capacity-constrained stores, addition due to competition or customers for less capacity-constrained stores, removal due to competition or customers for less capacity-constrained stores, and change due to competition or customers for less capacity-constrained stores. Robust standard errors are in parentheses. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.12

Robustness: Within-Firm Assortment Changes by Capacity Constraints (Alternative Trade Area Definition: 0.5 km)

	Assortment Changes: Competition/Customers					
	More Capacity-Constrained (small k)			Less Capacity-Constrained (large k)		
Dependant Variable:	Removed	Added	Changed	Removed	Added	Changed
	(1)	(2)	(3)	(4)	(5)	(6)
Chain Store Entered	0.19**	0.21**	0.27**	0.04	0.01	0.07
	(0.07)	(0.09)	(0.08)	(0.11)	(0.11)	(0.13)
Observations	142	142	142	83	83	83
Store Controls	Y	Y	Y	Y	Y	Y
Mean Dep. Var	0.1	0.33	0.40	0.13	0.16	0.22

Note: The table displays the assortment changes within stores resulting from the entry of large chains within a 0.5 km radius, estimating equation 5 using a linear probability model with observations at the store level. Our control group consists of stores who did not experience the entry of large chains within 1 km, for consistency with our main specification. Variation across columns denote diverse outcomes and subgroups, respectively: assortment addition due to competition or customers for more capacity-constrained stores, removal due to competition or customers for more capacity-constrained stores, change due to competition or customers for more capacity-constrained stores, addition due to competition or customers for less capacity-constrained stores, removal due to competition or customers for less capacity-constrained stores, and change due to competition or customers for less capacity-constrained stores. Robust standard errors are in parentheses. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.13
Robustness: Within-Firm Specialization by Capacity Constraints (Alternative Trade Area Definition: 2 km)

Dependant Variable: Fine-tuned Specialization Index			
	Full Sample	More Capacity-Constrained (small k)	Less Capacity-Constrained (large k)
	(1)	(2)	(3)
Chain Store Entered	0.16*	0.23**	0.01
	(0.09)	(0.11)	(0.14)
Observations	553	343	210
Store Controls	Y	Y	Y
Mean Dep. Var	0.66	0.63	0.74

Note: The table displays the specialization within stores resulting from the entry of large chains within a 2 km radius, estimated using OLS regressions with observations at the store level. The dependent variable is the specialization index described in Section 6.1. Variation across columns denote different subgroups, respectively: the full sample, more capacity-constrained stores, and less capacity-constrained stores. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.

Table A.14
Robustness: Within-Firm Specialization by Capacity Constraints (Alternative Trade Area Definition: 0.5 km)

Dependant Variable: Fine-tuned Specialization Index			
	Full Sample	More Capacity-Constrained (small k)	Less Capacity-Constrained (large k)
	(1)	(2)	(3)
Chain Store Entered	0.25*	0.33*	0.10
	(0.13)	(0.17)	(0.24)
Observations	223	141	82
Store Controls	Y	Y	Y
Mean Dep. Var	0.6	0.57	0.69

Note: The table displays the specialization within stores resulting from the entry of large chains within a 0.5 km radius, estimated using OLS regressions with observations at the store level. Our control group consists of stores who did not experience the entry of large chains within 1 km, for consistency with our main specification. The dependent variable is the specialization index described in Section 6.1. Variation across columns denote different subgroups, respectively: the full sample, more capacity-constrained stores, and less capacity-constrained stores. The stars next to the estimate, *, **, ***, represent statistical significance at the .10, .05, and .01 level, respectively.