One size fits all? The value of standardized retail chains*

Ben Klopack

Stanford University

For newest version see:

http://stanford.edu/~bklopack/

November 8, 2018

Abstract

Multi-outlet firms, or chains, make up a large and growing part of the US retail sector, with significant variation across store category and geographic location. This paper quantifies the demand-side incentives to operate a network of stores as a chain: various forms of economies of scale allow chains to generate higher demand than independent firms, but at the same time chains are less flexible in customizing product selection or prices across locations. To quantitatively assess this tradeoff, I develop a simple model and estimate it using a large transaction-level dataset from a payment card company, focusing on restaurant purchases. I find that on average chains could earn 22% more transactions if they could customize their product optimally to local tastes, but they would lose 13% of their transactions if they were to operate as unconstrained independents. Policies that ban chain restaurants would result in a loss of consumer welfare equivalent to 1.5% of restaurant spending and would disproportionately impact lower income consumers. Considering the endogenous entry of independent restaurants dampens the effects of chain bans by 30%.

Keywords: retail chains, standardization, market structure, demand heterogeneity

JEL Classification: L11, L25, R32

*I would like to thank Liran Einav, Tim Bresnahan, Rebecca Diamond, and Matt Genztkow for their invaluable advice and support. I am also grateful to Pascaline Dupas, Pete Klenow, Brad Larsen, Guido Martirena, Melanie Morten, Alessandra Peter, Nicola Pierri, Isaac Sorkin, and the participants of the Stanford Industrial Organization workshop and Stanford Applied Lunch for their insightful comments. I gratefully acknowledge financial support from the Patricia Liu McKenna and Kenneth McKenna Graduate Fellowship through a grant to the Stanford Institute for Economic Policy Research. All errors are my own.

1 Introduction

Large, national retail chains that operate in many markets may have large demand advantages relative to smaller firms. Chains may have economies of scale both in building an experience that consumers value, through product development and branding, and in communicating the reputation of that brand through marketing and advertising. The chain structure also enables a firm to share its reputation across stores, allowing it to more effectively compete for mobile consumers. However, chains also face a potentially important diseconomy of scope. In order to realize these reputational advantages, chains often standardize their product offering to create a consistent experience across stores. If consumer tastes differ across markets, standardization represents a costly strategic constraint for chain firms. The size of these two effects, and thus the payoffs to forming a chain relative to operating each store independently, is an empirical question.

This paper quantifies the value of the chain demand advantage in the restaurant industry and weighs it against the cost of standardization across a chain's stores. First, I examine the strategic problem that a firm manager faces in deciding how to organize a network of restaurants and the implications of this trade-off for market structure in retail industries. I leverage a large and novel dataset that includes all restaurant spending from a large payment cards company. I merge this data with a sample from Yelp that contains detailed restaurant-level information for seven mid-size US cities. I use this data to fit a discrete choice model of consumer choice over restaurants. In the model, restaurants are differentiated in their horizontal type (cuisine), vertical type (price and quality level), physical location, and chain affiliation. Consumer preferences over those restaurant attributes are allowed to vary both across cities and across consumers of different income groups. I categorize restaurants into eight cuisine types, which I observe from the Yelp data. To measure a restaurant's vertical type, I use the average dollars per transaction, defined at the brand level.[1]

While I focus on restaurants in this analysis, the basic forces I outline are present in standardized retail chains in many categories. Restaurants are an attractive setting to study consumer preferences for chains for several reasons. First, product differentiation in

[1] I assume a restaurant's price level is a measure of its quality. I further discuss this measure and its use in the demand model in Section 4.

restaurants is straightforward to measure using my data—Yelp records a measure of cuisine type, and I measure a restaurant's price level using the average dollar amount of its credit card transactions. Second, independent restaurants are available in the choice sets of most consumers; firms with only one location accounted for about a third of total restaurant sales, significantly more than other large consumer categories in the credit card data.[2] This variation in chain affiliation is important, since I use consumer choice patterns between chains and independents to identify the relative preferences of consumers for chain firms. Finally, restaurants make up a significant amount of consumer spending (approximately 14% of all transactions in my data).

I use the demand model to quantify the value of chains from consumer choices between observably similar chains and independent restaurants. I exploit variation in the spatial distribution of cardholders that changes how far a given consumer must travel to visit a particular restaurant. My estimates imply that chains with more than 1000 locations are highly valued by consumers relative to independent restaurants with a similar cuisine type and price level. The average consumer gets an increase in utility from eating at a chain that is equivalent to moving a restaurant 1-1.5 miles closer. This value is highest for the lowest income groups; cardholders with an annual household income less than $50,000 value chains about 50% more than those with annual household income above $250,000.

The costs of standardization across a chain's outlets depend upon the degree of heterogeneity in preferences across markets. I find evidence of significant taste differences across cities along both vertical and horizontal dimensions. For example, the average Cleveland consumer preferred Latin American restaurants to Burgers by about 1.1 mile equivalents, while a typical Phoenix consumer preferred Burgers by 0.9 mile equivalents. There is analogous heterogeneity in preferences over restaurant vertical types. These across-city taste differences suggest that the standardization constraint may be costly for chains that operate in multiple cities.

Both the reputational benefits of chain affiliation and the costs of the standardization constraint are large and economically significant. I use the estimates I recover from the

[2] In contrast, independent firms made up about 12% of sales in supermarkets and only 2% of sales in general merchandise.

demand model to quantify these two opposing effects. To quantify the cost that chains face, I ask how a given chain's demand would change if it could optimally choose its cuisine type and price level separately in each city, but keep the demand advantage afforded by its chain affiliation. I find that, on average, large chains in my data could get 12% more transactions if they could choose only their vertical type and 22% more transactions if they could choose both their vertical and horizontal characteristics.

I then consider the chain's hypothetical demand if it were to operate each outlet as an independent restaurant that chooses its horizontal and vertical type without constraint but loses its reputation advantage among consumers. I find that the average chain would lose about 21% of its transactions if it can only adjust its vertical type and 13% if it can adjust both its vertical and horizontal characteristics. The interpretation of some of the counterfactual exercises depends on why consumers prefer chains. If the primary benefits of chains for consumers act through investments made by the chain (for example, in advertising or product development), then a chain that decides to operate its restaurants independently would likely forfeit these advantages. If instead a chain's demand advantage comes from the unobserved quality of its particular restaurant concept or the skill of its manager, some of this advantage may be retained even if it were to operate each restaurant as an independent. I interpret the demand advantage primarily as the former and discuss this interpretation further in Sections 4 and 6.

I use these same demand estimates to quantify the welfare impact of chain bans on consumers in different income groups. In recent years, a number of cities have enacted regulations that ban or restrict the entry of chain firms, including restaurants. Chain bans are largely motivated by a desire to protect local businesses and maintain local character.[3] However, some have criticized these policies as primarily displacing stores that serve lower income consumers.[4] To evaluate the impact of these policies on consumers, it is important

[3] Local debates about chain retail restrictions routinely reference the standardized nature of chain stores. A regulation on chains in Jersey City, New Jersey targets retailers and restaurants with "multiple locations within the region that exhibit standardized characteristics such as logos, menus, store decor." (Retail Dive 2015)

[4] Dee Dee Workman, vice president of policy for the San Francisco Chamber of Commerce argues of the San Francisco ban on chain stores: " 'There used to be a lot more of neighbourhood-serving retail there ... retail that people could really use,' she says. The ban on chain stores changed that. 'What you have now are a lot of non-neighbourhood-serving retail: very, very high-end expensive little boutiques, selling super

to understand how consumers value chains, the distribution of those values across different groups, and the supply responses of independent firms.

When large national chains are replaced by independents with the same cuisine type and price and quality level, consumer surplus drops by between 1 and 2 miles per consumer, disproportionately impacting lower income consumers. This represents a loss in surplus equivalent to roughly 1.6% of restaurant spending. When national chains are instead replaced by a randomly drawn independent, consumer surplus drops by about 2 miles for consumers in the lowest income group, but actually benefits consumers in the two highest income groups—those with incomes between $200,000-250,000 and above $250,000. These consumers have weaker preferences for chains and tend to prefer both the vertical and horizontal characteristics of independent restaurants.[5]

In my primary demand specification, I allow restaurants to be differentiated along several dimensions, incorporating much of the richness from the data. However, I do not explicitly model the supply responses of other firms under the counterfactual scenarios, instead holding their entry and product characteristic decisions fixed at the values in the data. A change in chain strategy or a ban on chain restaurants might lead to supply changes by independent restaurants that partially offset the loss of chain brand value to consumers, an effect that is not accounted for by the demand model. To address this, I write down a stylized model of firm entry with one chain firm and a fringe of independent competitors. For tractability, I abstract away from spatial and horizontal differentiation, and allow restaurants to be differentiated only by their chain affiliation and their price level. In the supply model, one chain and a large fringe of independent entrants play a sequential entry game. The chain is constrained to choose the same location in every market, while independents can enter and locate flexibly. I follow much of the empirical entry literature and calibrate the model using a set of isolated small town markets.[6]

I then revisit the counterfactuals through the lens of the entry model using this small town sample. Relative to predictions made using only the demand model, counterfactuals

expensive shoes and purses and things.'" (Guardian 2017)

[5] Several of the arguments city planners make to justify chain bans lie outside of the scope of my paper, including concerns about local labor markets, city aesthetics, and positive externalities from local businesses. I discuss the magnitude of my estimates in the context of these arguments in later sections.

[6] See e.g. Bresnahan and Reiss (1990), Bresnahan and Reiss (1991), Mazzeo (2002), Seim (2006).

that incorporate the supply response find larger benefits to chain affiliation. If the chain were to give up its demand advantage and operate as an independent, my demand estimates predict that its transaction volume would drop by 28% in this sample, while calibrations of the entry model suggest that the effect is approximately twice as large. The reason for this is that the demand-based calculations hold fixed the behavior of other firms in response to a change in chain policy while the entry model predicts that a chain that changes to operating as an independent attracts entry in nearby product space, thus intensifying competition. Similarly, accounting for the supply response increases the estimate of the cost of the standardization constraint; when the chain moves to a new location in product space, some nearby independents change their product characteristics, reducing the level of competition that the chain faces. The demand counterfactuals predict that the chain could increase its transaction volume by 12%, while the entry model predicts an increase of 27%.

The model also suggests that the costs of chain bans are not as severe as those predicted by the demand-based estimates. In the demand model, I assume that each large chain in the data is replaced by an identical independent restaurant. The entry model suggests that a chain would be replaced by more than one independent, since the chain's demand advantage makes it a strong competitor that deters entry in equilibrium. Removing the chain thus encourages additional entry, partially offsetting the effects of a chain ban. Incorporating the endogenous supply response in this context reduces the cost to consumers by approximately 30%. Still, my estimates imply that these policies reduce consumer surplus by an amount equivalent to 1.2% of restaurant spending.

My estimates of consumer demand imply that chain restaurants are valuable to consumers, but do not distinguish between potential channels through which this demand advantage operates. In Section 6, I investigate several possible mechanisms. I first show evidence using new restaurant entries. I find that new restaurants that belong to large chains reach their steady state level of customers much more quickly after opening relative to independents. On average, outlets that are part of chains with more than 1000 locations grow about 10% annually in their first three years, compared to 26% annual growth for independent restaurants and 16% annual growth for small chains. These dynamics suggest that part of the value of chains for consumers is informational—consumers already recognize the chain's brand, and thus face less uncertainty about their utility from eating at the new restaurant. Second, I show evidence from consumers that move from one sample city to another. I find that the chain demand advantage is large and persistent after the move, but only for chains that had at least one location in the consumer's origin city—unfamiliar chains lost about 75% of the demand advantage. These results suggest that consumer familiarity and information—developed through chain branding and advertising—accounts for an important part of the chain advantage.

My results suggest a distinct, but complementary, explanation for the predominance of chains than that proposed by prior literature. Previous work has primarily focused on the cost advantages of chains that result from dense store networks and investment in cost-reducing technologies.[7] This paper provides evidence that these cost-side factors are only part of the story; chains also have important demand-side advantages over smaller firms.

Finally, my findings provide a lens through which to view chain ban policies. Chain bans are in general welfare-decreasing for consumers, and are particularly costly for the poor, who have the strongest preference for chains. This basic pattern is consistent with some who have criticized these policies as primarily benefiting the wealthy. It also helps to explain why cities that enact chain bans tend to have higher median incomes. These policies are frequently motivated by a desire to protect independent business owners. However, the magnitudes of my estimates suggest that bans of chain stores are a relatively costly instrument to transfer surplus to these firms. The estimates from the full entry model suggest that the amount of additional surplus that small firms receive only just offsets the efficiency loss in the form of reduced consumer surplus.

Several papers have examined the incentives to operate as part of a chain, in both the theoretical and empirical literature. My work is most closely related to Hollenbeck (2017), who studies the advantages of chains in the hotel industry. He finds that chain-affiliated hotels earn 20% higher revenues than similar independents, but that chains do not appear to have meaningful cost advantages. My paper also finds that chains earn a revenue premium

[7]Doms et al. (2004)) and Holmes (2001) find that large firms are more likely to invest in cost-reducing information technology like electronic scanners and bar code systems. Foster et al. (2006) find that large national firms have significantly higher labor productivity than single unit establishments. Jia (2008), Holmes (2011), and Ellickson et al. (2013) focus on estimating the cost-side benefits from economies in scale in distribution.

over independents, but focuses more explicitly on quantifying the potential costs to chains that arise from uniformity across outlets.[8]

My paper is also related to two papers that study uniform pricing policies of chains. DellaVigna and Gentzkow (2017) show that retail chains in the grocery, drug store, and mass merchandise categories set uniform prices across stores, despite significant heterogeneity in consumer demographics and competition across markets. Adams and Williams (2017) examine the effects of uniform pricing within geographic zones on competition and consumer welfare. I also quantify the cost of uniformity for retail chains that operate in heterogeneous markets, but I consider the effect of standardizing both price and other product characteristics like branding and cuisine type. I weigh these costs against the benefits the chain derives from brand recognition and familiarity among consumers.

There is a long literature in empirical industrial organization that studies market structure in retail industries. Ellickson (2007) applies the endogenous sunk cost models developed in Shaked and Sutton (1987) and Sutton (1991) to the market for grocery stores. He finds that this market fails to fragment even as the number of consumers grows large, consistent with a purely vertical model of competition. I expand on this literature by explicitly modeling how standardized retail chains that operate in many markets affect market structure and concentration. Shaked and Sutton (1987) makes the insight that markets in which consumers do not agree on the rankings of products will fragment as market size increases. The contribution of this paper is to highlight that when standardized chains are present that operate in multiple markets, not only within-market heterogeneity but also across-market heterogeneity in consumer taste affects the level of equilibrium concentration.

Finally, this paper builds on several recent papers that study the geography of consumption. Davis et al. (2017) use data on Yelp reviews to analyze the role that spatial and social frictions play in consumer choice of restaurants in New York City. Eizenberg et al. (2016) use aggregated credit card data to study the role of spatial frictions in explaining differences in grocery prices across neighborhoods in Tel Aviv. I bring a new rich dataset to look at

[8]In the theoretical literature, Loertscher and Schneider (2011) and Cai and Obara (2009) examine the reputational benefits of chains in attracting consumers. Additional relevant empirical papers include Mazzeo (2004), who studies the decision of motels to affiliate with a chain. In addition, there is a long literature that quantifies the cost advantages of chains, including Doms et al. (2004), Holmes (2001), Foster et al. (2006), Jia (2008), Holmes (2011), and Ellickson et al. (2013).

these questions. I use a similar empirical strategy to Davis et al. (2017), but focus on how consumers value chains and independent restaurants, and the effect on consumer welfare of chain ban policies.

The rest of the paper is organized as follows. In Section 2, I present details of the credit card and Yelp datasets. Section 3 presents statistics about the aggregate importance of chains and facts about market structure in the restaurant industry. Section 4 introduces a model of consumer demand for restaurants and shows results from estimation and accompanying counterfactuals. In Section 5, I present a model of restaurant entry and calibrate it using the credit card data. In Section 6, I investigate possible channels for the chain demand advantage. Section 7 concludes.

2 Data

2.1 Sources

My analysis of retail chains leverages a novel dataset provided by a large payment cards company that includes the universe of credit and debit transactions on the network in 2016. Each observation in the underlying data is a transaction between a cardholder and a merchant. On the merchant side, I see the merchant and store, which are mapped to a business category and location. On the card side, I see a unique card identifier. The data contain no information on the specific goods or services that were purchased, nor the prices of those items. The sample is completely anonymized, and I do not observe the name, address, or any other personally identifiable information about the cardholder. For about 50% of the active cards in 2016, the company has access to a measure of estimated household income, which I use in some of the analysis.

I further supplement this data with a sample from Yelp. Yelp is a major consumer review platform that allows users to review businesses and collects information about business category, hours, locations, and other characteristics. Yelp makes a sample publicly available for academic use that contains reviews and business attributes from a 2017 snapshot of all businesses in seven US cities: Las Vegas, NV; Phoenix, AZ; Charlotte, NC; Pittsburgh, PA; Cleveland, OH; Madison, WI; and Urbana-Champaign, IL. The data provide additional detailed merchant information, including a more granular business category, price level, and average user rating. I merge this with the credit card data using the merchant name, zipcode and address for these seven cities for the restaurant category. I report additional details of the merge in ??.

2.2 Variable construction

I lean on several data constructs throughout the paper to characterize merchants and cards. First, I define a consumer's location as the latitude and longitude coordinates corresponding to the weighted average of the zipcode centroids in which it transacts at least 20 times. In later sections of the paper, I use the distance between this consumer location and a restaurant's location as an explanatory variable.

To characterize each merchant, I construct three measures: the number of store locations within its brand, the average dollar amount of its transactions, and the merchant's restaurant category.[9] To construct the number of store locations l_j for each merchant j, I count the number of distinct store IDs belonging to a given merchant ID. I also make use of a merchant's average number of dollars per transaction—or its "ticket size"—as a measure of a merchant's price. I define ticket size at the merchant level. Finally, I assign each restaurant merchant to a category using the Yelp data. Yelp associates each restaurant with one or more tags that describe the type of food that it sells. I manually map this set of tags into eight categories and assign each restaurant to one of these.

2.3 Sample construction

In my final analysis sample, I use transactions made by cards active in 2016 for which I observe estimated household income. In Section 3, I show some basic facts about chains using all transactions that occurred in a set of augmented retail categories—in addition to traditional retail, I add restaurant and hotel transactions.[10] The largest categories within

[9] I use merchant to describe a specific retail brand. A merchant can be a chain that has many outlets under the same brand name, or an independent merchant with only one outlet.

[10] I include transactions at merchants classified by the payment card company as belonging to 2-digit North American Industry Classification System (NAICS) codes 44 and 45 (retail) and 72 (restaurants and hotels).

this group by total spending were restaurants (25%), grocery stores (16%), and general merchandise stores (13%).

In later sections, I limit the analysis to restaurants. In Section 4, I construct a sample of restaurant transactions made by urban consumers. For this sample, I use the set of transactions in 2016 that occurred at Yelp-matched restaurants in the seven US cities included in the Yelp sample. I further filter the sample by eliminating cards with a shopping location further than 25 miles away from the city and keeping only evening transactions (those made between 5pm and 11pm local time). Finally, I eliminate cards that transacted less than five times within the city. I limit the sample to cards that transacted relatively frequently to control for card-level unobserved heterogeneity in restaurant preferences. An observation in this dataset is a single transaction that occurred between a card and a restaurant. I show summary statistics for this sample by city in Table 1. In total, the sample included about 1.25 million transactions and $37 million dollars, spread over nearly 3000 restaurants across the seven cities.

In Table 2, I show card-level activity measures by bins of the estimated household income measure. Consumers in each income bin made about the same number of restaurant transactions during the sample period, but higher income consumers tended to spend more per transaction. They also tended to visit restaurants with higher average transaction sizes—the average transaction by a consumer with household income below $50k was at a restaurant with ticket size $28, compared to $35 for a consumer with income above $250k. The average transaction was made about 5 miles away from the consumer's location. In Table 3, I show the number of restaurants across the eight different cuisine type categories. The largest category is American, with 537 restaurants, while the smallest is European, with 115 restaurants.

In section 5, I analyze restaurant transactions in a set of small town markets. Starting from the set of places listed by the Census, I keep the set that meet the following criteria: each place must have between 1,000 and 10,000 residents, be at least ten miles from the nearest city or town of any size, and be at least 25 miles from the nearest city of 50,000 people or more. This leaves me with a sample of 640 towns. Of these, 418 towns had at least four restaurants in 2016. For these towns, I consider all evening restaurant transactions by consumers with a shopping location within 25 miles of the town and that made at least 5 restaurant transactions in the town in 2016. I report summary statistics by town in Table 4. The average town has 371 accounts making about 11,000 transactions at 16 restaurants.

3 Chains vs. independent stores: empirical patterns

3.1 Importance and growth of chains

To motivate my focus on retail chains, I first document the importance of chains across merchant categories. As I detail in the previous section, I use all 2016 credit card transactions for which I observe estimated household income in the set of augmented retail categories defined above. In Figure 1a, I show the share of spending by firm size. Less than 20% of spending in these categories is made up by single-establishment firms, with the rest spent at chains. In particular, nearly 40% of spending occurred at large, national chains with more than 1000 locations.

In Figure 1b, I show the breakdown of spending by firm size for four large merchant categories in the payment card data. Among the four, restaurants is the most fragmented, with about a third of spending at single-establishment firms, while in the general merchandise sector, chains accounted for over 98% of all spending.

Other work in the literature has documented the growth of chains during the 20th century. Analysis of longitudinal Census data shows that chains accounted for less than 30% of retail sales in 1948. That figure rose to 40% by 1976, and over 60% by 1997, with most of the growth at firms with more than 100 locations (Jarmin et al., 2009).

3.2 Characteristics of chain restaurants

I illustrate several stylized facts about chain and independent restaurants to motivate the demand and entry models I present later in the paper. First, chain restaurants are on average cheaper than independent restaurants. I illustrate this fact by relying on the average dollar amount of a transaction. In Figure 2, I show the mean and quantiles of average merchant ticket size separately for independents and restaurant chains of different sizes, with each merchant weighted by its number of transactions in 2016. The quantiles of the ticket size distribution are monotonically decreasing in the number of locations. The median dollar spent at independents went to a merchant with a ticket size of \$33, while the median dollar spent at chains with more than 1000 locations went to a merchant with a ticket size of about \$11. The table also shows that the distribution of average ticket sizes for large chains was quite compressed relative to independents and small chains. The difference between the 10th and 90th percentile of average ticket size for large chains was just \$13, compared to to over \$50 for independents.

Independent restaurants and small chains also tend to have a larger share among higher income consumers and in higher income places. In Figures 3a and 3b, I plot the share of transactions that went to restaurant chains of different sizes by quartile of county income and cardholder annual income. In the poorest quartile of counties, chains with more than 100 locations accounted for 44% of all restaurant sales, while in the richest quartile of counties those chains made up only 31%. Similarly, consumers in the highest income group, those with annual income above \$250k, conducted 37% of their restaurant transactions at chains with more than 100 locations, relative to 46% of transactions for the lowest income group.

Finally, independent restaurants tend to get a larger share of their business from repeat customers. For each restaurant, I calculate the share of transactions in a six month period that came from cards that only transacted once at the restaurant. To account for the fact that cheaper restaurants, which tend to be chains, are visited more frequently, I regress the share of transactions from one-time customers on a set of zipcode, ticket size quartile, and 6-digit NAICS fixed effects, as well as a set of fixed effects corresponding to firm size:

$$sh_{iz}^{1trans} = \alpha_m + \alpha_{ts} + \alpha_z + \beta_l$$

In Figure 4, I plot the firm size fixed effects β_l. The figure shows that chains with more than 1000 locations got about six percentage points more of their transactions from one-time customers relative to independent restaurants.

Together, these facts indicate that while chain restaurants are generally popular, there is a large fringe of independents that continue to operate. Chains also tend to target different types of consumers than independents. Chains are most popular among lower-income consumers who are likely to be more price sensitive. This heterogeneity suggests that the effects of chains on consumer welfare may not be equally distributed across demographic groups. It also suggests that some chains may have room to increase revenues in some markets if they could customize—for example, by offering a more expensive product in markets in which higher income consumers are more prevalent.

4 Costs and benefits of chains

Now I turn to quantifying the demand benefits from chain affiliation and weigh those against costs imposed by selling a uniform set of products in heterogeneous markets. I assume that chains optimize over both their vertical and horizontal product characteristics. In this section, I use the urban consumer sample described in Section 2 to recover consumer preferences for these two characteristics separately in each of the seven cities. Differences in consumer tastes across cities measure the degree of heterogeneity in demand. I then use the recovered parameters to conduct counterfactual analysis: I estimate the degree to which a chain could increase its revenues by choosing its vertical and horizontal characteristics flexibly in each city, and assess the impacts and distributional consequences of a ban on large chain restaurants.

4.1 Descriptive evidence

To illustrate some of the data patterns that drive my demand results, I first show some model-free evidence of several of the most important findings. First, the distribution of consumer preferences is heterogeneous across markets. The key variation that identifies the cost of standardization is the extent to which consumers in different cities prefer different types of restaurants. In the extreme case of perfect homogeneity across markets, the optimal choice of product characteristics—for restaurants, the level of price and quality and the cuisine type—is the same in each market, and thus the standardization constraint is not binding. In my sample, I find suggestive evidence of important heterogeneity across cities in both price and quality and cuisine type.

In Figure 5a, I show the per-restaurant average number of transactions (benchmarked to the average for that city) for restaurants of different ticket sizes by city. I measure price and quality using average ticket size, as in the figures above. I discuss this assumption further in Section 4.2. The figure shows that in Las Vegas, restaurants with ticket size below $25 received about 1.02 times the average number of transactions, while restaurants with a ticket size between $26 and $50 received about 0.84 times the average number of transactions. In Cleveland though, that pattern was reversed; restaurants in the $26-$50 bin received more than twice the number of transactions as those in the lowest bin.

The data show a similar degree of heterogeneity across restaurants of different cuisine types. In Figure 5b, I show the average number of transactions for restaurants of different cuisine types across cities. In Cleveland, Latin-American restaurants were extremely successful, getting about twice the average number of transactions, while in Las Vegas they were merely average. Instead, Burger restaurants were among the most popular in Las Vegas. This relatively large degree of variability in the success of restaurants of different types suggests that a chain that could be fully flexible in its choice of product characteristics across markets might choose to sell different types of food at different prices in each city.

Second, high income consumers tend to visit chains less than low income consumers, even conditional on ticket size. In Section 3, I show that consumers with annual household income above $250k spent about 9 percentage points more of their transactions on chains with at least 100 locations than did consumers with incomes below $50k. However, this pattern could be driven by preference among low income consumers for cheaper restaurants, which tend to be large chains. In Figure 6, I show that this relationship also holds conditioning on ticket size. Among transactions at restaurants that had a ticket size less than $25, consumers that made more than $250k visited chains with more than 100 locations about three percentage points less than consumers that made less than $50k, and about five percentage points less than those with incomes of $50-100k.

Finally, consumers in my sample are quite sensitive to distance. In Figure 7, I show the probability that a restaurant is chosen as a function of the distance between a consumer's shopping location (as defined above) and the restaurant, aggregated across consumers and cities. The figure shows that that a given restaurant is about 40% less likely to be chosen at two miles away versus one mile away, and an additional 40% less at three miles versus two miles. In total, over 70% of restaurant transactions in my sample were at restaurants that were less than five miles away from the consumer.

4.2 Econometric specification

In my main demand specification, I allow restaurants to be differentiated horizontally, vertically, and spatially. I define a restaurant's horizontal characteristic as its cuisine type, which I allow to take one of eight categories. I allow for preferences over horizontal types to vary by consumer income group and city. A restaurant's vertical type corresponds to its price and quality level, which I estimate separately for each consumer. I also allow a consumer's choice to depend on her distance in physical space from each restaurant. I assume that the set of restaurant trips is given for each consumer, and the consumer decides which restaurant to visit among the set of restaurants available in the city.

Restaurants differ in both price and quality. If producing a higher quality restaurant experience is costly, the observed quality and price offered by restaurants in equilibrium will be correlated.[11] More expensive restaurants are likely to use pricier ingredients and offer more service than casual restaurants.[12]

I proxy for a restaurant's price and quality level with a restaurant's average transaction size x_j. I do not observe price and quality separately in my data. I instead assume that restaurants choose from a schedule of price and quality bundles. Consumers have single-peaked preferences over these bundles; each card prefers a certain level of quality given the equilibrium relationship between price and quality. This is consistent with a vertical model in which some consumers value quality more highly than others, and prices and qualities are monotonically ordered in equilibrium. Cremer and Thisse (1991) show that every vertical model (in which all consumers agree on the ranking of products) within a broad class can be equivalently formulated as a horizontal Hotelling-like model (in which consumers differ in their rankings). This formulation helps to rationalize the empirical observation that some consumers in my data repeatedly choose higher priced restaurants when low-priced options

[11] See, for example, Mussa and Rosen (1978).

[12] Toast, a provider of inventory management software for restaurants, reports that average food cost was about 35% for fine dining restaurants and 25% for quick service restaurants, implying that more expensive restaurants spend significantly more on raw ingredients than cheaper competitors.

are available. I also assume that the underlying relationship between price and quality is unchanged in the counterfactuals.

The link between price and quality may be different for chains and independent restaurants. If chains have a demand advantage, then they may charge higher markups on average than independents. On the other hand, chains may have lower marginal costs to produce a given level of quality because of economies of scale in procurement or distribution. If the relationship between price and quality differs systematically across firm size, these differences will show up as part of the difference in consumer utility from chains. In the restaurant industry, cost differences between chains and independents appear to be of secondary importance. While high quality data on the marginal costs of independent restaurants is difficult to obtain, industry sources suggest that marginal costs are comparable to those reported by publicly held restaurant chains.[13]

Average transaction size is the product of price and quantity, and thus could be higher simply because consumers at some restaurants tend to purchase more items or visit in larger parties. To test this, I compare average transaction size to an alternative measure of price from the Yelp data. Yelp assigns each restaurant a price rating between one and four dollar signs. In the Appendix, I plot the distribution of log ticket size by Yelp dollar sign rating. The four distributions are monotonically ordered by ticket size and largely non-overlapping, consistent with the notion that most variation in average transaction size results from differences in prices. In modeling consumer choice and restaurant entry in later sections, I prefer the ticket size measure for the additional granularity it contains over the coarser Yelp measure.

I assume that consumer i's utility from visiting restaurant j belonging to a merchant with $l(j)$ locations in trip t is:

$$u_{ijt} = \delta_{l(j),y(i)} + \theta_{y(i)} z_j - \tau(\alpha_i - x_j)^2 - \gamma dist_{ij} + \epsilon_{ijt} \tag{1}$$

[13] I collect cost data (where available from public company financial reports) from the top 10 fast food chains and top 5 fine dining and casual dining chains, as ranked by Restaurant Business Magazine. For these large firms, food costs as a share of revenue from company-operated restaurants are between 28% and 31% of revenue, while labor was between 28% and 32%. Resources for independent restaurant operators suggest a similar range, with the sum of labor and food costs lying between 55% and 65% of revenue (see e.g. https://therestaurantexpert.com/figure-your-restaurants-prime-cost/).

where $\delta_{l(j),y(i)}$ is the taste of card i's income group $y(i)$ for chain size $l(j)$, z_j is restaurant j's horizontal category, x_j is the restaurant's average ticket size, and $dist_{ij}$ is its distance from consumer i's shopping location (the weighted centroid of the zipcodes in which it transacts). ϵ_{ijt} is a random preference shock that is distributed iid extreme value type I. Cards are divided into six bins $y(i)$ based on annual household income: \$0-50k, \$51-100k, \$101-150k, \$151-200k, \$201-250k, and >\$250k. Restaurants are divided into four bins $l(j)$ based on the number of US locations belonging to that brand: 1 location, 2-100 locations, 101-1000 locations, and more than 1000 locations. $\theta_{y(i)}$ is a vector of preferences for consumers of income group $y(i)$ over restaurant categories. α_i is card i's "bliss point" ticket size that gives her preferred level of quality. I assume α_i is fixed for each consumer and place no restrictions on its distribution in the population.[14] τ is a parameter that scales the cost of visiting a restaurant far from an individual's ticket size preference and γ is a parameter that describes the utility cost of traveling one mile. All parameters are allowed to vary freely across cities.

The purpose of the demand model is to make counterfactual predictions about consumer utility and restaurant demand if a restaurant were to change its chain affiliation or product characteristics. The specification in equation 1 makes several assumptions about the form of consumer utility that may affect these counterfactual predictions. First, it assumes that the difference in utility that consumers receive from eating at a chain restaurant relative to an independent enters additively in a consumer's utility, and does not interact with other restaurant characteristics like price, cuisine type, or location. This assumption would be violated if consumers value lower ticket size chains differently than higher ticket size chains, for example. I test this assumption by looking at the differences in the average number of transactions for chains of different price ranges and cuisine types. In general, I find that the difference in demand between chains and non-chains does not vary systematically by ticket size. I show these results in the Appendix.

[14]This allows for α_i to be correlated with other explanatory variables—for example, a consumer's willingness to pay for restaurant quality may be correlated with household income or the location in which the card transacts.

4.3 Estimation and results

My demand estimation uncovers significant heterogeneity in preferences across the seven sample cities. I estimate the parameters of demand in equation 1 for each of the seven cities independently via maximum likelihood. Equation 1 allows for card-specific preferences over restaurant ticket sizes α_i. These preferences are identified by cards that make many restaurant transactions during my sample and help to account for unobserved heterogeneity within demographic groups. While the model can be estimated using standard maximum likelihood methods, the inclusion of individual-specific preferences requires nonlinear estimation of thousands of fixed effects and presents computational challenges. I proceed by constraining the values of α_i to a discrete grid and estimating the model via the two-step estimation procedure in Heckman and MaCurdy (1980). I describe the estimation further in the Appendix.

I show the main parameter estimates in Table 5. The parameters γ and τ, the utility cost of traveling one mile and of eating at a restaurant with a ticket size far from a consumer's preference, are both negative, as expected. The average estimate for γ across the seven cities in my sample is -0.4, implying that moving a restaurant one mile further away from a consumer decreases the probability it is chosen by 33%.[15]

My estimation does not include a traditional price coefficient, since I assume that average restaurant transaction size measures price and quality together. Instead, I quantify the parameters in terms of "mile equivalents"—the amount of physical distance that a consumer would travel to visit a restaurant with a particular characteristic. To understand how consumers trade off physical distance and eating at a restaurant with a ticket size that is far from that consumer's preferences, I calculate the indifference curve between restaurant ticket size and physical distance for a hypothetical consumer with a ticket size blisspoint $\alpha_i = 40$ using the average of the parameters I estimate across cities.[16] I plot this relationship in Figure 8a. It shows that consumers are sensitive to ticket size—the average τ parameter

[15]The ratio of choice probabilities for two restaurants j and j' that are identical except that j is one mile further from consumer i than j' is $P_{ij}/P_{ij'} = \exp(\gamma(dist_{ij} - dist_{ij'}))$. When $\gamma = -0.4$, $P_{ij}/P_{ij'} = \exp(-0.4 \cdot 1) = 0.67$.

[16]To calculate the number of miles a consumer would travel to make her indifferent between eating a restaurant with ticket size x_j and a restaurant with ticket size \$40, I divide the ticket size penalty term $\tau(x_j - \alpha_i)^2$ by γ.

implies that consumers are indifferent between traveling about 0.5 miles of physical distance to eat at a restaurant with ticket size at their blisspoint or eating at an otherwise-identical restaurant with ticket size \$20 above or below their blisspoint.

I use a similar approach to quantify the firm size coefficients. In Figure 8b, I show the chain fixed effects $\delta_{y(i),l(j)}$ (with preferences for restaurants with only one location normalized to zero) averaged across markets and divided by γ. The chart shows that consumers strongly preferred chains to independents, and preferred large chains to medium and small chains. The parameter estimates imply that consumers are indifferent between eating at an independent restaurant and traveling an additional 1 to 1.5 miles to eat at a large chain. However, this "chain premium" was much lower for high income consumers than for low income consumers. Consumers with annual income below \$50,000 value chains about 50% more than consumers with annual income above \$250,000.

Importantly, my estimates of $\theta_{y(i)}$ suggest that there is significant taste heterogeneity in restaurant categories across markets. In Figure 9, I show the rankings of preferences different horizontal categories by city (averaged across income groups within a city). In Cleveland, Pittsburgh, and Charlotte, the most-demanded category was Latin-American, while in Las Vegas and Champaign, Burger restaurants were the most successful. Within a city, the difference in demand between the most popular and least popular categories was between two and four mile equivalents.[17]

4.4 Counterfactuals

I use these estimates to conduct counterfactuals to quantify the returns to operating as a chain and to measure the consumer welfare impact of a ban of chain restaurants. The costliness of the standardization constraint relative to the additional demand that a chain receives has important implications for the prevailing market structure. As markets become more heterogeneous, the costs to operating as a chain relative to an independent become larger, and thus concentration (at the national level) will be limited.

[17] To calculate this, I divide each component of the vector $\theta_{y(i)}$ (the income group-specific preferences over restaurant categories) by γ and average over income groups. The intra-city range in demand between categories is given by the largest element of this vector minus the smallest.

4.4.1 Chain customization

I first quantify the costs and benefits of being part of a chain. I consider the vertical and horizontal product choices of large chains with more than 1000 locations nationwide, holding the actions of all other firms fixed. First, I compute the baseline predicted transaction volume for each chain under the status quo—in which it chooses a single vertical and horizontal characteristic across all outlets and cities in the sample. Second, I compute the chains' predicted transaction volumes if they were to pick their product characteristics separately in each market, while keeping the chain demand advantage δ_{1001+}. Finally, I compute each chain's revenues if it were to dissolve the chain and operate each store independently, with each store free to choose a different price and horizontal category, but without the demand advantage δ_{1001+}. I conduct this exercise for the 22 large chains in the sample separately and report averages across the merchants.[18] I consider a scenario in which each chain can only adjust their ticket size and one in which each firm can choose both ticket size and horizontal category.

In the model, the additional utility that consumers receive from visiting a chain is identified from the differences in choice probabilities between observably similar chains and independents. However, the interpretation of the counterfactuals depends on why consumers receive higher utility from chains. If the primary benefits of chains for consumers act through investments made by the chain (for example, in advertising or product development), then a chain that decides to operate its restaurants independently would likely forfeit these advantages. If instead a chain's demand advantage comes from the unobserved quality of its particular restaurant concept or the skill of its manager, some of this advantage may be retained even if it were to operate each restaurant as an independent. Two pieces of evidence suggest that much of this demand advantage would be lost if the restaurant were to abandon its chain affiliation. First, the largest restaurant chains advertise extensively. In 2018, 8 of the 25 most-advertised brands on television were restaurant chains (IdenTV, 2018), with McDonalds spending approximately \$1.5B on US advertising in 2017, or about 25% of its

[18] My demand estimation uses only transactions during dinner hours. Since preferences over both vertical and horizontal dimensions may be different for breakfast or lunch, I consider only large chains for which dinner hours were the largest share of transactions.

US revenue (Ad Age, 2018).[19] These patterns suggest that the largest firms reap substantial benefits from advertising. Second, this interpretation is broadly consistent with the findings of the literature on branding.[20] This distinction is less important for quantifying the impact of chain bans, since the net effect is that the chain is induced to exit and the restaurant is no longer available to consumers in that market.

In Table 6, I show the average effect of the two counterfactual policies on the transaction volume of the chains in each city. I consider two scenarios: one in which each chain can change only its vertical characteristic—for example, to sell higher quality hamburgers at higher prices—and one in which it can change both its vertical and horizontal characteristics. The first column shows the percentage change in transactions (across all outlets) if each chain can flexibly choose its quality level in each market, but keep its demand advantage. Relative to the constrained optimum, the chain could increase its transaction volume by an average of 12% across markets. In some markets, the chain's ticket size at the constrained optimum is close to the unconstrained optimum—in Pittsburgh and Cleveland, for example, chains would not benefit at all from customization. In Champaign and Madison, on the other hand, my estimates of consumer taste distributions imply that most diners prefer dinner restaurants with a ticket size lower than the constrained optimum for most chains as calculated from my demand estimates.

The second column shows the change in the chain's transaction volume if it were to give up its demand advantage and operate each outlet as an independent restaurant. In a market like Pittsburgh, where the chain stands to gain nothing from customization, it would do substantially worse by operating as an independent, losing 42% of its transactions. On average, the chain would lose about 21% of transaction volume across all seven markets.

Compared to a chain that could fully optimize both price and product category, adopting flexible pricing achieves more than half of the total theoretical benefit from customization. In columns (3) and (4), I allow chains to change both their ticket size and horizontal

[19] These numbers are somewhat higher than leading consumer packaged goods companies, whose products have been studied extensively in the branding literature. For example, Procter & Gamble, the world's second largest advertiser, spent $4.4B in advertising on about $28B of revenue (15%) (Ad Age, 2018).

[20] In particular, see Hollenbeck (2017) and Tsai et al. (2015) who analyze the hotel industry. Hollenbeck (2017) finds that a hotel that switches from independent to chain gets a 21% increase in their revenue, while Tsai et al. (2015) find that hotels that rebrand from one chain to another increase their revenues about 4%.

category. Relative to their current category, a chain that could customize but keep its demand advantage could increase its transaction volume by 22% relative to the constrained optimum. If it were to lose its demand advantage δ and operate each location independently, it would lose 13% relative to the constrained optimum.

4.4.2 Chain bans

In the second set of counterfactuals, I consider the effect of a hypothetical policy to ban large chains with more than 1000 locations. Chain bans or entry restrictions have been implemented in a number of small and large cities across the US. I will consider a "hard ban" most similar to the policy enacted in the Hayes Valley neighborhood in San Francisco.[21] In this section of the paper, I only model the demand side of the market in each city. As a result, the counterfactual analysis ignores general equilibrium effects (for example, other firms' competitive responses to a chain ban) and makes no rigorous prediction about market structure under alternate policies. I proceed by considering three market configurations: in the first, chain restaurants disappear and are not replaced. In the second, chains are replaced by independent restaurants with identical characteristics (both cuisine type and ticket size). In the third, I replace each chain with a randomly drawn independent restaurant from the population of independents in the city. In Section 5, I present a stylized model of entry to consider some of these potential general equilibrium effects.

I show the results of the counterfactual analysis in Table 7. In general, chain bans have large, negative effects on consumer surplus. Completely removing all chain options without replacement by another choice cuts consumer welfare by between 5.4 and 7.2 mile-equivalents per consumer, with the largest impacts falling on low income consumers, who in general have strong preferences for chains. The total loss to consumers across the seven markets is equivalent to the disutility from travelling 564,719 miles. I convert this loss into dollar terms by multiplying the number of (one-way) miles between a consumer and a restaurant by two to get the roundtrip distance and assuming that each mile costs \$0.91 in time costs and \$0.81 in direct costs, for a total of \$3.44 for each one-way mile between the consumer

[21] The San Francisco policy is slightly more restrictive than my simulated policy, as it bans chains that have more than 11 locations from operating in Hayes Valley. The hard ban operates only in certain neighborhoods but other restrictions on chain retail are in place in most San Francisco neighborhoods.

and the store.[22] This conversion gives me an estimated loss in consumer surplus equivalent to about 5.2% of all restaurant spending in these seven markets.

However, large chains (which make up between 20 and 30% of all restaurants in these seven cities) are likely to be replaced through additional entry by independents and smaller chains in the event of a ban. First, I assume that each large chain outlet is replaced by an independent that sells the same type of food and has the same average ticket size. In this scenario, consumer surplus drops by between 1.1 and 2.2 miles, with a dampened effect on the highest income consumers, implying a loss equivalent to 1.5% of spending in aggregate.

I show in earlier sections that chains tend to have different characteristics than independents—in particular, they have lower average ticket size and tend to sell different types of food. Thus, it may be unrealistic to assume that each low-end chain burger restaurant would be replaced by an identical independent burger restaurant. To address this, I simulate a third scenario, where I assume that each large chain restaurant is replaced by a randomly drawn independent. I show that the effect on consumer surplus in this case is slightly positive for the highest income consumers—on average, consumers with annual household incomes above $250,000 would benefit by about half a mile, as they tend to prefer the more expensive restaurants that replace chains under this scenario.

Arguments for chain bans tend to focus on three themes: aesthetic concerns, positive local externalities from independent businesses, and preference for independent ownership. The first two justifications—preserving the character of downtown areas and positive labor market and other externalities from independents—are difficult to quantify and outside the scope of my model. My estimates suggest that chain bans are a costly tool for redistributing surplus to local business owners. Assuming that the overall level of restaurant spending remains fixed under the policy, a chain ban would redirect the 25% of restaurant spending going to large chains to independents and small chains. Industry sources estimate average accounting profit margins at about 5% of revenues, implying that these policies would increase independent

[22] To obtain the monetary cost of a mile, I follow Einav et al. (2017) in using estimates from Einav et al. (2016), who report summary statistics for a large number of short-distance trips of breast cancer patients. They report that an average trip takes 10.9 minutes to travel 5.3 straight-line miles, with an actual driving distance of 7.9 miles. The BLS reports that the average after-tax hourly wage in 2016 was $26 per hour. As an estimate for the driving cost, I use the IRS 2016 reimbursement rate of $0.54 per mile, which considers the cost of fuel and depreciation of the car. Thus, the time cost of driving one mile is given by $\$26/60 \cdot 10.9/5.3 = \$0.91$ and the driving cost of of one mile is $\$0.54 \cdot 7.9/5.3 = \$0.81$.

profits by roughly 1.25%, less than the implied loss of consumer welfare of 1.5% when chain restaurants are replaced one for one by similar independents.

5 Supply responses by independent firms

In the previous section, I present and estimate a model of consumer demand for restaurants. I use the estimates that I recover from this model to assess the effect of chain bans, and to estimate the costs and benefits to operating a set of establishments as a chain with a standardized price and product strategy. However, this analysis did not consider potential supply responses by other firms. If a large chain chooses different product characteristics, it changes the competitive environment faced by independent firms. New firms may enter in the parts of the product space vacated by the chain, while firms that are similar in product space to the chain's new choice of characteristics may be induced to exit or choose new characteristics.

To account for these effects, I write down a stylized model of restaurant entry with endogenous product characteristic choices. For tractability, I abstract away from spatial and horizontal differentiation between restaurants, as well as strategic interactions between chains, and focus on a single chain's choice of quality level. In the model, one chain and many potential independent entrants make sequential entry decisions and vertical characteristic choices in a set of small town markets.[23] The chain moves first and enters in all markets, but must choose the same quality level in each town. I then calibrate this entry model using data from evening restaurant purchases made in this set of small towns and revisit my counterfactual analysis from the previous section through the lens of the model.

5.1 Model of entry

I consider a modified version of the sequential location game developed in Loertscher and Muehlheusser (2011). Consumers in market m choose between a set of restaurants that are differentiated on quality. One firm in each market is a chain, while all others are

[23] I follow much of the entry literature by focusing on geographically isolated small-town markets (e.g. Bresnahan and Reiss (1990), Bresnahan and Reiss (1991), Mazzeo (2002)) and assume that each market is self-contained and unaffected by restaurant offerings outside the town.

independents. Consumer preferences α are distributed on the interval $[0, \infty)$ according to the distribution function $f_m(\alpha)$, which I parameterize with a lognormal distribution. A consumer's preference α describes her bliss point in quality space, which I measure with a restaurant's average ticket size. A consumer i with bliss point α_i receives the following utility from eating at chain restaurant j with ticket size x_j:

$$u_{ijt} = \delta - \tau(\alpha_i - x_j)^2 \tag{2}$$

As in the previous section, the parameter δ describes the difference in utility that a consumer gets from eating at a chain relative to an independent restaurant for a consumer of a given income group and τ determines the cost of visiting a restaurant far from a consumer's ticket size bliss point. When j is not a chain, her utility is instead:

$$u_{ijt} = -\tau(\alpha_i - x_j)^2 \tag{3}$$

One chain and many potential independent entrants observe consumer preferences in each market and play a sequential entry game. Each firm decides whether to enter and, upon entry, chooses a vertical characteristic. If the firm enters, it pays a sunk cost K_m which varies across markets. The chain moves first in each market and faces different demand relative to similarly positioned independent firms because of the additional utility $\delta_{y(i)}$ that consumers receive. However, the chain is constrained to pick the same ticket size in each market. I assume that the chain enters in every market. Each firm j earns profits equal to the mass of consumers it serves, Q_j^m, less the sunk cost of entry K_m:

$$\Pi_j^m = Q_j^m(x_j^m, x_{-j}^m) - K_m \tag{4}$$

After the chain enters and choose its ticket size, independent restaurants make sequential entry and ticket size decisions. Firms continue to enter until no further entry is profitable. Each consumer with ticket size preference α_i then chooses a restaurant to maximize her utility.

Loertscher and Muehlheusser (2011) show that, for some classes of distributions f_m, there is a unique set of subgame-perfect equilibrium locations. In particular, they show that this result holds for both monotone densities and symmetric hump-shaped densities (under some conditions). In equilibrium, independent firms choose locations to maximally deter entry of future firms in the nearby interval. In Appendix **??**, I extend this result to my setting and show that it can be used to characterize the best response functions of independent firms. The resulting equilibrium has independent firms more tightly spaced where demand is thickest and more sparse in the tails.

The chain makes the first entry decision in the game, correctly anticipating the equilibrium ticket size decisions of the independent firms that move later. Rather than characterizing the chain's strategy analytically, I perform a grid search over a set of chain locations to find the optimum. When the chain competes in only one market, it earns the highest profits by locating near the mode of f, where most consumers are located. Intuitively, its demand advantage δ is most valuable where demand is thickest, since it deters entry by subsequent independent firms in an interval around its location. Absent the chain's entry in this region, independent firms would be tightly clustered together to serve the relatively large mass of consumers.

When the chain enters in many markets at the same ticket size, it optimizes the sum of its profits across all markets. When demand is heterogeneous, it must trade off profits in one market with those in another, choosing a point where demand is thickest in the average market. As heterogeneity across markets increases, the strategic constraint that the chain must offer the same product and pricing everywhere becomes more costly.

The model allows for two supply margins that the demand model does not incorporate. First, it allows incumbent firms to change their position in product space in response to a move by the chain. In the demand-based counterfactuals, if the chain sets a higher quality level in a particular market, all incumbent product positions remain fixed. The entry model instead predicts that some high quality incumbents will change their product choices to avoid the intensified competition as a result of the chain's adjustment. Second, the model allows for adjustment along the extensive margin through entry or exit of independent restaurants. A change in the chain's product characteristic choice may cause the total number of restaurants in the markets to adjust, depending on the size of its demand advantage and thus the degree to which it deters potential entrants.

5.2 Calibration

I now turn to calibrate the entry model using the sample of small town restaurant transactions described in Section 2.

5.2.1 Demand

Each consumer chooses between the set of restaurants available in the town. I estimate the following:

$$u_{ijt} = \delta 1(j = chain) - \tau(\alpha_i - x_j)^2 + \epsilon_{ijt} \tag{5}$$

via maximum likelihood, pooling observations from all towns and using the two-step estimation procedure from Heckman and MaCurdy (1980). I show the parameter results in Table 8. The estimates of $\delta_{l(j),y(i)}$ (the chain premium) and τ (the cost of visiting a higher or lower priced restaurant than an individual's bliss point) using this simplified model are similar in magnitude to those obtained from the richer demand specification in Section 4.

5.2.2 Entry Model

I estimate the entry model for each small town. I set δ to the average δ_{1001+} (averaging across income groups) shown in Table 8, and assume that $\log(\alpha_i^m) \sim N(\mu_m, \sigma_m^2)$. To estimate the parameters of the consumer taste distribution μ_m and σ_m^2, I fit a lognormal distribution to the values of α_i, the consumer-specific ticket blisspoints I estimate, for each town.

Conditional on the utility parameters and the set of consumers in a market m, my model implies a monotonic mapping between the sunk cost of entry K_m to the number of firms N_m that enter in equilibrium. Since N_m is discrete, there is a range of K_m consistent with the observed N_m. In estimating the model, I set $K_m = (K^{min} + K^{max})/2$. All cost in my model is fixed and consumer demand in each market integrates to one, so K_m can be interpreted as the minimum market share a firm must get to enter.

5.2.3 Counterfactual Results

I now revisit the economic questions posed in the introduction through the lens of the model. First, I evaluate the costs and benefits of the chain's strategy to sell a common set of products across locations under a unified brand. Next, I examine a hypothetical ban of large chains in each of the small-town markets I study. In practice, many such bans have been implemented in small, relatively isolated places.[24] The effects of these policies on consumers may be particularly stark in this setting, where nearby towns that permit chains are geographically distant.

For each counterfactual, I first compute a version based solely on the demand-side estimates, following the approach described in Section 4 in which I hold the ticket size decisions and entry locations of other firms fixed. I then solve for the full equilibrium using the entry model and compare the estimates under the two approaches to assess the importance of the endogenous supply response.

Using the estimated parameters, I consider two alternate scenarios to quantify the cost of standardization for a chain relative to its demand-side benefits. As a benchmark, I calculate the chain's profits under the "status quo"—the chain gets some benefit in attracting consumers as a result of its higher demand (reflected in δ), but it is constrained to pick the same ticket size in each market. Then I compare the chain's constrained profits to two counterfactuals. First, I consider a chain's profits if it could maintain its demand advantage but choose its ticket size unconstrained in each market. Second, I calculate the chain's profits if were to operate in each market as an independent, forfeiting its demand advantage but choosing unconstrained. In this scenario, the chain keeps its first mover advantage, which allows it to occupy the most profitable location in the no-chain equilibrium.

In Figure 10, I show the chain's sum of profits across all markets as a function of its chosen ticket size. At the parameters I estimate, the chain maximizes profits by choosing a ticket size of about \$13, close to the mean ticket size I observe in the data across all large chains shown in Figure 2.

The cost of being a chain within the framework of the model comes from the heterogeneity

[24]For example, the Institute for Local Self Reliance reports that policies to limit retail chains are in place in Arcata, California; McCall, Idaho; Port Townsend, Washington; and Coronado, California.

of consumer tastes across markets. This heterogeneity is quantitatively important within the set of small town markets I study. In Figure 11, I show the histogram of optimal ticket sizes that the chain would choose if it picked unconstrained, with the dotted line indicating the optimal ticket size when the chain chooses under the constraint. The figure shows that the constrained choice of $13 is optimal in the modal market, but that there is significant dispersion. In particular, there is substantial mass at the right tail; in 44 of the 418 towns, the chain would set a ticket size higher than $25 if it could choose unconstrained.

Next I quantify the effect of this heterogeneity in consumer preferences on the chain's revenues. In Figure 12a, I plot the histogram of the ratio of constrained to unconstrained transactions in each town. The plot shows that the chain gets more than 80% of its unconstrained maximum revenue in about half of the markets, and more than 90% in about 20% of markets. However, there are also towns in which the chain is very poorly positioned relative to consumer tastes; the chain gets less than 50% of the unconstrained maximum level of transactions in about 10% of the towns. On average, the chain operating with the standardization constraint gets 79% of its unconstrained revenues across all sample towns.

Figure 12b plots the ratio of competitive transactions to unconstrained chain transactions and Figure 12c plots the ratio of competitive transactions to constrained transactions in each small town market. On average, if the chain were to give up its demand advantage δ and operate as a set of (unconstrained) independent restaurants instead of a constrained chain, it would keep only about 44% of its overall transactions. However, Figure 12c shows that its transaction volume would actually increase in about 15% of markets.

To quantify the role of the supply response of independent firms, I recompute the counterfactuals using this sample with the demand-based approach used in Section 4. In doing the demand-based counterfactuals, I predict transaction volumes for each chain if it were to choose its product characteristics without constraint but keep its demand advantage δ or operate as an independent in each market, holding fixed the ticket size and entry decisions of all other firms. I compare the predicted number of transactions under these scenarios to the constrained optimum, in which each chain chooses one ticket size across all markets in which it has entered. I show the results in Table 9.

Using the entry model, I find that the ratio of constrained to unconstrained profits is lower than in the demand model, suggesting that the cost of standardization is higher. The intuition for this result is that the chain has a significant entry deterrence effect on other firms that is shut down in the demand-based approach. In the entry model, when the chain picks unconstrained it not only chooses the most profitable location, it also induces other firms not to locate nearby, since its demand advantage δ will attract nearby consumers and thus make it less attractive for independent firms to enter.

In a similar vein, the demand counterfactuals find a lower benefit of being a chain; the estimates from the entry model in Table 9 imply that the chain would keep only 35% of its transactions if it were to operate as a network of independent stores, relative to 63% under the demand scenarios. Again, the discrepancy results from the fact that the demand counterfactual does not account for the stiffer competition the chain would face, as other independent firms would enter at similar ticket sizes absent the deterrent effect of the highly demanded chain firm.

I also revisit the chain ban counterfactuals using the entry model. I compute the difference in consumer welfare between a scenario in which the chain enters in every market (choosing one ticket size to maximize its total number of transactions across all markets) and one in which the chain does not enter. I again quantify the difference in welfare in consumer mile-equivalents.[25] I show the results of this counterfactual exercise compared to the demand-based approach in Section 4 in Table 10. Averaged across all 418 small towns, chain bans decrease welfare by about 0.09 mile equivalents per transaction, compared to 0.13 mile equivalents from the demand estimates. The average consumer in this sample transacts about 30 times, implying an annual loss of 2.27 mile equivalents per consumer. Using only the demand estimates, I estimate a larger cost of these policies—between 3.5 and 10.7 miles per consumer.

[25] Restaurants in the entry model are not spatially differentiated. To transform utility into mile-equivalent units, I instead assume that the ratio between the cost of physical distance from the model I estimate in Section 4 γ and τ, the utility cost of visiting a restaurant far from a consumer's blisspoint α_i, are the same in the small town sample. I scale differences in consumer surplus by the ratio $\gamma^{urban} \cdot \tau^{town} / \tau^{urban}$.

6 Channels

My demand estimates imply that chain restaurants have a large and important demand advantage over smaller firms. In my empirical specification, I estimate this demand advantage using a fixed effect that is shared across all chain restaurants with a given number of locations. Under the assumptions of the model, this fixed effect measures the total additional utility that consumers receive from eating at chain restaurants relative to an independent. However, this approach does not provide a way to determine what aspect of the chain is valuable for consumers.

In this section, I try to distinguish between two broad categories of channels. First, chains may have informational advantages in attracting consumers; consumers may be more likely to recognize the chain's brand and know about its products because of advertising.[26] Second, upon visiting the restaurant, consumers may enjoy the food or experience more at a chain relative to a similar independent. I show suggestive evidence that a key driver of chain success is their familiarity among consumers. First, new chain restaurants acquire customers more quickly than independent restaurants, suggesting that consumers take time to learn about new independents. Second, among consumers that move cities, the chain premium persists only for chains that were familiar in their new city.

6.1 Evidence from entry

First, I show evidence from new restaurant entries. If the underlying utility that a consumer gets conditional on visiting a particular restaurant is constant over time, then growth in the sales of a new restaurant is primarily driven by changes in information. If chain restaurants have an advantage over independents in communicating their reputation and product attributes to consumers, then they may acquire customers more quickly.

To study this, I look at a sample of about 600 restaurants that opened in 2015 in the seven cities included in my urban consumer sample and calculate their monthly sales through the

[26] Existing evidence suggests that incomplete information may be an important reason why some restaurants are more successful than others. For example, Luca (2011) finds evidence that favorable restaurant reviews on Yelp lead to significantly higher revenue for independent restaurants, but do not affect chains, about which consumers conceivably have more familiarity.

first three years after entry. In Figure 13a, I plot the average monthly sales in year 0, year 1, and year 2 for restaurants that were still open by the end of the third year, with average year 0 sales normalized to one. The plot shows that large chains quickly reach their steady state level of sales in the first 12 months after opening, but that independent restaurants grow over the first three years. The revenue of independent restaurants grew nearly 40% between the year 1 and year 0, and an additional 13% between year 1 and year 2. In contrast, large chains grew only 11% between year 0 and year 1, and 7% between year 1 and year 2. Figure 13b shows the sales in each month since opening for this same sample. The takeaway is similar—independents grow steadily in each month, while the trajectory of medium chains (with 101-1000 locations) and large chains (more than 1000 locations) are relatively flat.

6.2 Evidence from movers

Second, I study the behavior of consumers that move from one city to another. Because there is variation in the availability of chains across cities, moving sharply changes a consumer's choice set while her information and accumulated brand awareness of these new choices adjusts more slowly. If the chain demand premium is primarily driven by selection, then we should expect large chains to be of equal quality on average, and consumers should be as likely to visit chains who only have locations in the new city as those that were also available in the old city. If the main driver of the chain premium is the branding of the chain, restaurants that were unavailable in the consumer's prior city should have a smaller demand premium in the first few months after moving.

To test this hypothesis, I find a sample of cards that move between two of the seven cities included in my urban consumer sample. I define a mover as a card that primarily transacts in one MSA for a period of at least three months before beginning to transact in a new MSA for at least three months, without returning to the first MSA, and without going to a third MSA. Using these criteria, I identify approximately 2500 cards that move between the seven cities between 2015 and 2017 that made at least 5 transactions at burger restaurants, for a total of about 35,000 transactions between the pre-move period and the first twelve months after moving. I focus on burger restaurants in this analysis; burgers are one of the largest restaurant categories in my data, are widely available in every sample city, and exhibit substantial variation in the availability of large chains across cities. Even among burger chains with more than 1000 nationwide locations, about half were unavailable in at least one city.

I estimate the following equation using the sample of transactions at burger restaurants by cards that move:

$$u_{ijt} = \delta_{l(j),c} + \delta^{avail}_{l(j)} + \delta^{unavail}_{l(j)} - \tau(\alpha_i - x_j)^2 - \gamma dist_{ij} + \epsilon_{ijt} \tag{6}$$

$\delta_{l(j),c}$ is a chain size fixed effect that varies across location bins $l(j)$ and city c. $\delta^{avail}_{l(j)}$ and $\delta^{unavail}_{l(j)}$ are indicators that are equal to one in the twelve months after the move for chains that were available and unavailable in the origin city, respectively. I consider a chain available if it operated at least one location. As in previous sections, α_i is a consumer-specific preference over restaurant ticket size, x_j is the average ticket size of restaurant j, and $dist_{ij}$ is the distance in miles between card i and restaurant j. I define a card's location separately before and after the move as the transaction-weighted centroid of the zipcodes in which it transacts within its current MSA (defined as the origin in the pre-move period and the destination after).

I show the coefficients $\delta_{l(j),c}$, $\delta^{avail}_{l(j)}$, and $\delta^{unavail}_{l(j)}$ from equation 6 in Figure 14. The figure shows that the chain premium for burger chains that were available in both the consumer's origin and destination cities was virtually unchanged between the pre-move and post-move period. However for unfamiliar chains, those that were unavailable in the origin city, the chain premium fell by about 75%. This result suggests that the bulk of the advantage of being a chain comes from the power of the chain's brand.

7 Conclusion

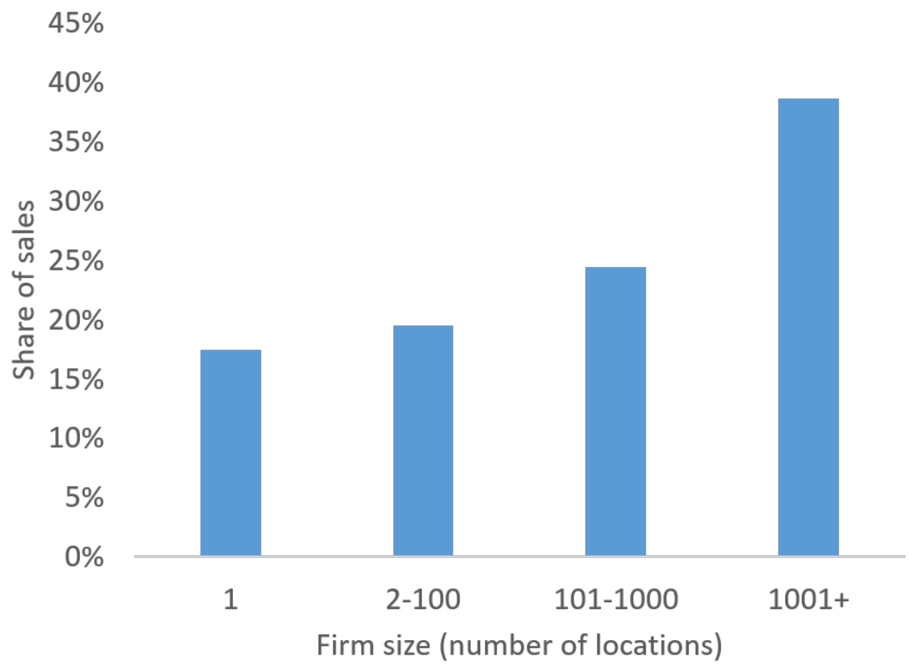
In this paper, I study the role of large chains in the restaurant market. In particular, I consider the costs and benefits to standardization across chain outlets. Using a novel dataset that includes about half of credit and debit transactions in the United States, I find that standardized chains have a large demand advantage compared to smaller firms, but that the standardization strategy has an important cost. Because consumers are heterogeneous across markets, a chain that standardizes sacrifices a significant amount of potential revenue in some markets. I find that if chains could choose their product characteristics flexibly in each market but keep their demand, they would get between 12% and 22% more transactions. However, if they were to give up that demand advantage in exchange for flexibility, they would lose between 13% and 21%. Taking into account the endogenous supply response of competing firms suggests that the costs and benefits to being a chain are roughly twice the size as counterfactuals that hold the behavior of other firms fixed.

Despite the advantages of chains, small chains and independent firms still account for a significant portion of restaurant sales. My work suggests that differences in consumer tastes across markets are an important reason for this. Absent heterogeneity across markets, we might expect to see differentiated large chains covering nearly all of demand, even when consumer tastes are dispersed within a market. Restaurants are naturally limited in their ability to cater to different tastes within the same outlet; restaurants that offer food from many cuisine types at many price levels are uncommon. Retail categories in which chains account for nearly all of sales, like in the general merchandise category, tend to offer a wide array of different products, and thus the natural constraint that standardization imposes may not be as important in those sectors.

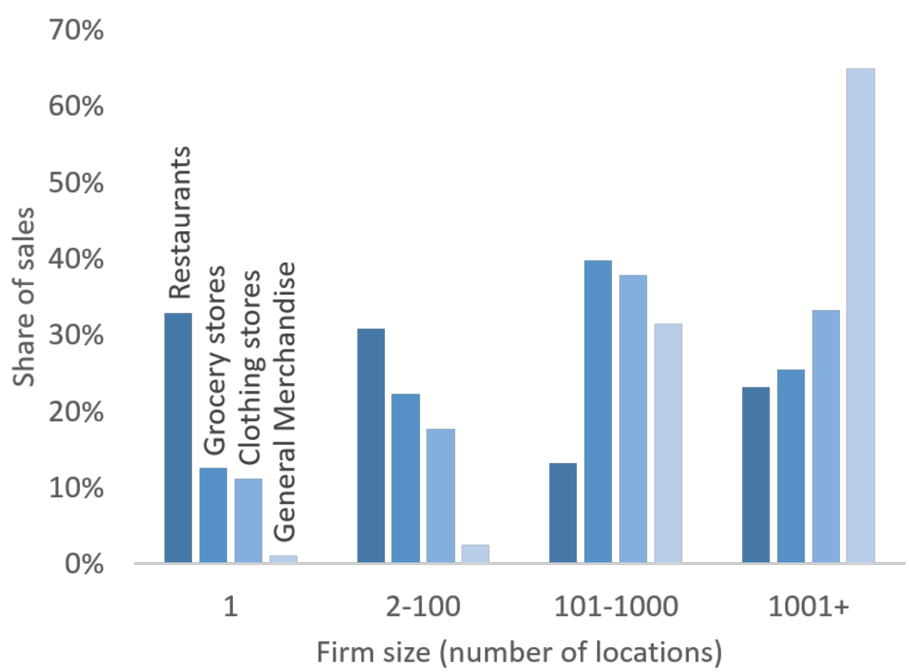
I also quantify the effects of a ban on chain restaurants on consumer welfare. I find that such a ban would decrease consumer welfare by an amount equivalent to making each consumers in my sample travel an additional one to two miles. Further, these bans disproportionately impact lower income consumers, who have the strongest preferences for chains. The magnitudes of my estimates suggest that bans of chain stores are an inefficient way to transfer surplus to small firms, and thus are justified only if independent firms are associated with large externalities.

Figure 1: Share of chains

(a) Aggregate share of spending by number of firm locations in augmented retail categories

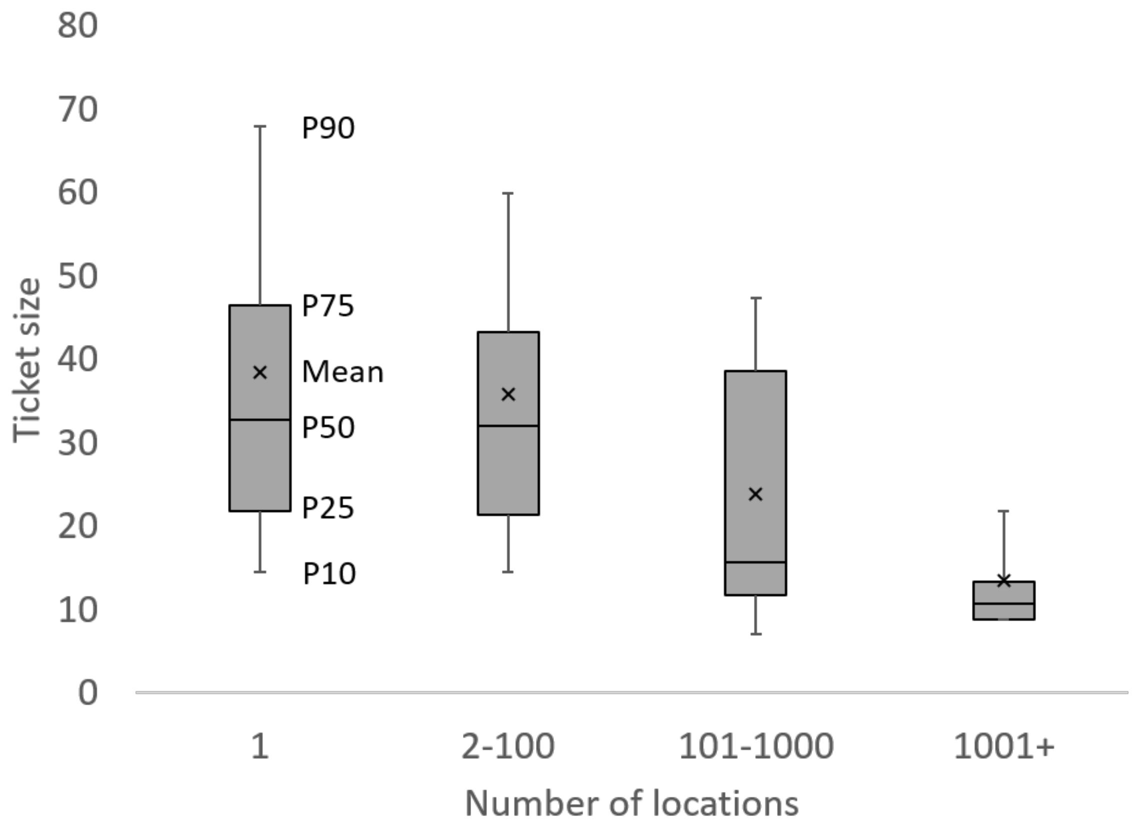


(b) Share of spending by number of firm locations by merchant category



The figure shows the share of US payment card spending in 2016 that went to merchants with the number of locations given on the x-axis. Panel (a) shows the aggregate shares by firm size in the retail, restaurants, and hotel categories (corresponding to NAICS codes beginning with 44, 45, and 72) . The largest categories wihin this set by share of spending were restaurants (25%), grocery stores (16%), and general merchandise stores (13%). Panel (b) shows the share of spending by firm size for four large retail categories. Each group of bars corresponds to a firm size bin, while each bar within the group gives the share of spending within a firm category that went to firms with that number of locations.

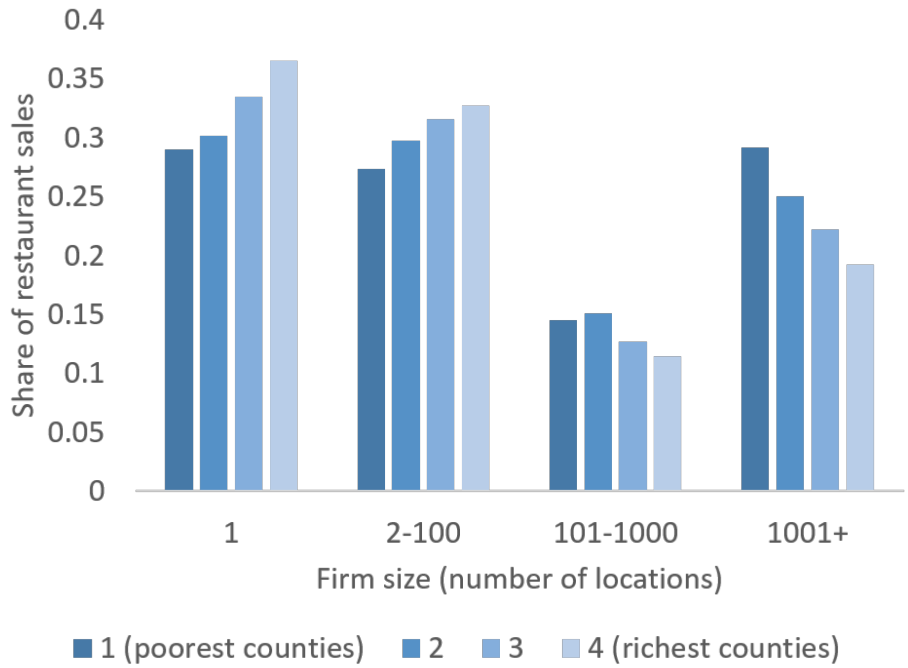
Figure 2: Quantiles of average restaurant ticket size by chain size



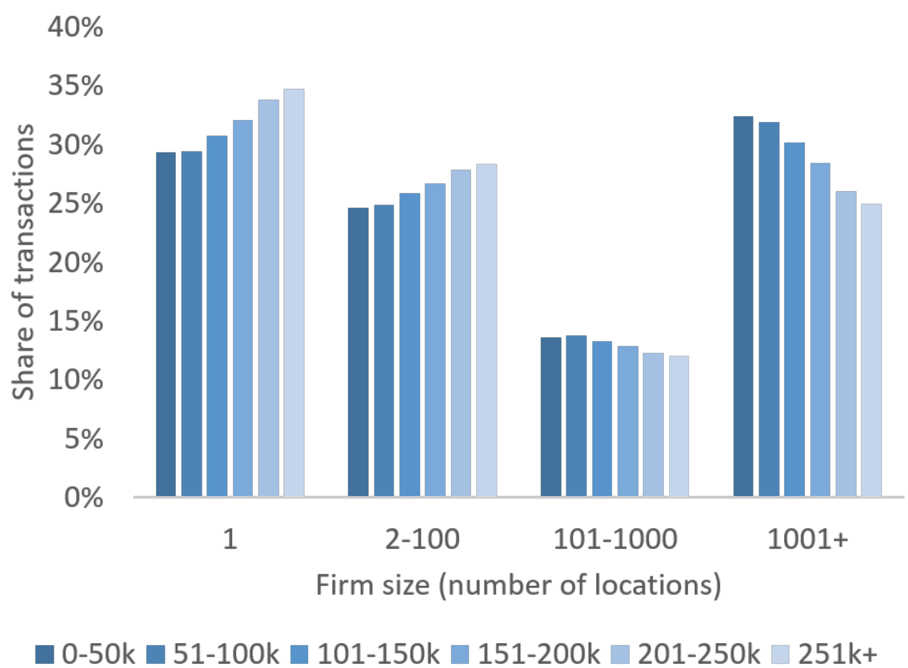
The figure shows the mean and several quantiles of the distribution of average restaurant ticket size. An observation in the underlying dataset is a restaurant in 2016. I compute average ticket size for each restaurant as the sum of dollars spent divided by the number of transactions. I compute the average and quantiles of restaurant ticket size across all US restaurants, weighting each restaurant by its number of transactions in 2016. The x-axis gives the number of nationwide locations belonging to that restaurant brand. The edges of the middle shaded box correspond to the average ticket size of the restaurant that accounted for the 25th, 50th, and 75th percentiles of transactions. The lines extending below and above the box give the 10th and 90th percentiles of restaurant ticket size. The x mark in the interior of the box indicates the mean of the distribution.

Figure 3: Chain share of restaurants by income

(a) Share of chain restaurants by quartile of county income

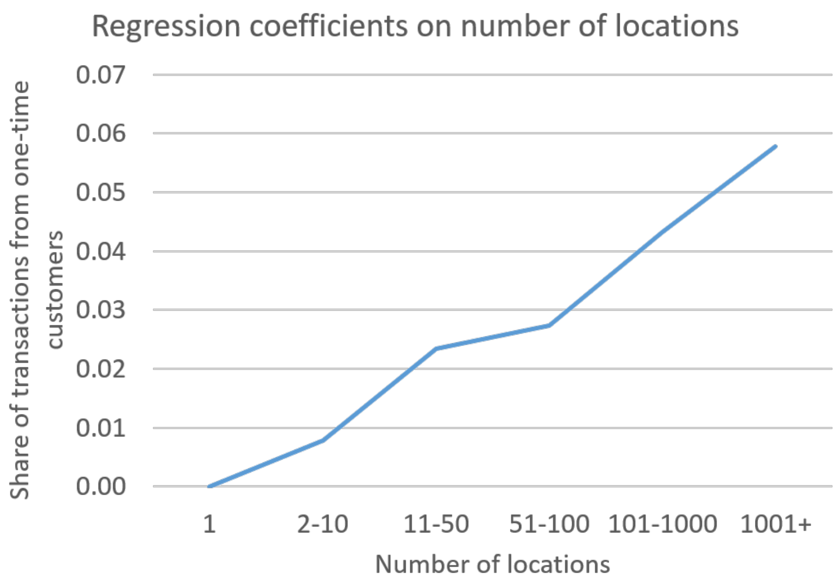


(b) Share of chain restaurants by annual cardholder income



The figure shows the share of all US restaurant spending in the payment card data in 2016 that went to merchants with the number of US locations given on the x-axis by quartile of county income. Each group of bars corresponds to a firm size bin. In Panel (a), each bar within the group gives the share of spending within a firm category by the quartile of that county's income. Counties are placed in quartiles so that each quartile contains about 25% of the US population. In Panel (b), each bar within the group gives the share of spending within a firm category by the card's annual income.

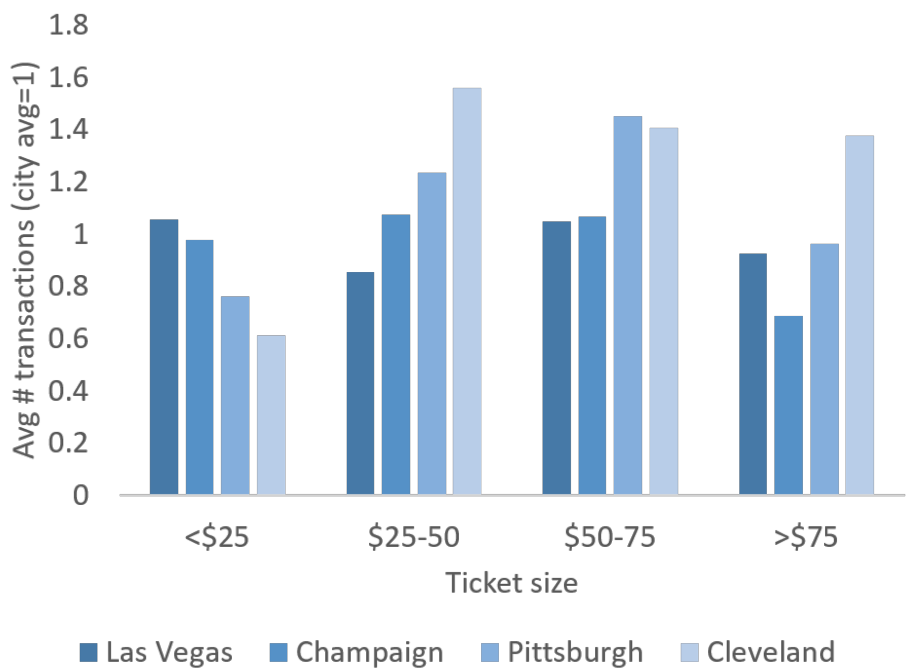
Figure 4: Share of transactions from one-time customers by firm size



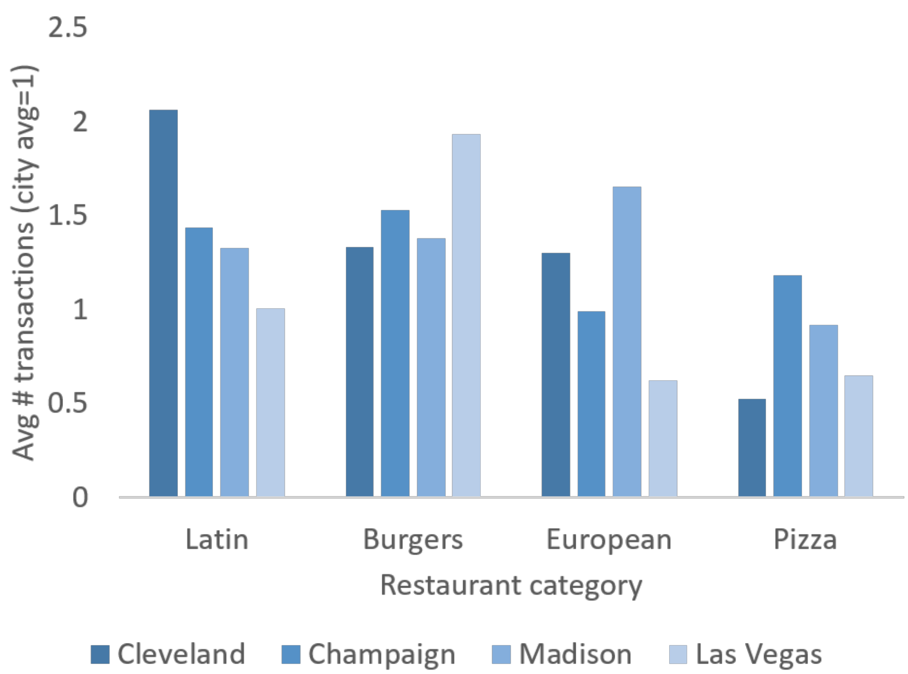
The figure plots the differences in the share of transactions that were made by cards that only transact once at a given restaurant in a six-month period in 2016, averaged across all US restaurants and calculated separately for each firm size bin. The y-axis shows the firm sized fixed effects β_l from the regression $sh_{iz}^{1trans} = \alpha_m + \alpha_{ts} + \alpha_z + \beta_l$, where sh_{iz}^{1trans} is the share of one-time customers for restaurant outlet i in zipcode z. α_m, α_{ts}, and α_z are 6-digit NAICS, ticket size quartile, and zipcode fixed effects, respectively. I normalize the share of one-time customers at restaurants with one location to zero.

Figure 5: Average number of transactions by ticket size and cuisine type across cities

(a) Per-restaurant transactions by average transaction size and city

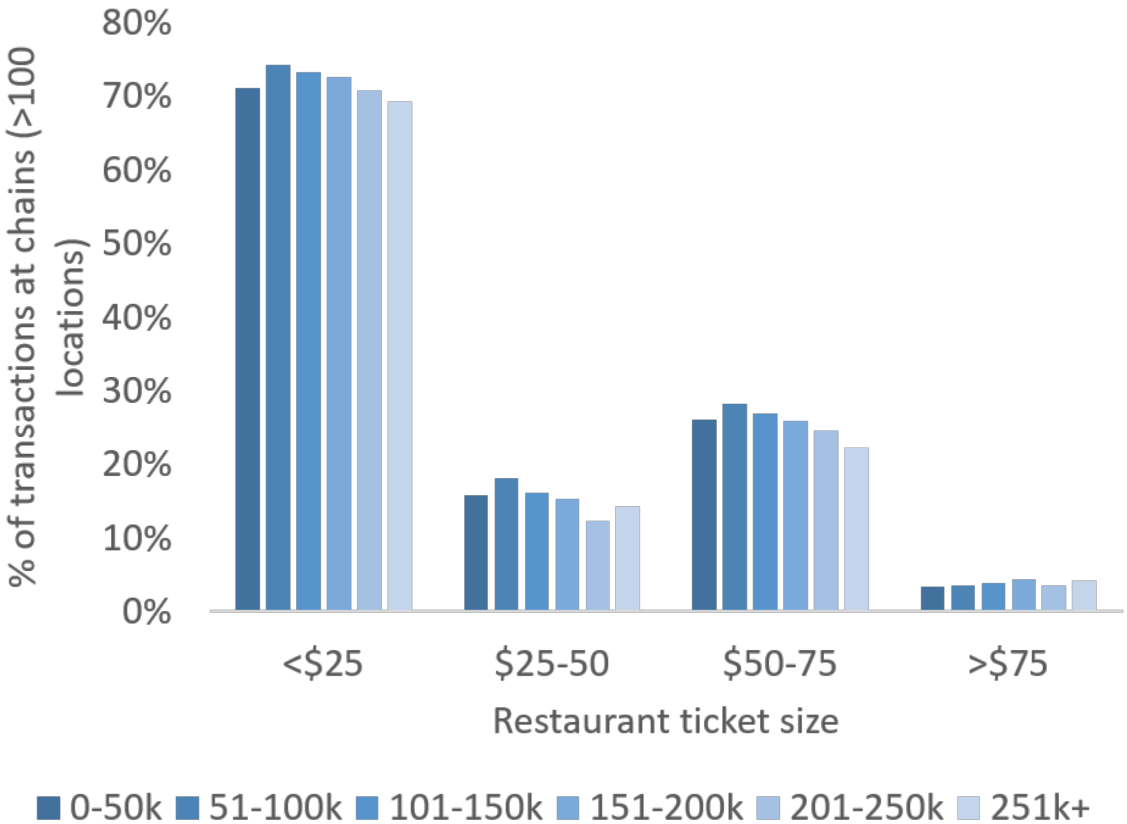


(b) Per-restaurant transactions by cuisine type and city



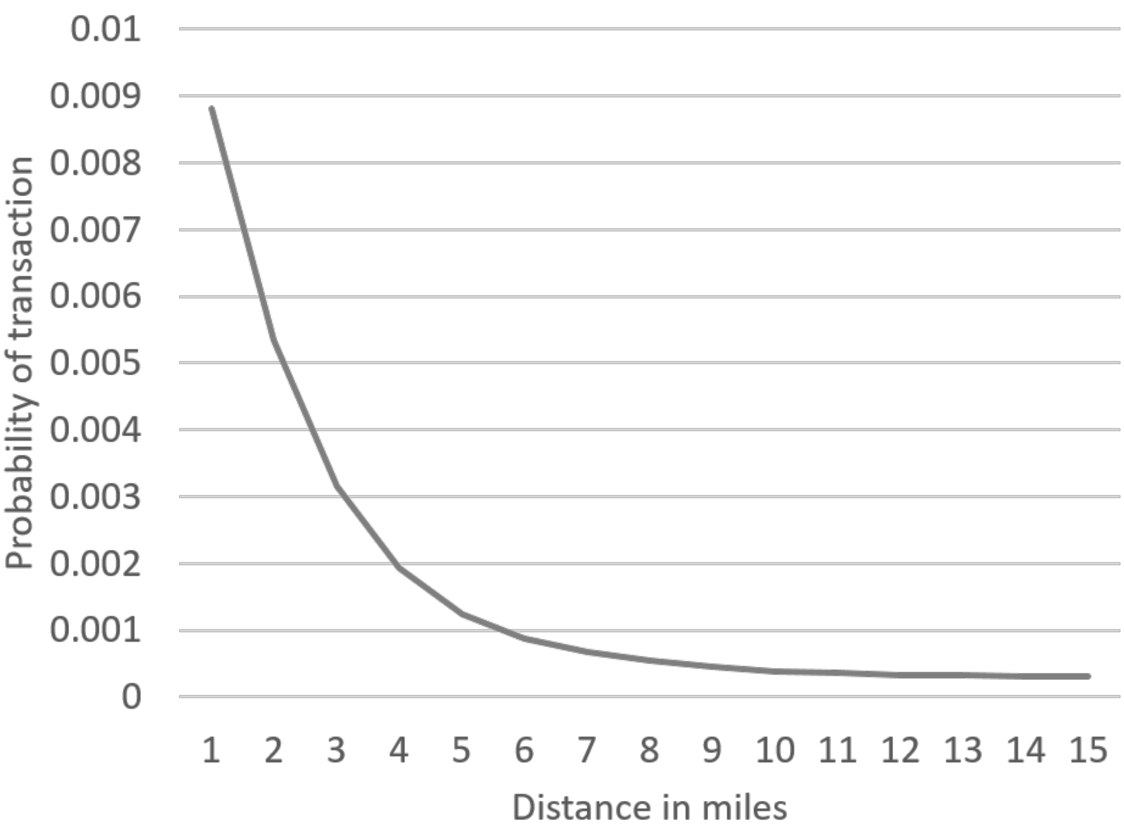
The figure shows the relative success of restaurants with different vertical and horizontal types in four cities included in the urban consumers sample described in Section 2. The value on the y-axis gives the average number of transactions in 2016 for a restaurant in that category divided by the transactions that the average restaurant received in that city. Panel (a) divides restaurants into different vertical types, measured by their average transaction size (defined at the brand level). Each group of bars denotes a restaurant ticket size bin, with each bar within the group corresponding to a city. Panel (b) divides restaurants into horizontal types, corresponding to their cusine type. Each group of bars denotes a cuisine type, with each bar within the group corresponding to a city.

Figure 6: Share of transactions at chains with more than 100 locations by income group and ticket size



The figure shows the share of transactions for each cardholder income group that went to chains with more than 100 US locations for restaurants of different average transaction sizes within the urban consumer sample. Each group of bars corresponds to the restaurant's average transaction size in $25 bins, which is defined at the chain level. Each bar within the group gives the share of transactions among restaurants in that ticket size bin that were at chains with more than 100 locations by cardholder income group.

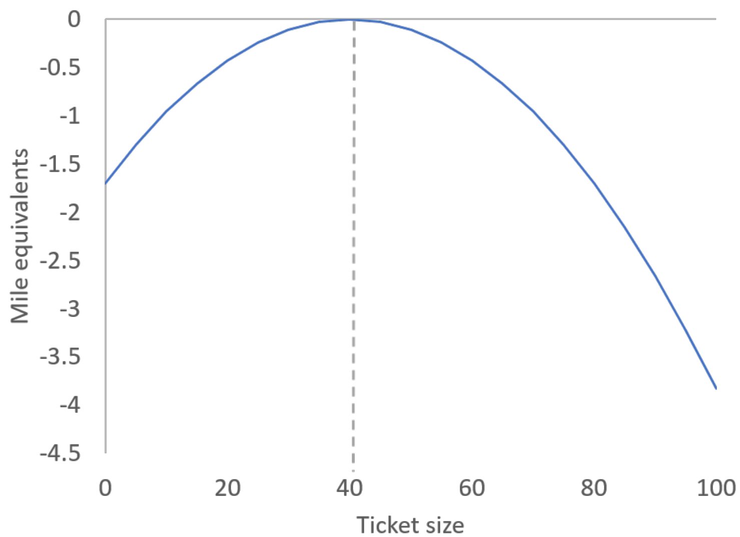
Figure 7: Probability that a restaurant is chosen as a function of distance



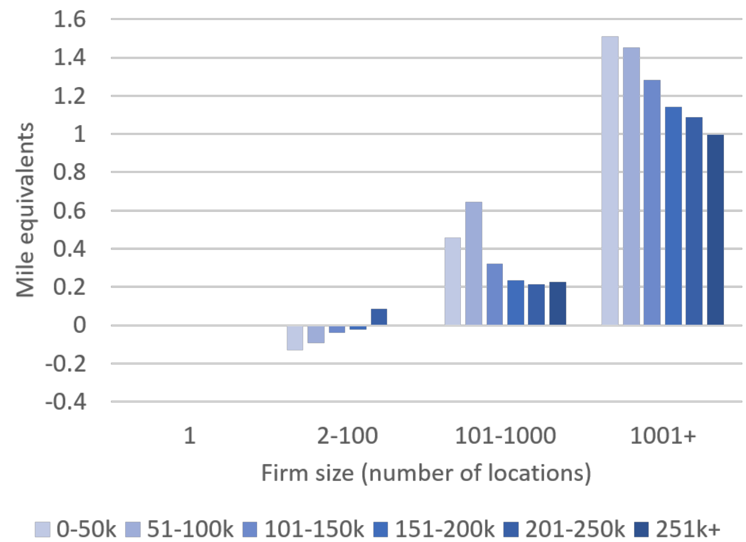
The figure shows the probability that a given restaurant is chosen as a function of the distance between the restaurant and the consumer's shopping location (defined as the coordinates corresponding to the transaction-weighted average of the zip centroids in which it transacts). Each observation in the underlying data is a restaurant-consumer-trip combination for a given cardholder in my urban consumer sample. I average across all possible restaurant choices and consumers in the sample for each one mile bin. For example, the first point in the graph shows that restaurants that were between zero and one miles away from the consumer were chosen about 0.9% of the time.

Figure 8: Demand estimates

(a) Ticket size penalty in mile equivalents



(b) Chain premium in mile equivalents



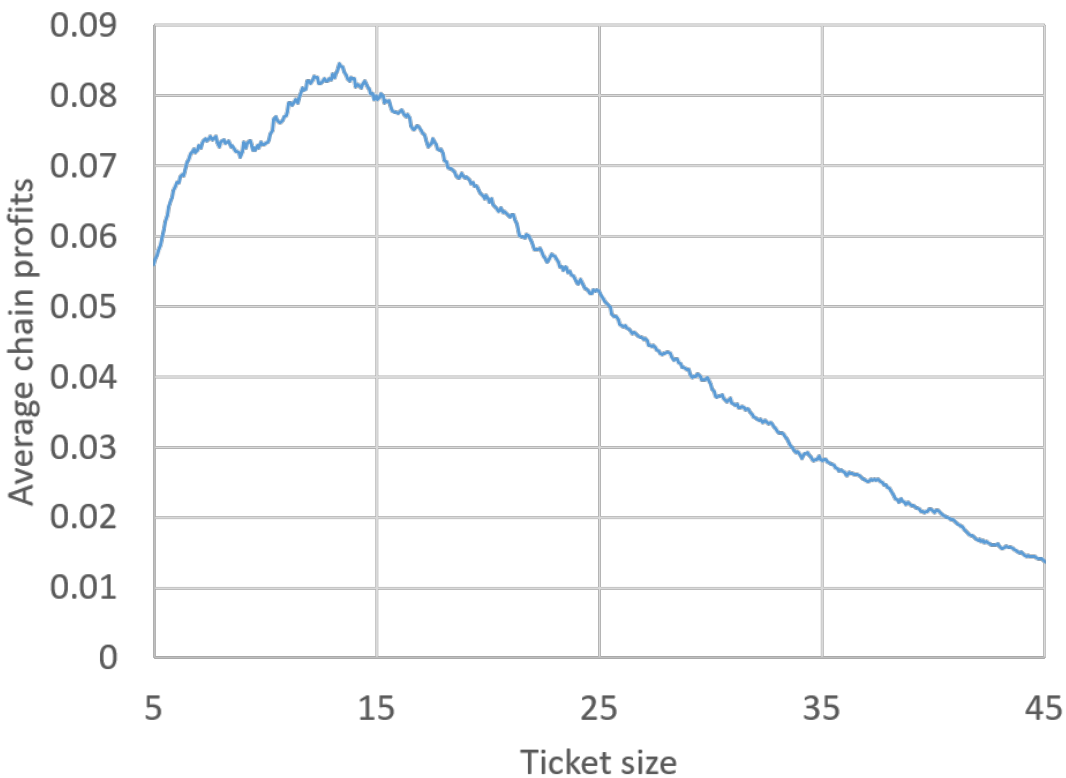
Panel (a) shows the estimated disutility that a consumer gets from visiting a restaurant with a ticket size different from her blisspoint in units of miles of physical distance. For a card with ticket size blisspoint $\alpha_i = \$40$ that eats at a restaurant with a given ticket size x_j, I calculate the ticket size penalty as $\tau(\alpha_i - x_j)^2$ divided by the utility cost of traveling one mile γ, where each parameter is set to the average value across cities estimated in Section 4. The vertical dotted line shows the ticket size blisspoint α_i of the card. For example, the point on the solid line at $x = 20$ indicates that a consumer with a ticket size blisspoint of \$40 is indifferent between eating at a restaurant with ticket size \$30 or traveling 0.5 additional miles to eat at a restaurant with ticket size \$40. Panel (b) shows the additional utility that a consumer receives from visiting a chain restaurant relative to an independent restaurant as a function of her income. To calculate this, I take the firm size fixed effect estimates of $\delta_{y(i),l(j)}$ (where $y(i)$ indexes consumer income and $l(j)$ indexes number of locations) recovered from the demand estimation described in section 4, divided by the cardholder cost of traveling one mile γ. I use the average parameter value taken across cities. The fixed effect for firms with one location is normalized to zero. Each group of bars corresponds to a firm size bin, while each bar within the group gives the normalized fixed effect corresponding to cardholders with a given estimated household income.

Figure 9: Category heatmap with ranks

	American	Asian	Burgers	European	Latin	Pizza	Sandwiches	Other
Champaign	6	7	1	5	2	3	8	4
Charlotte	2	6	3	7	1	5	8	4
Cleveland	5	7	2	6	1	8	3	4
Las Vegas	4	3	1	7	2	6	8	5
Madison	4	7	3	1	2	6	8	5
Phoenix	1	4	2	5	3	6	8	7
Pittsburgh	3	6	2	4	1	7	8	5

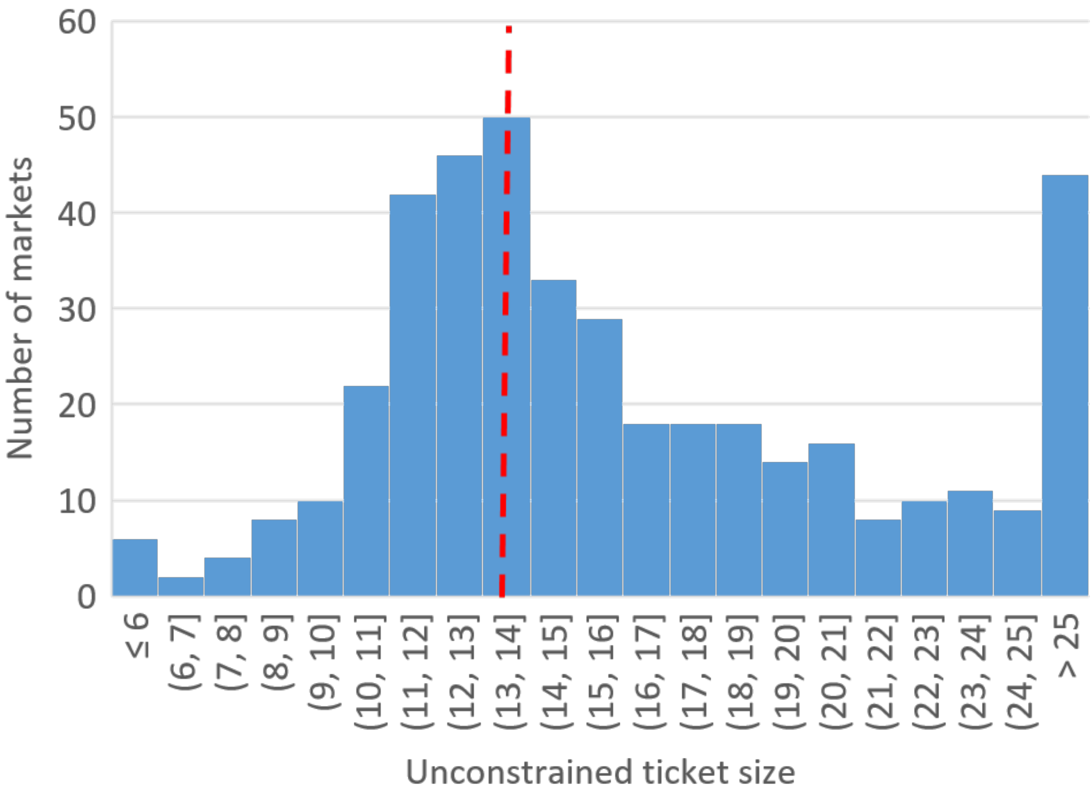
The figure shows the rankings of the horizontal category fixed effects $\theta_{y(i)}$ in each city, averaged across income groups. The ranking of each category is done for each city, so that a rank of 1 indicates that the category was the top-ranked category in the city and a rank of 8 indicates that it was the lowest ranked category in the city. The heat map is color coded with dark blue boxes corresponding to higher rankings (more demanded categories) while white boxes correspond to lower rankings (less demanded).

Figure 10: Sum of chain profits across markets as a function of ticket size



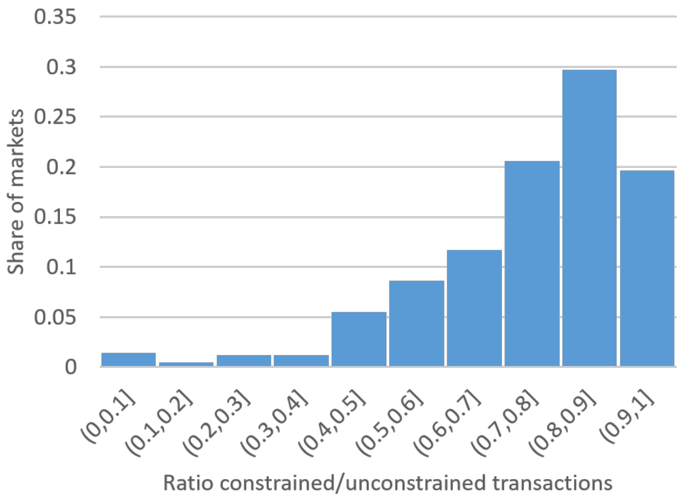
The figure shows the weighted average of chain profits across all 418 markets in the small town sample as a function of the chain's constrained ticket size. The entry model used to calculate chain profits under the constraint is given in Section 5. In computing the average across markets, each market is weighted by its total number of restaurant transactions.

Figure 11: Distribution of unconstrained chain location across markets

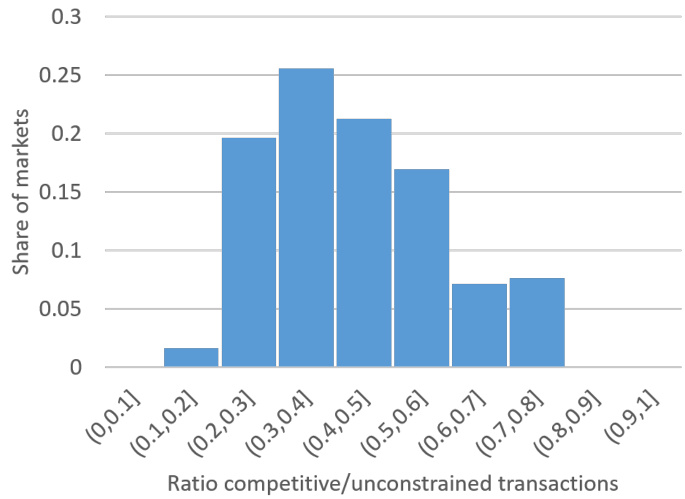


The figure shows the distribution of the unconstrained optimal ticket size in each market in the small town sample computed using the entry model described in Section 5. The vertical dotted line shows the constrained optimum ticket size when the chain must pick the same ticket size in each market. Each blue bar gives the number of markets in which the optimal unconstrained ticket size lies in that ticket size interval.

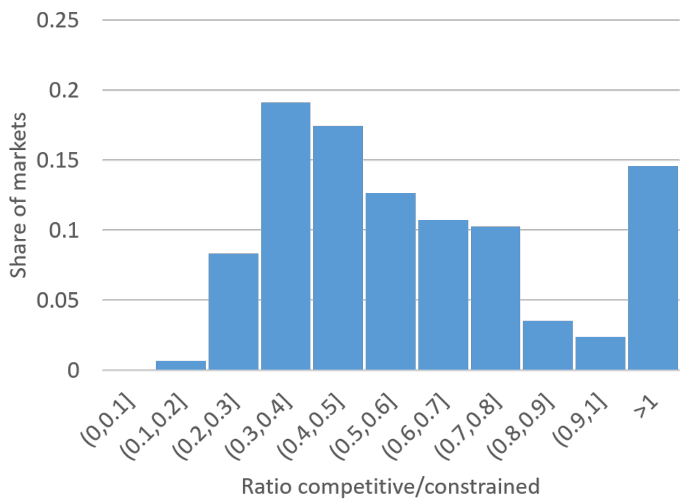
Figure 12: Distribution of chain transaction volume by market from entry model



(a) Ratio of constrained to unconstrained transaction volume by market



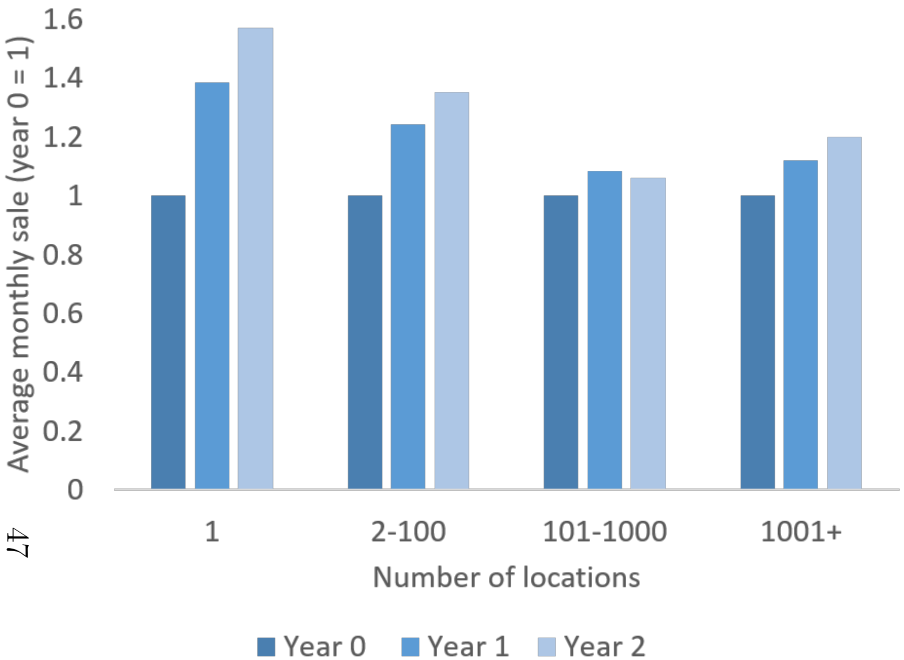
(b) Ratio of non-chain to chain unconstrained transaction volume by market



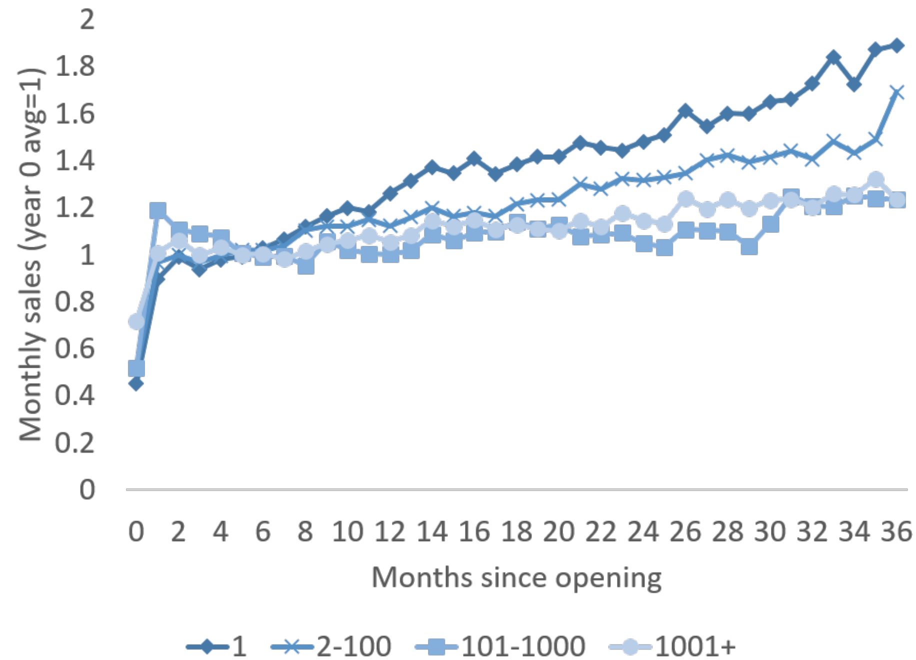
(c) Ratio of non-chain to chain constrained transaction volume by market

The figure shows the distribution of the ratios between predicted transactions using the entry model described in Section 5 under three scenarios: the chain chooses one ticket size in all markets (constrained), chooses flexibly in each market but maintains the chain demand advantage δ (unconstrained), and operates as an independent that chooses ticket size separately in each market but loses δ (competitive). Each data point represented in the histogram is one market in the small town sample.

Figure 13: New restaurant sales during first three years after opening by chain size



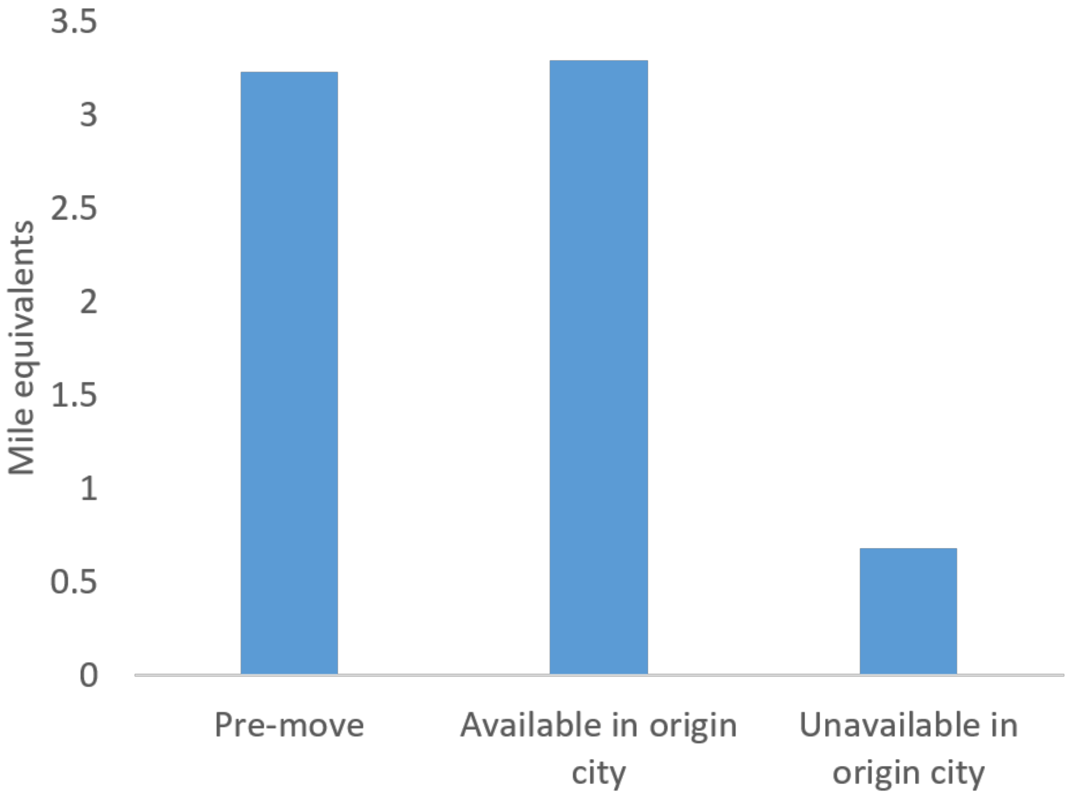
(a) Average monthly sales during first three years



(b) First 36 months of new store sales

The figure shows average monthly sales for new restaurants during the first three years after entry. The sample of entries represented in the figure includes new restaurant entrants that opened in 2015 in one of the seven midsize cities described in Section 2. I plot their sales between the month of their first transaction in 2015 until the 36th month after they open (until May 2018). I identify entrants as restaurant locations that had both their first transactions and their first Yelp review within three months of each other. I drop firms that did not survive until the third year, including firms that were marked closed by Yelp in 2017 or averaged less than 100 transactions per month in either their second or third years open. Panel (a) shows the average monthly sales by year, where year 0 indicates average sales in months 1 to 12, year 1 between months 13 and 24, and year 2 between months 25 and 36. I normalize the average year 0 sales to 1 for each store and plot the average ratio of year 1 to year 0 sales and year 2 to year 0 sales, weighting each store by their total dollars over the three years. Each group of bars corresponds to a firm size bin, where the number of locations gives the nationwide number of restaurants belonging to the brand. Panel (b) shows monthly sales over the first 36 months after opening. Each point gives the sales after a given number of months divided by the average first year sales for that restaurant. I average across restaurants within a firm size bin, weighting each restaurant by its total sales over the first 3 years.

Figure 14: Chain premium in miles for burger restaurants by availability in mover origin city



The figure shows the chain premium δ for burger chains with more 1000 than locations, calculated from a set of cards that move from one city in the urban consumer sample to another. The sample includes transactions by this set of cards at restaurants in the burger category prior to the card's move date and in the first 12 months after the move date. Each bar gives the estimated utility in miles that a card receives from visiting a chain with more than 1000 locations relative to an independent. The first bar in the figure shows the demand premium δ_{1001+} for burger chains with more than 1000 locations during the months before the card moved, averaged across cities and divided by γ, the coefficient on physical distance. The second bar shows the chain premium in the 12 months post-move for large chains that had at least one location in the mover's origin city, calculated as $(\delta_{1001+} + \delta_{1001+}^{avail})/\gamma$. The third bar shows the chain premium in the 12 months post-move for large chains that did not have locations in the mover's origin city, calculated as $(\delta_{1001+} + \delta_{1001+}^{unavail})/\gamma$.

Table 1: Summary statistics for urban consumer sample by city

City	Accounts (K)	Transactions (K)	Dollars (M)	Restaurants
Champaign	4.45	68.59	1.47	128
Charlotte	17.48	247.83	7.54	529
Cleveland	7.36	83.43	3.02	254
Las Vegas	9.80	109.87	2.98	455
Madison	12.88	183.76	5.10	270
Phoenix	25.38	339.48	8.77	771
Pittsburgh	16.26	219.54	8.15	456
Total	93.61	1,252.49	37.03	2,863

The table shows the number of accounts, transactions, dollars and restaurants included in the urban consumer sample used in Section 4 by city. I drop consumers that live more than 25 miles away from the city, as well as those that made fewer than five restaurant transactions during 2016.

Table 2: Summary statistics for urban consumer sample by household income

Household Income	Accounts	Avg transactions	Avg dollars	Avg restaurant ticket size	Avg distance
<$50k	31,889	13.59	347.80	28.21	4.45
$51-100k	30,855	13.23	384.52	29.93	4.86
$101-150k	16,361	13.16	423.73	31.31	5.03
$151-200k	6,618	13.28	446.49	32.31	5.10
$201-250k	3,023	13.79	506.84	33.64	4.91
>$250k	4,861	13.51	545.73	35.05	4.86

The table shows summary statistics on the activity of cardholders included in the urban consumer sample used in Section 4 by bin of household income. Each observation used to create the table is a card. The third and fourth columns give the average number of transactions and dollars conducted by cards in a given income group, averaged across cards and cities. The fifth column gives the average ticket size of a restaurant visited by a card in that income group. For example, the value in the first row indicates that the average transaction made by a consumer with household income under $50k occurred at a restaurant with average ticket size of $28. The sixth column gives the average distance between a consumer's location and the restaurant. The consumer's location is defined as the transaction-weighted centroids of the zipcodes in which it transacts.

Table 3: Summary statistics for urban consumer sample restaurants by cuisine type

Category	# Restaurants	Avg transactions	Avg dollars	Avg accounts
American	537	540	22,016	293
Asian	507	365	12,844	188
Burgers	449	451	7,382	221
European	115	487	26,763	297
Latin	345	545	12,160	265
Other	238	417	9,703	212
Pizza	322	379	11,790	195
Sandwiches	350	313	5,715	155

The table shows summary statistics on restaurants included in the urban consumer sample used in Section 4 by restaurant cuisine type. Each observation used to create the table is a restaurant. Dollars, accounts, and transactions are computed from sample cards (those within 25 miles of the city that made at least 5 restaurant transactions in 2016).

Table 4: Summary statistics for small town markets

	2010 Population	Accounts	Transactions	Dollars	Restaurants
Average	4,084	371	11,234	230,318	16
P10	1,495	62	1,011	21,155	5
P50	3,387	210	5,136	87,769	12
P90	8,216	889	24,319	542,045	35
Total	-	154,902	4,695,659	96,272,823	6,827

The table shows summary statistics calculated over the 418 small town markets used for the estimation in Section 5. Each observation used to create the table is a town included in the sample. Each sample town has between 1,000 and 10,000 residents, is at least ten miles from the nearest city or town of any size, is at least 25 miles from the nearest city of 50,000 people or more, and has at least four restaurants.

Table 5: Demand parameter estimates with horizontal differentiation

Parameter	Description	Average estimate	Average SE
γ	Distance cost (1 mile)	-0.400	0.001
τ	Ticket size cost	0.001	0.000
$\delta_{2-100,<50k}$	Chain utility 2-100 locations, income <50k	-0.053	0.011
$\delta_{2-100,100k}$	Chain utility 2-100 locations, income 100k	-0.037	0.012
$\delta_{2-100,150k}$	Chain utility 2-100 locations, income 150k	-0.015	0.016
$\delta_{2-100,200k}$	Chain utility 2-100 locations, income 200k	-0.009	0.026
$\delta_{2-100,250k}$	Chain utility 2-100 locations, income 250k	0.034	0.038
$\delta_{2-100,>250k}$	Chain utility 2-100 locations, income >250k	0.001	0.030
$\delta_{101-1000,<50k}$	Chain utility 101-1000 locations, income <50k	0.183	0.016
$\delta_{101-1000,100k}$	Chain utility 101-1000 locations, income 100k	0.257	0.016
$\delta_{101-1000,150k}$	Chain utility 101-1000 locations, income 150k	0.128	0.023
$\delta_{101-1000,200k}$	Chain utility 101-1000 locations, income 200k	0.094	0.038
$\delta_{101-1000,250k}$	Chain utility 101-1000 locations, income 250k	0.085	0.056
$\delta_{101-1000,>250k}$	Chain utility 101-1000 locations, income >250k	0.091	0.044
$\delta_{1001+,<50k}$	Chain utility 1001+ locations, income <50k	0.604	0.013
$\delta_{1001+,100k}$	Chain utility 1001+ locations, income 100k	0.581	0.014
$\delta_{1001+,150k}$	Chain utility 1001+ locations, income 150k	0.514	0.019
$\delta_{1001+,200k}$	Chain utility 1001+ locations, income 200k	0.457	0.031
$\delta_{1001+,250k}$	Chain utility 1001+ locations, income 250k	0.435	0.047
$\delta_{1001+,>250k}$	Chain utility 1001+ locations, income >250k	0.399	0.038

The table shows the point estimates of the parameters in equation 1 that describe the preferences of consumers over restaurants. γ is the cost of traveling one mile and τ is a parameter that scales the cost of eating at a restaurant with a ticket size higher or lower than a consumer's preference. $\delta_{y(i),l(j)}$ gives the utility that a consumer of income group $y(i)$ receives from eating at a restaurant with $l(j)$ locations (normalized to zero for restaurants with one location). I report the averages of the coefficients and standard errors (averaged across the seven sample cities) weighting each city by its total number of transactions. The standard errors given in the table above are computed using the standard MLE formula and are incorrect, as they do not account for the two-step estimation procedure I implement. Correct standard errors based on a bootstrapping procedure are in progress.

Table 6: Counterfactual change in chain transactions after vertical and horizontal reoptimization

	(1)	(2)	(3)	(4)
	Ticket size only		Ticket size + cusine type	
City	With chain δ	Without chain δ	With chain δ	Without chain δ
Champaign	38%	-13%	52%	-1%
Charlotte	11%	-6%	9%	-7%
Cleveland	0%	-4%	0%	-5%
Las Vegas	18%	-31%	48%	-10%
Madison	27%	-27%	58%	-8%
Phoenix	11%	-20%	41%	3%
Pittsburgh	0%	-42%	0%	-40%
Average	12%	-21%	22%	-13%

The table shows the counterfactual change in transactions for large chains if each chain with more than 1000 locations could flexibly choose its quality level and cusine type in each city in the urban consumer sample. I show the change relative to its choice of optimal characteristics if it chooses only one value across all cities. In columns (1) and (2), I allow each chain to change only its ticket size x_j, while in (3) and (4), I allow each chain to reoptimize over both its ticket size and horizontal category z_j. In columns (1) and (3), the chain keeps its demand advantage $\delta_{y(i),l(j)}$, but chooses its characteristics independently in each city. In (2) and (4), each chain chooses its characteristics optimally but loses its demand advantage—consumers view the chain as an independent restaurant with one location, whose δ is normalized to zero. Transaction volume in each counterfactual scenario is compared to the predicted number of transactions that the chain would receive if it set its characteristics at the "constrained optimum"—the vertical and horizontal characteristics that it would choose if it were to choose one choice of characteristics to maximize the sum of transactions across the seven cities. To construct the figure, I perform the exercise described above for each large chain present in the urban consumer sample for which evening transactions account for the largest share of its sales (compared to sales during breakfast or lunch hours). I average across chains, weighting each chain by its predicted number of transactions.

Table 7: Change in consumer welfare from ban of large chains

	(1)	(2)	(3)	(4)	(5)	(6)
				Change in CS in miles per account		
Income group	Accounts	Transactions	Baseline CS in miles	No replacement	Replacement with independent with same characteristics	Replacement with randomly drawn independent
<50k	29,032	396,311	994,510	-7.24	-2.23	-1.75
50-100k	27,794	364,926	791,080	-6.86	-2.06	-1.46
100-150k	14,520	189,428	380,580	-6.24	-1.72	-0.68
150-200k	5,753	75,926	152,427	-6.05	-1.47	-0.08
200-250k	2,673	36,554	75,305	-5.68	-1.30	0.25
>250k	4,312	58,244	110,391	-5.39	-1.10	0.56

The table shows the effect of a hypothetical ban of large chains with more than 1000 locations in each city in the urban consumer sample by consumer income bin, computed using the logit formula for expected consumer surplus. The first and second columns shows the total number of accounts and transactions across all seven sample cities. The third column shows the baseline level of expected surplus in mile equivalents. The fourth through sixth columns show the change in consumer surplus from the baseline in miles per account under different assumptions about the supply side response. In column (4), I assume all large chain restaurants disappear and are not replaced by another restaurant. Column (5) assumes that each large chain restaurant observed in the data is replaced by an independent restaurant with the same ticket size and restaurant category as the chain it replaced. Column (6) assumes that each large chain would be replaced by a randomly selected independent, where the independent is drawn with replacement from the population of independent restaurants in that city.

Table 8: Demand parameter estimates from small town markets

Parameter	Estimate	Standard Error
τ	0.001	0.000
$\delta_{2-100,<50k}$	-0.008	0.003
$\delta_{2-100,100k}$	0.012	0.003
$\delta_{2-100,150k}$	0.049	0.004
$\delta_{2-100,200k}$	-0.044	0.008
$\delta_{2-100,250k}$	-0.154	0.014
$\delta_{2-100,>250k}$	-0.039	0.011
$\delta_{101-1000,<50k}$	0.379	0.004
$\delta_{101-1000,100k}$	0.378	0.004
$\delta_{101-1000,150k}$	0.312	0.006
$\delta_{101-1000,200k}$	0.275	0.013
$\delta_{101-1000,250k}$	0.202	0.023
$\delta_{101-1000,>250k}$	0.308	0.018
$\delta_{1001+,<50k}$	0.591	0.002
$\delta_{1001+,100k}$	0.638	0.002
$\delta_{1001+,150k}$	0.642	0.004
$\delta_{1001+,200k}$	0.583	0.007
$\delta_{1001+,250k}$	0.538	0.013
$\delta_{1001+,>250k}$	0.635	0.010

The table shows the demand parameter estimates from the estimation of equation 5 via maximum likelihood using restaurant transactions from the small town sample described in Section 2. τ is a parameter that scales the cost of eating at a restaurant with a ticket size higher or lower than a consumer's preference. $\delta_{y(i),l(j)}$ is the utility that a consumer of income group $y(i)$ receives from eating at a restaurant with $l(j)$ locations (normalized to zero for restaurants with one location for each income group).

Table 9: Entry model-based and demand-only reoptimization counterfactuals

	Demand-based counterfactual	Entry model counterfactual
Ratio constrained/unconstrained	89%	79%
Ratio competitive/unconstrained	63%	35%
Ratio competitive/constrained	72%	44%

The table compares the chain reoptimization counterfactuals conducted using only the demand estimates to those performed with the entry model described in Section 5. Each number in the table gives the ratios of total predicted transactions for the chain under three scenarios: the chain chooses one ticket size in all markets (constrained), chooses separately in each market but maintains the chain demand advantage δ (unconstrained), and operates as an independent that chooses ticket size separately in each market but loses δ (competitive). In the demand-based counterfactuals, I calculate chain profits under each of the three scenarios for all chains with more than 1000 locations operating in the set of sample towns that receive the largest share of their transactions during dinner hours (between 5pm and 11pm). I hold the ticket size and entry decisions of all other firms fixed. In the entry model, the ticket size and entry decisions of other firms are determined endogenously.

Table 10: Entry model-based and demand-only chain ban counterfactuals

	Change in consumer welfare in mile equivalents		
	Miles per transaction	Miles per consumer	Total miles (K)
Demand-based counterfactuals			
No replacement	-0.40	-10.66	-1514
Replacement with independent characteristics	-0.21	-5.56	-790
Replacement with same characteristics	-0.13	-3.51	-499
Entry model based counterfactuals	-0.09	-2.27	-322

The table compares the chain ban counterfactuals conducted using only the demand estimates to those performed with the entry model described in Section 5. Estimates of the change in expected consumer surplus using the demand-based approach are computed under the assumptions of the logit demand model described in Section 4, using the consumers and restaurants in the small town sample. To convert utility to mile equivalents in this model, I assume that the ratio of ticket size cost τ to physical distance cost γ is the same in the small town sample as in the urban consumer sample and divide the change in expected consumer surplus by $\gamma^{urban} \cdot \tau^{town}/\tau^{urban}$. The first three rows of the table show the change in consumer surplus from a chain ban under three different assumptions about the supply-side response: chain restaurants with more than 1000 locations are not replaced, they are replaced with a randomly drawn independent, or they are replaced with an independent with the same ticket size as the chain. The last row of the table shows the change in consumer surplus calculated using the entry model. To compute this, I find an equilibrium of the entry model when the chain enters all markets and an equilibrium when only independent firms enter and calculate the expected difference in consumer utility.

References

Ad Age (2018). 200 Leading National Advertisers 2018 Fact Pack. Technical report.

Adams, B. and K. R. Williams (2017). Zone pricing in regional oligopoly. Working paper, Cowles Foundation Discussion Paper.

Bresnahan, T. F. and P. C. Reiss (1990). Entry in monopoly market. *The Review of Economic Studies 57*(4), 531–553.

Bresnahan, T. F. and P. C. Reiss (1991). Entry and competition in concentrated markets. *Journal of political economy 99*(5), 977–1009.

Cai, H. and I. Obara (2009). Firm reputation and horizontal integration. *The RAND Journal of Economics 40*(2), 340–363.

Cremer, H. and J.-F. Thisse (1991). Location models of horizontal differentiation: a special case of vertical differentiation models. *The Journal of Industrial Economics*, 383–390.

Davis, D. R., J. I. Dingel, J. Monras, and E. Morales (2017, September). How segregated is urban consumption? Working Paper 23822, National Bureau of Economic Research.

DellaVigna, S. and M. Gentzkow (2017, November). Uniform pricing in us retail chains. Working Paper 23996, National Bureau of Economic Research.

Doms, M. E., R. S. Jarmin, and S. D. Klimek (2004). Information technology investment and firm performance in us retail trade. *Economics of Innovation and New Technology 13*(7), 595–613.

Einav, L., A. Finkelstein, and H. Williams (2016). Paying on the margin for medical care: Evidence from breast cancer treatments. *American Economic Journal: Economic Policy 8*(1), 52–79.

Einav, L., P. J. Klenow, B. Klopack, J. D. Levin, L. Levin, and W. Best (2017). Assessing the gains from e-commerce. Working paper, Stanford University.

Eizenberg, A., S. Lach, and M. Yiftach (2016). Retail prices in a city. Working paper.

Ellickson, P. B. (2007). Does sutton apply to supermarkets? *The RAND Journal of Economics 38*(1), 43–59.

Ellickson, P. B., S. Houghton, and C. Timmins (2013). Estimating network economies in retail chains: a revealed preference approach. *The RAND Journal of Economics 44*(2), 169–193.

Foster, L., J. Haltiwanger, and C. J. Krizan (2006). Market selection, reallocation, and restructuring in the us retail trade sector in the 1990s. *The Review of Economics and Statistics 88*(4), 748–758.

Heckman, J. J. and T. E. MaCurdy (1980). A life cycle model of female labour supply. *The Review of Economic Studies 47*(1), 47–74.

Hollenbeck, B. (2017). The economic advantages of chain organization. *The RAND Journal of Economics 48*(4), 1103–1135.

Holmes, T. J. (2001). Bar codes lead to frequent deliveries and superstores. *RAND Journal of Economics*, 708–725.

Holmes, T. J. (2011). The diffusion of wal-mart and economies of density. *Econometrica 79*(1), 253–302.

IdenTV (2018). Leading brands advertised on TV in the United States in 1st quarter 2018, by number of ad occurrences).

Jarmin, R. S., S. D. Klimek, and J. Miranda (2009). The role of retail chains: National, regional and industry results. In *Producer dynamics: New evidence from micro data*, pp. 237–262. University of Chicago Press.

Jia, P. (2008). What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica 76*(6), 1263–1316.

Loertscher, S. and G. Muehlheusser (2011). Sequential location games. *The RAND Journal of Economics 42*(4), 639–663.

Loertscher, S. and Y. Schneider (2011). Chain stores, consumer mobility, and market structure. *Journal of Institutional and Theoretical Economics JITE 167*(2), 236–246.

Luca, M. (2011). Reviews, reputation, and revenue: The case of yelp. com. harvard business school nom unit working paper 12-016, 1-40.

Mazzeo, M. J. (2002). Product choice and oligopoly market structure. *RAND Journal of Economics*, 221–242.

Mazzeo, M. J. (2004). Retail contracting and organizational form: Alternatives to chain affiliation in the motel industry. *Journal of Economics & Management Strategy 13*(4), 599–615.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic theory 18*(2), 301–317.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics 37*(3), 619–640.

Shaked, A. and J. Sutton (1987). Product differentiation and industrial structure. *The Journal of Industrial Economics*, 131–146.

Sutton, J. (1991). *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration.* MIT press.

Tsai, Y.-L., C. S. Dev, and P. Chintagunta (2015). What's in a brand name? assessing the impact of rebranding in the hospitality industry. *Journal of Marketing Research 52*(6), 865–878.