Online appendix: One size fits all? The value of standardized chains

Appendix A Data

A.1 Payment card data overview

My primary source of data is the universe of 2016 transactions on a major payments card network. The payments card provider is among the largest in the US. Total transaction volume on the network in 2016 was approximately 20% of all US consumption.

An observation in the underlying data is a transaction between a card and a merchant. The key variables used from this dataset are a unique card identifier, date and time of transaction, merchant identifier (defined at the brand level), store identifier (defined at the outlet level), latitude and longitude of the store location, and the dollar amount of the transaction. For 55% of active 2016 credit cards issued by the payment cards network, the company has access to a measure of estimated household income and the cardholder's billing zipcode. Household income is estimated by a third party from information available in a credit report. I use all credit card transactions for which the payment card company observes estimated household income. I further exclude transactions that do not meet the following criteria: the transaction must be a sales draft, occur at a US merchant, and take place on the payment card provider's credit network. This excludes prepaid cards and debit cards, as I do not observe cardholder income transactions for these cards. Together, these filters eliminate less than 10% of credit spending in 2016. To describe the aggregate importance of chains and the variation across category and income group (Section 3), I use the sample described above. In the demand estimation and entry model analysis, I further restrict this sample to specific geographies, described below.

A.2 Card location

In my empirical analysis, a key determinant of consumer choice is the distance the consumer must travel to visit a particular restaurant. My primary measure of consumer location is constructed from the zipcodes in which the card transacts. I define a card's "shopping location" as follows:

$$l_i = \frac{\sum l_z \cdot N_{iz}}{\sum N_{iz}}$$

where l_i is the latitude, longitude of the shopping location for card i, l_z is the latitude, longitude of zipcode z, and N_{iz} is the number of transactions card i made at non-restaurant merchants in z conditional on $N_{iz} \geq 20$. In other words, I place a card at their transaction-weighted centroid of the zipcodes in which it transacts at least 20 times. As a result, this measure is only available for cards that transacted relatively frequently. I check the robustness of the "shopping location" described above by comparing it to an external measure of card location - the billing zipcode of the card at the zip+4 level. The two measures are highly correlated. I use the shopping location in my baseline estimation.

A.3 Yelp data

Yelp publishes a sample of data for academic purposes.[27] The data contain reviews and business characteristics for 11 cities, 7 of which are in the US. I match restaurants in the payment card data to this sample. I require a restaurant to match on the following fields between the payment card and Yelp datasets: the restaurant names (stripped of numbers, punctuation, and spaces), restaurant address number (excluding the street name) OR latitude and longitude within 0.25 miles.

I show summary statistics on the result of the merging process in Table A2. Overall, as a share of all businesses in Yelp, I am able to match 47% of Yelp restaurants across the seven cities to an entity in the payment card data, containing 43% of the total Yelp reviews. Out of all restaurants that appeared in the payment card data, about 31% were matched to Yelp, accounting for 43% of total transactions in the category. There are several potential reasons that businesses may not be matched, including discrepancies in the recorded business name between the two data sources, unpopulated or incorrect address information (e.g. food trucks), the presence of Yelp businesses that do not accept credit cards, businesses that

[27]The current version can be freely downloaded here: https://www.yelp.com/dataset/challenge

opened during 2017 (which would not register 2016 transactions in the Visa data but would appear in the 2017 Yelp snapshot), or businesses that closed in 2016 (and thus would have Visa transactions but would be marked as closed by Yelp in 2017). To reduce the presence of false positives that could bias my demand results, I require an almost exact match on the business name and location for inclusion in the analysis sample, which is likely to lower the absolute number of businesses I am able to match.

I include all matched restaurants where the listed city in the Yelp dataset is Champaign, IL; Charlotte, NC; Pittsburgh, PA; Cleveland, OH; and Madison, WI. In the two largest cities, Las Vegas and Phoenix, the number of available restaurants make my demand estimation exercise infeasible using all matched restaurants. Las Vegas has 1,615 matched restaurants and Phoenix has 1,115 matched restaurants. For these two cities, I limit my analysis to restaurants close to the downtown area. In Las Vegas, I restrict the sample to restaurants in the Downtown, Spring Valley, and Southeast neighborhoods, according to a map of Las Vegas neighborhoods, which I reproduce in Figure C.1. These neighborhoods correspond to the following zipcodes: 89101, 89102, 89103, 89104, 89106, 89107, 89109, 89119, 89123, 89146, 89147, and 89169. In Phoenix, I limit restaurants to those available within a 10 mile radius of the downtown area, as defined by Google Maps.

A.4 Ticket size

The payments card data does not contain information on prices of the items purchased. One element of restaurant differentiation is along vertical attributes - some restaurants charge higher prices and have higher quality food and service. In my data, I measure this with average transaction size, or ticket size. I calculate ticket size x_j for each merchant j as the sum of all dollars spent at merchant j divided by the number of swipes at merchant j, summed across all outlets belonging to j.

This measure is highly correlated with Yelp's measure of price, defined on a dollar sign scale that goes from one to four dollar signs. The Yelp measure is based on a survey of Yelp users that review a restaurant. The survey asks about the approximate price for a meal for one person. The translation of the dollar sign measure to prices is as follows: one \$ implies a cost under \$10, two \$\$ between \$11 and \$30, three \$\$\$ between \$31 and \$60, and four $$$$ above $60. I show the distribution of log(ticket size) by Yelp dollar sign rating in Figure C.2. The four distributions are monotonic and largely non-overlapping, implying that almost no one dollar sign restaurant has a higher average ticket size than a two dollar sign restaurant, and the same for two and three dollar sign restaurants. There are few four dollar sign restaurants in the data, and thus the distribution of their average ticket size tends to be noisier.

A.5 Locations

I define the number of locations for each merchant using data from the payments card company. The company records a unique merchant and store identifier associated with each transaction. For each merchant, I define the number of nationwide locations as the number of distinct store identifiers that had at least 100 transactions in 2016.

A.6 Urban consumer sample

In Section 4, I study the seven markets included in the Yelp data. In addition to the filters described above, I keep only transactions that occurred at Yelp-matched restaurants by cards that had a shopping location within 25 miles of the city and that had at least 5 transactions at Yelp-matched restaurants in that city in 2016. Of these remaining cards, I take a 50% random sample.

A.7 Restaurant categories

I use the category "tags" from the Yelp data to assign each restaurant to a one of eight categories: Latin-American, European, Pizza, Sandwiches, Asian, Burgers, American, and Other. Restaurants in Yelp are assigned detailed category tags that describe the type of food they sell. Tags can be assigned by restaurant owners or Yelp users, and are sometimes populated algorithmically from review content by Yelp. Restaurants are frequently assigned multiple tags. I show a screenshot from Yelp highlighting these tags for one restaurant in Figure C.3.

I manually map each tag into one of these eight categories. Within a merchant, I assign all outlets to the same category. I find the modal category for each merchant and assign all outlets belonging to that merchant with that category. In practice, there are few cases in which outlets from the same merchant would be assigned to different categories. I force this standardization to correct for these rare errors in tag assignment. In Table A1, I briefly describe the types of restaurants in each category and show some of the most common Yelp tags for each category.

Appendix B Estimation

I face the computational challenge that estimation of equation 1 with individual-specific quality blisspoints α_i with fixed effects requires the estimation of thousands of parameters. I estimate the α_i parameters as fixed effects to allow for correlation between quality preferences and other observables, such as cardholder income. I implement the estimation procedure described in Heckman and MaCurdy (1980) (see Appendix A in that paper).[28]

I proceed as follows. I first maximize the log likelihood arising from equation 1 without the individual-specific ticket size preferences, substituting the income-group mean $\alpha_{y(i)}$ for the card-specific α_i. Then, conditional on the parameters δ, θ and γ, I solve for the likelihood-maximizing $\hat{\alpha}_i$ for each card i, exploiting the fact that the α_i are independent across i and can be found in isolation for each i. I substitute these $\hat{\alpha}_i$ into the likelihood and repeat the estimation procedure until the parameters converge.[29] For computational convenience, I discretize α_i and allow it to take values between \$5 and \$100 in \$5 increments for a total of 20 grid points.

Appendix C Model

The model in Section 5 is based on Loertscher and Muehlheusser (2011). In their setup, a large number of potential entrants choose a one-dimensional product characteristic between

[28] Similar procedures have also been used in Polachek and Yoon (1996). See Greene (2002) for additional discussion.

[29] In practice, convergence happens relatively quickly. I find that there is no appreciable change in the parameter values after four iterations.

0 and 1. Firms move sequentially, deciding whether or not to enter and making product characteristic choices upon entering. I make several modifications to their original setup. First, I assume that there is one chain firm that moves first and faces higher demand than subsequent independent entrants. Second, I assume that the firms compete in many markets and the chain firm is constrained to choose the same ticket size in every market, while each entrant can choose their ticket size without constraint. Third, I assume that the distribution of demand in each market is distributed lognormally. In this section, I summarize key results from Loertscher and Muehlheusser (2011) (hereafter LM) and extend them to my context where necessary.

C.1 Model setup

Consumers in each market, indexed by m, choose between a set of restaurants available in their market. Restaurants are indexed by j and are differentiated by their one-dimensional vertical characteristic x_j. If consumer i with preference parameter α_i purchases from a chain firm with vertical characteristic x_j, she get utility:

$$u^c_{ijt} = \delta - \tau(\alpha_i - x_j)^2$$

If she purchases from an independent firm k, she gets:

$$u_{ikt} = -\tau(\alpha_i - x_j)^2$$

Consumer preferences α_i are distributed in each market according to the lognormal distribution f_m, with $\log(f_m) \sim N(\mu_m, \sigma^2_m)$.

On the firm side, one chain competes with a fringe of potential independent entrants who make sequential entry and product characteristic decisions. The chain moves first and chooses its product characteristic x_j to maximize the weighted sum of profits across markets, where each market is weighted by its total number of restaurant transactions. Each independent entrant then considers entry into the market. If it enters, it pays a sunk cost K_m, which varies across markets, and chooses a location in characteristic space x_k. Firms continue to enter until further entry is no longer profitable. Once the entry process is complete, each firm j gets profits $\Pi_j^m = Q_j^m(x_j^m, x_{-j}^m) - K_m$, where $Q_j^m(x_j^m, x_{-j}^m)$ is equal to demand for firm m. All cost in the model is included in the fixed cost parameter K_m.

C.2 Firm demand

Consumers maximize utility by choosing a restaurant $j^* = \arg\max_{j \in J_m} u_{ijt}$, where J_m is the set of restaurants that enter in market m. The utility-maximizing choice for each consumer is either the independent firm closest to the consumer's preference in ticket size space or the chain. The consumer picks the chain j over the independent k closest to her preference parameter α_i if $\delta - \tau(\alpha_i - x_j)^2 \geq -\tau(\alpha_i - x_k)^2$.

The demand functions for each firm can be simply characterized as a function of its own product characteristic choice and the choices of the firms on its left and right. Consider first the demand of the chain firm j, with its closest neighbors in product characteristic space l and r, with $x_l < x_j < x_r$.[30] Between firm j and its left neighbor l, there is an indifferent consumer with preference $\tilde{\alpha}_l$, defined by $\delta - \tau(x_j - \tilde{\alpha}_l)^2 = -\tau(x_l - \tilde{\alpha}_l)^2$. Expanding this expression and simplifying yields $\tilde{\alpha}_l = \frac{x_l + x_j}{2} - \frac{\delta}{2\tau(x_j - x_l)}$. Similarly denote the indifferent consumer between j and r as $\tilde{\alpha}_r$. Then firm j attracts all consumers with $\alpha_i \in [\tilde{\alpha}_l, \tilde{\alpha}_r]$, and thus has demand $Q_j = F_m(\tilde{\alpha}_r) - F_m(\tilde{\alpha}_l)$, where F_m is the CDF of the distribution of consumer preferences α_i.

Now consider the demand for an independent firm k with neighboring independent firms located at x_l and x_r on its left and right. The indifferent consumer between l and k is located at $(x_l + x_k)/2$, and k faces demand of $Q_k = F((x_r + x_k)/2) - F((x_l + x_k)/2)$. If instead k's neighboring firm is the chain, say the firm to its left l, the indifferent consumer is given by $\tilde{\alpha}_l = \frac{x_l + x_k}{2} + \frac{\delta}{2\tau(x_k - x_l)}$, and the firm faces demand $Q_k = F((x_r + x_k)/2) - F(\tilde{\alpha}_l)$. I summarize the demand functions for chain and independent firms in Table A3.

[30] I restrict attention to neighboring locations x_l and x_r such that $\delta - \tau(x_j - x_l)^2 < 0$ and $\delta - \tau(x_j - x_r)^2 < 0$. This is without loss of generality; When $\delta > 0$ and $K_m > 0$, there is a set of locations around the chain firm in which independent firms earn strictly negative profits, since even a consumer with $\alpha_i = x_l$ prefers chain firm j when $\delta - \tau(x_j - x_l)^2 \geq 0$.

C.2.1 Equilibrium with no chain

A key insight from LM is 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). The intuition for this result is that, in equilibrium in their model with no chain, independent firms choose locations to maximally deter entry of future firms in the interval around them.

I first characterize the equilibrium in the LM model with one market and no chains when f is lognormal and then discuss the strategy of the chain. I introduce two definitions from LM. For a location y occupied by an independent firm, let $\lambda(y)$ and $\rho(y)$ be such that $\max_{x_j \in [y,\lambda(y)]} Q_j(x_j) = K$ and $\max_{x_j \in [\rho(y),y]} Q_j(x_j) = K$ for another firm j that chooses location x_j and gets revenue $Q_j(x_j)$, respectively. In other words, given that a firm has located at y, $\lambda(y)$ is the maximum distance from y (in ticket size space) that the nearest independent firm can occupy while still deterring entry in the interval between y and $\lambda(y)$. $\rho(y)$ is analogously defined as the maximum entry-deterring distance on the left. Define the entry-deterring boundary locations on the left and right respectively as $\lambda_b \equiv F^{-1}(K)$ and $\rho_b \equiv F^{-1}(1-K)$. These locations give the largest and smallest locations in product characteristic space that independent firms can occupy while still deterring entry above or below, respectively.

LM show two important results that make characterizing the best-response functions of independent firms particularly tractable.

Lemma C.1 *1. If f is quasiconcave and $x_0 < \rho(x_1) \leq \lambda(x_0) < x_1$, then $N(x_0, x_1) = 1$.*

2. *Given neighboring equilibrium locations at x_0 and $x_1(x_i)$, i's optimal location is $x_i^* = \lambda(x_0)$ if f is increasing over $[x_0, x_1(x_i)]$ and $\frac{dx_1(x_i)}{dx_i} \geq 0$. Given neighboring locations $x_0(x_i)$ and x_1, i's optimal location is $x_i^* = \rho(x_1)$ if f is decreasing over $[x_0(x_i), x_1]$ and $\frac{dx_0(x_i)}{dx_i} \geq 0$.*

3. *Given neighboring locations at x_0 and x_1, i's optimal location is $x_i^* = \lambda(x_0)$ if f is symmetric and hump shaped and $\lambda(x_0) < Mo(f)$, where $Mo(f)$ is the mode of the distribution f. Analogously, given neighboring locations at x_0 and x_1, i's optimal location is $x_i^* = \rho(x_1)$ if f is symmetric and hump shaped and $\rho(x_1) \geq Mo(f)$.*

In other words, in any monotonic region of f, firm i's optimal location will be the furthest entry-deterring location from the firm on its left (right) when f is increasing (decreasing). I refer the interested reader to Loertscher and Muehlheusser (2011) for the proof of the above lemma (proof of Lemma 3, p. 21).

I show that a version of part 3 of this result also holds for lognormal f, which is hump shaped, but not symmetric, with an additional condition.

Lemma C.2 *Given neighboring locations at x_0 and x_1, i's optimal location is $x_i^* = \lambda(x_0)$ if f is lognormal, $\lambda(x_0) < Mo(f)$ and $f(\frac{x_i+x_1}{2}) \geq f(\frac{x_i+x_0}{2})$ for $x_i \in [x_0, \lambda(x_0)]$, where $Mo(f)$ is the mode of the distribution f. Analogously, given neighboring locations at x_0 and x_1, i's optimal location is $x_i^* = \rho(x_1)$ if f is lognormal, $\rho(x_1) \geq Mo(f)$, and $f(\frac{x_i+x_1}{2}) \geq f(\frac{x_i+x_0}{2})$ for $x_i \in [\rho(x_0), x_1]$.*

Proof (modified from LM): I consider only the first claim, since the proof for the second is nearly identical. For x_1 to be the equilibrium neighboring location to the right of x_i, it must be the case that $\rho(x_1) \leq x_i$, since otherwise an additional firm would enter between x_i and x_1. For x_0 to be the neighboring equilibrium location on the left, it must be that $x_i \leq \lambda(x_0)$. Firm i's profit is given by $\Pi_i = F(\frac{x_i+x_1)}{2}) - F(\frac{x_i+x_0}{2}) - K$. Taking the derivative with respect to x_i yields $\frac{\partial Pi_i}{\partial x_i} = \frac{1}{2}(f(\frac{x_i+x1}{2}) - f(\frac{x_i+x_0}{2}))$. By the assumed condition $f(\frac{x_i+x_1}{2}) \geq f(\frac{x_i+x_0}{2})$, $\frac{\partial Pi_i}{\partial x_i} \geq 0$ for $x_i \in [x_0, \lambda(x_0)]$. Thus, in the interval $[x_0, \lambda(x_0)]$, i's optimal location is $\lambda(x_0)$. By Lemma C.1 part (i), choosing $x_i > \lambda(x_0)$ is not consistent with equilibrium, since it induces another firm to enter to $\lambda(x_0)$ and earn weakly positive profits.

The additional condition $f(\frac{x_i+x_1}{2}) \geq f(\frac{x_i+x_0}{2})$ for $x_i \in [x_0, \lambda(x_0)]$ requires that x_1 be "close enough" to the mode of the lognormal distribution. In calibrating the model, I check numerically that this condition holds.

With this result, I now reproduce a key theorem from LM that characterizes the equilibrium locations of independent firms. Define $\lambda^1 \equiv \lambda(\lambda_B)$, $\lambda^2 \equiv \lambda(\lambda^1)$, $\ldots$ $\lambda^{j+1} = \lambda(\lambda^j)$ and $\rho^1 \equiv \rho(\rho_B)$, $\rho^2 \equiv \rho(\rho^1)$, $\ldots$ $\rho^{j+1} = \rho(\rho^j)$. Now define two sets: $\lambda(s) \equiv \{\lambda_B, \lambda^1, \lambda^2, \ldots \lambda^s\}$ and $\rho(r) \equiv \{\rho^r, \ldots \rho^2, \rho^1\}$.

Theorem C.3 *1. If f is monotone over $[a, b]$, the set of equilibrium locations is unique. If f is increasing, it is given by $\lambda(n) \cup \rho_B$. If f is decreasing, it is given by $\lambda_B \cup \rho(m)$.*

2. *If f is hump shaped and K is such that no entry occurs at the mode of the distribution, then the unique set of equilibrium locations is given by $\lambda(s) \cup \rho(s)$.*

C.2.2 Best response of independents to chain entry

Given the previous result, the equilibrium locations of independent firms in ticket size space can be characterized simply. Suppose first that the chain enters at location $x_c \leq Mo(f)$. By Theorem C.3, since f is increasing on $[0, x_c]$, the unique set of equilibrium locations in this interval is given by $\lambda_b, \lambda^1, \ldots \lambda^s \cup \rho(x_c)$. Stated differently, the first s locations to the left of the chain are identical to the no-chain equilibrium. By the same theorem, the locations to the right of the mode are the same as in the no-chain equilibrium, since this region of f is also monotonic (but decreasing). These locations are given by the set $\rho(r)$.

Finally, I look for additional locations right of x_c but left of $Mo(f)$. Since this region of f is monotonically increasing, Lemma C.1 can be applied. The first location to the right of x_c, λ_{s+1} is given by $\lambda(x_c)$ (when $\lambda(x_c) < Mo(f)$ - otherwise no such locations exist). Additional left-of-the-mode locations can be found be found by by applying the λ function repeatedly: $\lambda^{s+2} = \lambda(\lambda^{s+1})$, $\lambda^{s+3} = \lambda(\lambda^{s+2})$, $\ldots \lambda^{s+t} = \lambda(\lambda^{s+t-1})$. By Lemma C.2, the right-most location that is still left of the mode is pinned down by the location to its left.

An analogous result holds when the chain enters at $x_c > Mo(f)$. All locations left of the mode are identical to the no-chain equilibrium, and the firms to the right of x_c are given by $\lambda(x_c), \rho(t), \rho(t-1) \ldots \rho(B)$. as long as firms at those locations earn non-negative profits. Locations right of the mode but left of the chain are spaced at ρ distance from x_c.

C.2.3 Strategy of the chain

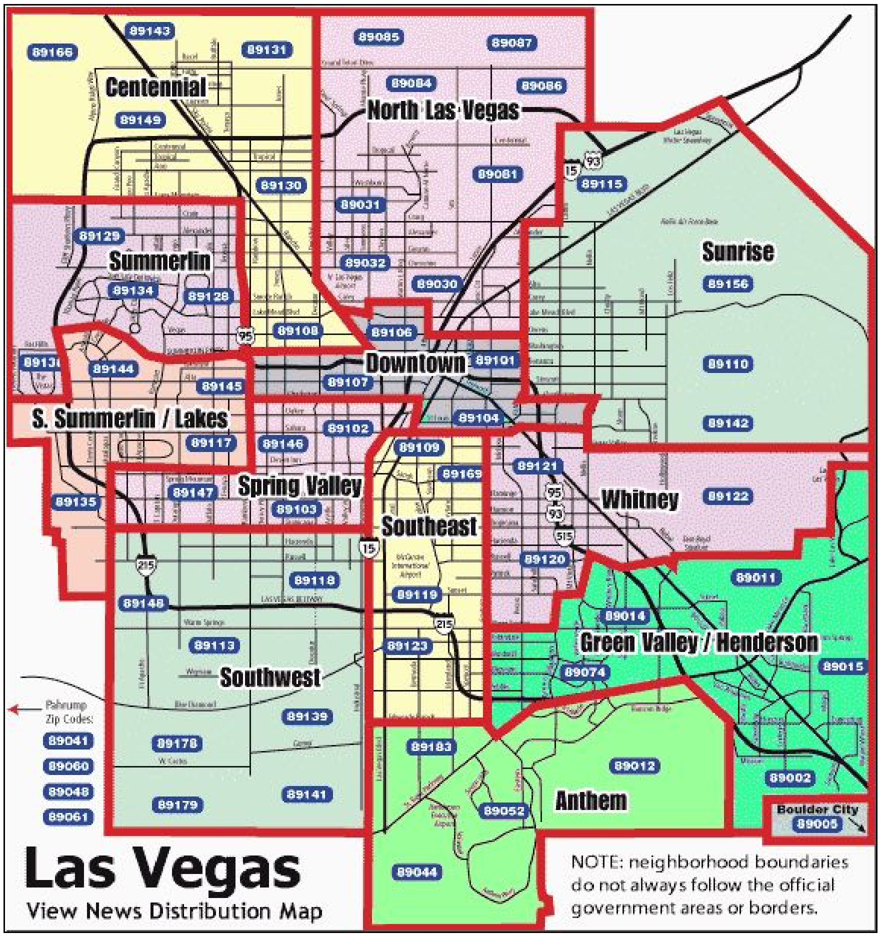
The chain makes the first entry decision in the game, correctly anticipating the equilibrium ticket size decisions of the independent firms that move later. Rather than characterizing the chain's strategy analytically, I perform a grid search over a set of chain locations. When the chain competes in only one market, it earns the highest profits by locating near the mode of f. Intuitively, its demand advantage δ is most valuable when 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. I plot the equilibrium locations and profits for the set of independent firms and the chain for a particular distribution f in Figure C.6.

When the chain enters in many markets at the same ticket size, it optimizes the sum of its profits across all markets. Recall that δ is the additional utility a consumer receives from eating at a chain. When δ is positive, the chain maximizes its profits by choosing a point where demand is thickest in the average across markets.

If the distribution of consumers is identical across markets, the constraint that the chain must pick the same ticket size does not decrease the chain's profits. If markets are very heterogeneous, then this constraint may significantly disadvantage chains relative to independent firms. To illustrate this point using the model, I simulate two sets of markets - a "low heterogeneity" case in which consumers in each markets have relatively similar taste distributions and a "high heterogeneity" case where consumer tastes vary more.[31] Figures C.5 and C.4 show the "boundary" distributions for the two sets of markets, with all simulated markets lying in between the blue and red lines. The solid black line in the figure shows the profits of the chain as a function of its chosen ticket size. In both cases, the chain maximizes its profits by choosing a point between the modes of the boundary distributions. In the low heterogeneity case, however, the chain earns more than twice the profits than in the high heterogeneity case.

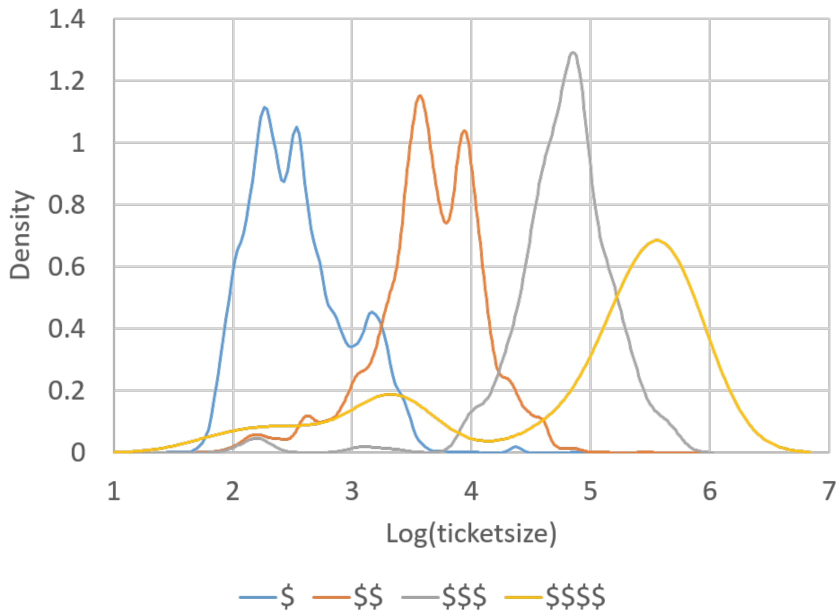
[31] I assume that the blisspoints of consumers x_i in each market are distributed lognormally with $log(x_i) \sim N(\mu, \sigma^2)$. μ is linearly spaced between 3 and 4 in the high heterogeneity case, and between 3.4 and 3.6 in the low heterogeneity case. σ^2 is fixed across markets at X.

Figure C.1: Las Vegas Neighborhood Zipcode Map



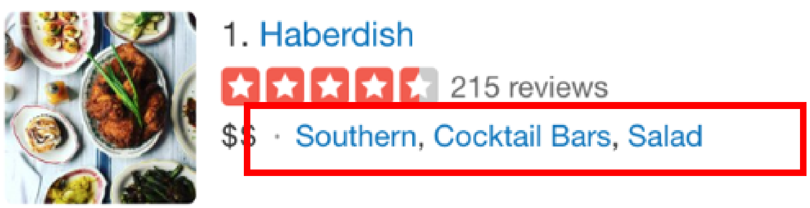
The figure shows a map of Las Vegas by neighborhoods, with their accompanying zipcodes. In my empirical analysis of Las Vegas, I restrict attention to the downtown area, including the Downtown, Spring Valley, and Southeastern neighborhoods on the above map. This includes the official Downtown area, as well as the Strip (primarily in Southeast) and adjacent neighborhoods in Spring Valley.

Figure C.2: Distribution of log(ticket size) for restaurants by Yelp dollar sign rating



The figure shows the distribution of log(ticket size) for restaurants included in my urban consumer sample by their dollar sign rating on Yelp. I calculate ticket size at the merchant level (so that all outlets within a chain have the same ticket size).

Figure C.3: Sample of Yelp review with category tags



The figure shows a screenshot for the Yelp listing of a restaurant in Charlotte, NC. The red box shows the three category tags for this restaurant. I map these category tags to eight restaurant categories, described above.

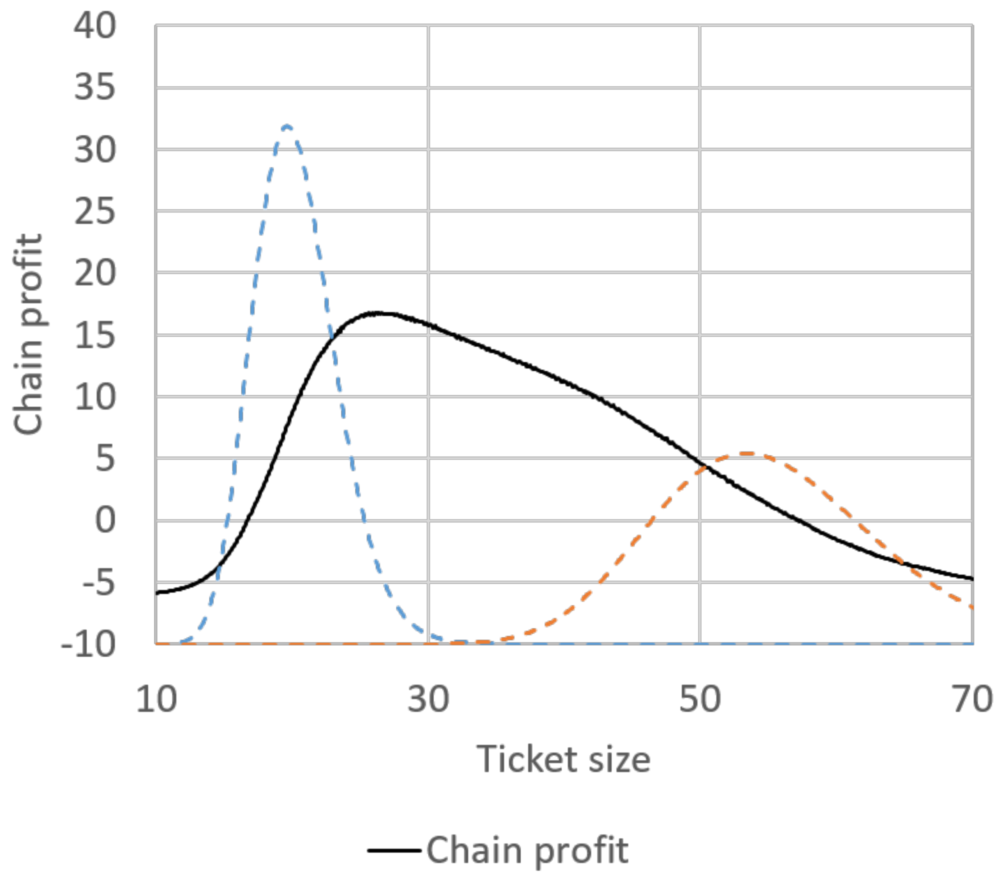


Figure C.4: Optimal chain location in high heterogeneity case

The figure shows simulations of the model described in Section 5 and C. I simulate 100 markets with different demand distributions. The 100 simulated markets have distributions of demand that lie in between the red and blue dotted lines. I solve the model by calculating the sum of the profits of the chain across all 100 markets for a grid of possible chain locations. The black line shows the sum of chain profits for each location on this grid.

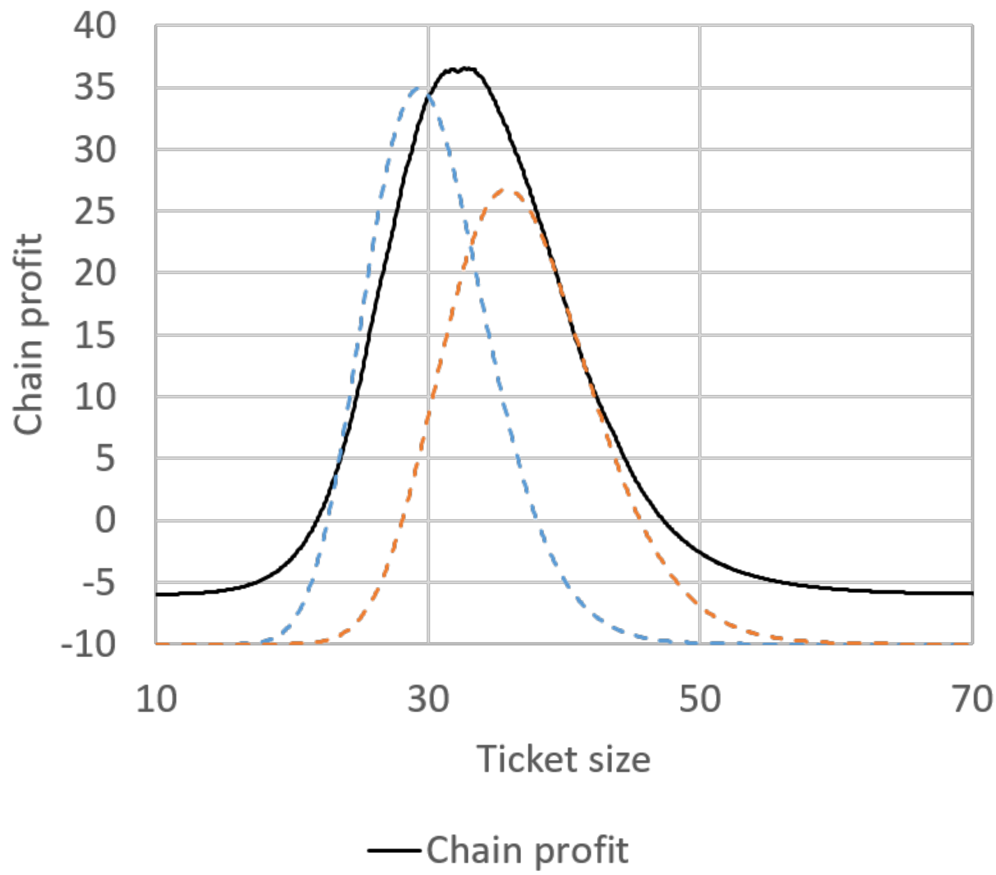


Figure C.5: Optimal chain location in low heterogeneity case

The figure shows simulations of the model described in Section 5 and C. I simulate 100 markets with different demand distributions. The 100 simulated markets have distributions of demand that lie in between the red and blue dotted lines. I solve the model by calculating the sum of the profits of the chain across all 100 markets for a grid of possible chain locations. The black line shows the sum of chain profits for each location on this grid.

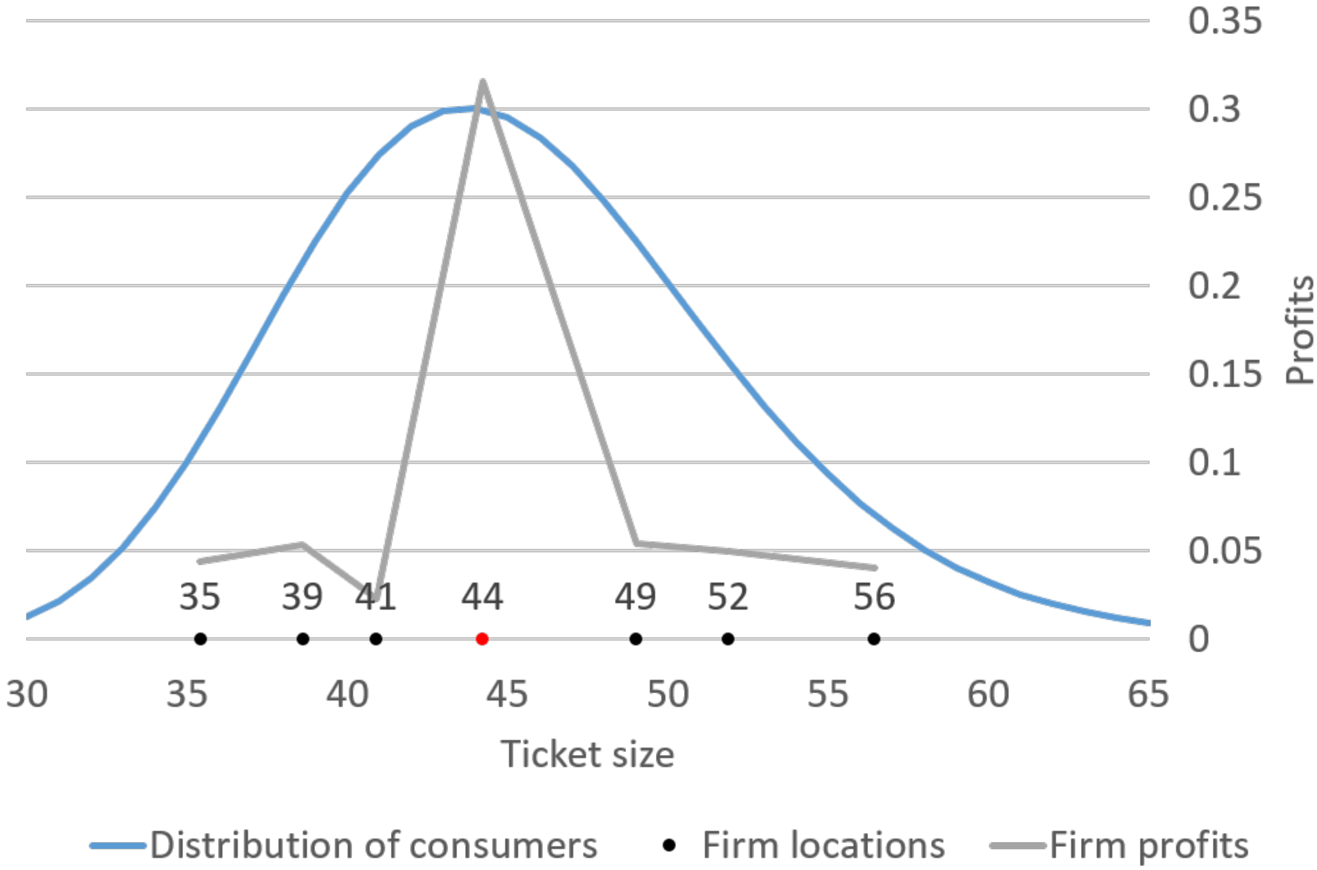


Figure C.6: Solution in one simulated market

The figure shows an equilibrium to the model described in Sections 5 and C in one market. The blue line gives the distribution of consumer demand, the black dots show the equilibrium locations of independent firms, and the red dot shows the equilibrium location of the chain firm. The grey line shows the profits of each firm.

Table A1: Description and Popular Tags of Cuisine Types

Category	Description	Popular Tags
Latin American	Restaurants specializing in cuisines from South and Central America and the Caribbean.	Mexican, Tex-Mex, Latin American, Seafood, Caribbean, Cuban
American	Restaurants serving American cuisine, but excluding restaurants specializing in Burgers and Sandwiches, and excluding restaurants that were also tagged as another type.	American (Traditional), American (New), Breakfast & Brunch, Chicken Wings, Diners
Asian	Restaurants specializing in cuisines from South Asian, East Asian, and Southeast Asian countries, as well as Pacific Islands.	Chinese, Japanese, Sushi Bars, Asian Fusion, Thai, Indian, Hawaiian
Burgers	Restaurants with tag "Burgers".	Burgers, Hot Dogs, Sports Bars, Steakhouses
European	Restaurants specializing in Italian, French, or other European cuisines, except for restaurants also tagged "Pizza".	Italian, French, Irish, Wine Bars, Noodles, Mediterranean
Pizza	Restaurants with tag "Pizza".	Pizza, Italian, Salad
Sandwiches	Restaurants with tag "Sandwiches", "Deli", or "Cheesesteaks".	Sandwiches, Deli, Cheesesteaks
Other	Restaurants not tagged as any of the above categories.	Cafe, Bar, Breakfast & Brunch

Table A2: Share of Yelp restaurants matched to credit card data

	Businesses			Reviews		
City	Matched	All	% Matched	Matched	All	% Matched
Champaign	207	387	53%	8580	16643	52%
Charlotte	1299	2775	47%	67017	142842	47%
Cleveland	1662	3452	48%	56552	120700	47%
Las Vegas	2031	4720	43%	299323	758181	39%
Madison	448	968	46%	24371	52403	47%
Phoenix	3356	6953	48%	287409	634523	45%
Pittsburgh	1170	2509	47%	51717	110910	47%
Total	10173	21764	47%	794969	1836202	43%

The table shows summary statistics for the share of Yelp businesses and reviews by city that I am able to match to an entity in the Visa data in 2016. I match businesses in Yelp and in the credit card data if their exact names (stripped of spaces, special characters, and numbers), addresses, and zipcodes are identical in the two datasets. In the table above, I report all entries in the Yelp data that were categorized as restaurants. The Yelp data also include some restaurants outside of city boundaries in nearby suburbs that I exclude in my final analysis sample. The statistics above show the outcome of the match process before this filtering occurs.

Table A3: Demand functions for chains and independent firms

Left neighbor \ Right Neighbor	Independent	Chain	None
Independent	$F(\frac{x_r+x_j}{2}) - F(\frac{x_l+x_j}{2})$	$F(\frac{x_r+x_j}{2} - \frac{\delta}{2t(x_j-x_r)}) - F(\frac{x_l+x_j}{2})$	$1 - F(\frac{x_j+x_l}{2})$
Chain	$F(\frac{x_r+x_j}{2}) - F(\frac{x_l+x_j}{2} + \frac{\delta}{2t(x_j-x_l)})$	N/A	$1 - F(\frac{x_l+x_j}{2} + \frac{\delta}{2t(x_j-x_l)})$
None	$F(\frac{x_j+x_r}{2})$	$F(\frac{x_r+x_j}{2} - \frac{\delta}{2t(x_j-x_r)})$	1

(a) Demand for independent firm j

Left neighbor \ Right Neighbor	Independent	None
Independent	$F(\frac{x_r+x_j}{2} + \frac{\delta}{2t(x_j-x_r)}) - F(\frac{x_l+x_j}{2} - \frac{\delta}{2t(x_j-x_l)})$	$1 - F(\frac{x_l+x_j}{2} - \frac{\delta}{2t(x_j-x_l)})$
None	$F(\frac{x_r+x_j}{2} + \frac{\delta}{2t(x_j-x_r)})$	1

(b) Demand for chain firm j

The table shows the demand functions in the entry model described in Appendix C for firm j as a function of the locations of its neighboring firms on the left and right. Panel (a) shows the demand functions for independent firms and panel (b) shows the demand function for the chain. In each panel, the label on the rows gives the closest neighbor to the left and the label on the column gives the closest neighbor on the right.