Journal of Economic Geography (2014) pp. 1–39 doi:10.1093/jeg/lbu040

Cities and product variety: evidence from restaurants

Nathan Schiff

Sauder School of Business, University of British Columbia and School of Economics, Shanghai University of Finance and Economics
Corresponding author: *email* <nschiff@gmail.com>

Abstract

This article measures restaurant variety in US cities and argues that city structure directly increases product variety by spatially aggregating demand. I discuss a model of entry thresholds in which market size is a function of both population and geographic space and evaluate implications of this model with a new data set of 127,000 restaurants across 726 cities. I find that geographic concentration of a population leads to a greater number of cuisines and the likelihood of having a specific cuisine is increasing in population and population density, with the rarest cuisines found only in the biggest, densest cities. Further, there is a strong hierarchical pattern to the distribution of variety across cities in which the specific cuisines available can be predicted by the total count. These findings parallel empirical work on Central Place Theory and provide evidence that demand aggregation has a significant impact on consumer product variety.

Keywords: Product variety, Central Place Theory, spatial models, consumer cities
JEL classifications: R12, L10
Date submitted: 4 June 2013 **Date accepted:** 9 September 2014

1. Introduction

Easy access to an impressive variety of goods may be one of the most attractive features of urban living. A 'love of variety' is a central feature of many canonical economics models (e.g. Dixit and Stiglitz, 1977) and more recent work in urban economics suggests that consumption amenities, especially local or nontradable goods, are a significant force driving people to live in cities (Glaeser et al., 2001; Chen and Rosenthal, 2008; Lee, 2010). The amount of product variety in a market may also provide insight into the intensity of competition as firms in differentiated markets may face less direct price competition (Mazzeo, 2002). Despite the potential importance, we have little evidence on consumer product variety in cities[1] and therefore cannot explain why some cities seem to have more variety than others. Cities vary widely in the characteristics of their residents, government, and culture and it could be that these idiosyncratic features explain all differences in consumer product variety. On the other hand, the Central Place Theory of cities suggests a hierarchical pattern in the number of goods available in different size markets with larger markets having all of the goods of smaller markets

1 A recent paper by Handbury and Weinstein (2012) uses grocery store data to show that residents of cities have greater access to *tradable* varieties.

as well as additional higher order goods (Christaller and Baskin, 1966; Lösch, 1967). While this theory has not been applied to consumer product variety, recent empirical work has shown that industrial composition varies systematically with population size (Mori et al., 2008; Mori and Smith, 2011; Hsu, 2012). In this article I will show that a key feature of the Central Place Theory of cities, demand aggregation, suggests that even if all cities were identical in their characteristics, differences in population size and land area could lead to significant differences in consumer product variety. I will argue that cities aggregate demand on two margins—population size (scale) and land area (transportation cost)—and thus by concentrating groups of consumers with the same preferences in a small geographic space, large, dense cities provide the necessary demand for a firm catering to that taste.

I then measure product variety across a large set of cities for an important nontradable consumption amenity—restaurants—and examine how demand aggregation affects variety across locations. Using a unique data set of over 127,000 restaurants across 726 US cities I find that population size and population density have a substantial effect on the amount of product variety in a city. I estimate the elasticity of restaurant variety with respect to population as between 0.35 and 0.49. The elasticity with respect to population density, independent of population size, is between 0.16 and 0.21, suggesting that geographically concentrating a population also increases restaurant variety. However, I only find a significant effect of density alone for cities with large land areas, defined as those in the top quartile by land area in my data (182 cities, mean population 331,000). The way in which variety increases with population and land area is consistent with a simple model of demand aggregation and many of the characteristics of the distribution of cuisines across cities parallel findings from the empirical work on Central Place Theory. The specific cuisines found in each city follow a hierarchical structure in which cities with relatively rare cuisines often have all of the more common cuisines and cities with few cuisines tend to have only common cuisines. Rare cuisines are only found in cities with many restaurants while common cuisines can be found in cities with few restaurants. These findings cannot be explained by the empirical distribution of restaurants across cuisines ('balls and bins' models) and suggest that the demand aggregation mechanism of Central Place Theory can have a significant effect on a city's consumer product variety.

1.1. Consumer cities and nontradable product variety

In their 'Consumer City' paper, Glaeser et al. (2001) write that there are 'four particularly critical amenities' leading to the attractiveness of cities and that 'first, and most obviously, is the presence of a rich variety of services and consumer goods'. Noting that most manufactured goods can be ordered and are thus available in all locations, they write that it is the local nontradable goods that define the consumer goods of a city. I will continue with this notion of local nontradable consumer goods and suggest that it is especially for products characterized by significant consumer transportation costs, heterogeneous tastes, and a fixed cost of production, that the ability of cities to agglomerate people with niche tastes will lead to greater variety. Examples of this type of product would include bars, concert halls, hair salons, movie theaters, museums, restaurants and any other location-based service or good that is differentiated and patronized by consumers with a specific set of preferences. This idea would also hold for retailers that aggregate specific collections of tradable goods and

where visiting the store itself provides some substantial benefit to the consumer, such as specialty bookstores, niche toy stores, or clothing boutiques.

Theoretical models of product differentiation often specify that each firm produces a unique product, such as in Dixit–Stiglitz models with constant elasticity of substitution utility functions, circular city models or discrete choice models (Dixit and Stiglitz, 1977; Salop, 1979; Anderson et al., 1992). Additionally, product varieties are often assumed to be symmetric in that every variety is valued equally by the representative consumer, making only the count of varieties important and not the actual labels or identities of the products.[2] While a symmetric view of variety has been quite useful in tackling many economic problems, such as in the new economic geography (Fujita et al., 1999), in the consumer cities literature variety is often described as the availability of product subcategories. For example, Glaeser et al. (2001) describe variety as a 'range' of services' and 'specialized retail'; Lee, (2010) notes that 'large cites have museums, professional sports teams, and French restaurants that small cities do not have'. The idea that bigger cities have specialized or rarer varieties of products suggests taking an asymmetric approach to describing variety: there are some varieties that may be preferred only by small subsets of consumers, or consumed less often by a representative consumer. Under this framework it becomes necessary to identify the specific labels of varieties, and not just the count of firms.

However, this categorization approach to variety is difficult to use empirically. While data on tradable goods often has quite granular categorization (e.g. SITC codes), the data on nontradable consumption amenities seldom has precise information on characteristics of the goods; further, it might not even be clear what is the appropriate classification scheme. For this reason restaurants are particularly well suited to the aims of this article. Product differentiation is easily measured: restaurant varieties can be identified by cuisine and this categorization is fairly uncontroversial.[3] Moreover, transportation costs are often significant in the decision to eat at a restaurant,[4] a factor that could make the variety in cities different from that in places with more dispersed populations. Conveniently, there exists a wealth of information about restaurants on the Internet, including precise address information allowing accurate geographic matching of firm counts to demographics data. And finally, restaurants are arguably one of the most prominent and important examples of a city's nontradable consumer goods. Couture (2013) estimates the consumption benefit of greater access to restaurants, both across and within Metropolitan Statistical Areas (MSAs), and finds large differences when comparing the most dense areas to the least dense.

The difference between the firm count and categorization approach to variety can be readily seen in Figure 1. The left-hand panel plots the number of restaurants against population for the cities (Census Places) in my data set and shows a clear log-linear pattern with a slope very close to one. This result was also found in Berry and Waldfogel (2010) and Waldfogel (2008) using different data and illustrates a strong

2 Here I use the wording of Dixit–Stiglitz to describe symmetry. Note that there are also asymmetric versions of these models; the authors discuss an asymmetric case of their model in chapter 4 of the compilation 'The Monopolistic Competition Revolution in Retrospect' (Brakman and Heijdra, 2004).

3 Obviously there is some flexibility in categorization but for the purpose of this article it doesn't matter whether a restaurant is classified as Southern Italian or Italian as long as the categorization schema is consistent across cities.

4 Intuitively, consumers are not willing to travel very far for a meal, as evidenced by the large number of identical fast food franchises in the same city.

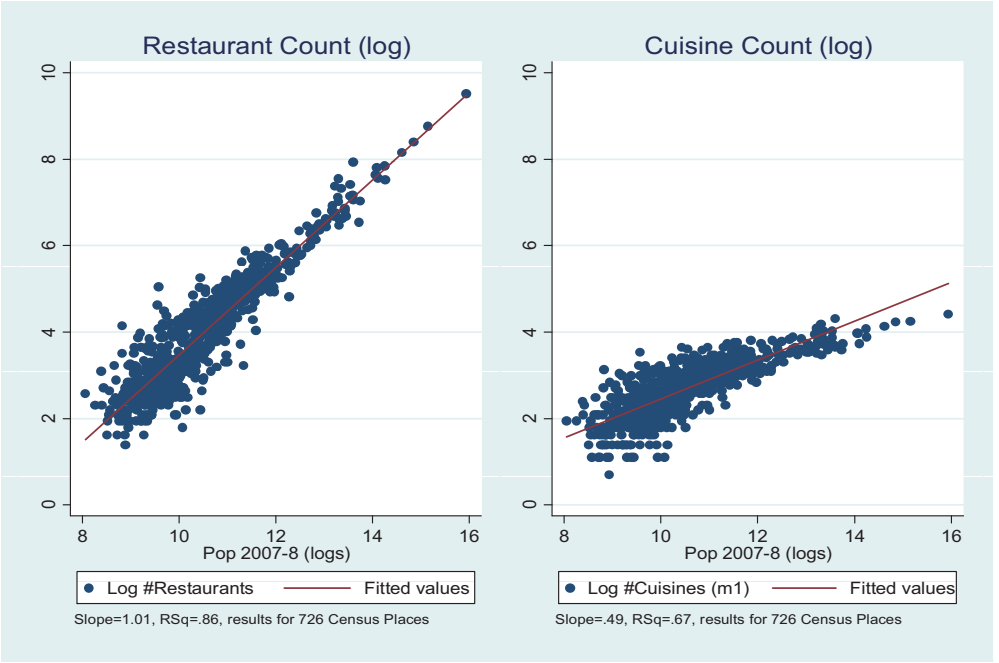


Figure 1. Restaurants and cuisines versus population.

proportional relationship between population and restaurant count. The right-hand panel shows the count of cuisines plotted against population for the same cities and displays a very different pattern with far more variance. In this article I will show that a theory of demand aggregation suggests these patterns could be quite different.

1.2. Related literature

This article is related to a large literature on market size and firm characteristics. Syverson (2004) proposes a model in which large markets with many producers allow consumers to easily switch their patronage, thus increasing competition and leading to more efficient firms—a result borne out in the author's data on the concrete industry. Campbell and Hopenhayn (2005) investigate the link between market size, average revenue, and average employment, and also conclude that competition is tougher in larger markets. Notably, they find that 'eating places', defined as places to eat with and without table service, have larger average size and greater dispersion of sizes in larger markets. The authors use population and population density as alternative measures of market size but find that their results are roughly the same. While this would seem to downplay the importance of space in market measurement their data set does not include any measures of horizontal differentiation. In my article I will set aside the issue of firm efficiency and instead focus on how market size affects differentiation, making the assumption that efficiency does not affect the degree of horizontal differentiation.

Berry and Waldfogel (2010) study vertical differentiation, or differentiation in quality, and contrast the effect of market size in industries where quality is produced with fixed costs (e.g. newspapers) to industries where quality is produced with variable

costs (e.g. restaurants). In industries where quality is tied to fixed costs, increasing market size allows a high-quality firm to undercut lower-quality firms and thus the market remains concentrated with just a few high-quality firms. On the other hand, if quality is produced with variable costs then higher-quality firms cannot undercut lower-quality competitors and larger markets will show a greater range of available qualities. The authors find that higher-quality restaurants are found in bigger cities, similar to my finding that relatively rarer restaurants are found in bigger cities. Apart from this similarity, it is rather difficult to compare their results on vertical differentiation to mine on horizontal differentiation.[5]

Two recent papers consider horizontal differentiation in the restaurant industry and look at the effects of specific populations on the cuisines of local restaurants. Waldfogel (2008) combines several datasets, including survey data on restaurant chain consumers, to show that the varieties of chain restaurants in a zip code correspond to the demographic characteristics of the population. Mazzolari and Neumark (2012) use data on restaurant location and type in California combined with Census data to show that immigration leads to greater diversity of ethnic cuisines. This article will also advance an argument about the importance of clustered groups with a particular taste but differs in subject and aim from these papers in several important ways. First, while Waldfogel is interested in showing the relationship between restaurant type and demographics at a zip code level and Neumark looks at commuting-defined markets in California, this article will focus on the city level and use a national set of data. I can then make general statements about cities and variety and relate my findings to Central Place Theory. Second, while Waldfogel looks specifically at fast food restaurants and Neumark at a larger set of mostly ethnically defined restaurants, I will use data covering all restaurants with very precise categorization (90 cuisines) so that I can provide a general measure of product diversity for a city. Finally, I develop a model to formalize how entry thresholds are affected by population and space. While this model is quite simple, it allows me to show how the product diversity of a city fluctuates with changes to overall city density. For these reasons I view this article as complementary in that I provide further evidence of the role of specific preference groups in the location choice of corresponding firms[6] but extend this finding to a general theory of demand aggregation that helps to explain why cities differ in the variety of these goods they offer.

2 Population, land area and entry

The essence of my argument is that given positive transport costs there must be enough demand in a small enough space to support a given variety. This general idea is straightforward and was mentioned in the context of restaurants in both Glaeser et al., (2001) and Waldfogel (2008). Glaeser et al. note that 'the advantages from scale

5 In vertical differentiation models all consumers agree upon standards of quality and quality differences across firms often stem from differences in costs or efficiency. Horizontal differentiation views all firms as equals that cater to different parts of the consumer taste distribution, meaning quality would not be an appropriate measure of horizontal differentiation. Additionally, the authors focus on the availability of high-quality restaurants in different markets, limiting their data to the restaurants appearing in the *Mobil* and *Zagats* guides, while I am concerned with the entire range of cuisines, but not quality.

6 Both my article and Waldfogel discuss a demand-driven explanation, but Mazzolari and Neumark (2012) argue that the supply-side, through comparative advantage in ethnic restaurant production, is the more important channel through which local ethnic groups lead to local ethnic restaurants.

economies and specialization are also clear in the restaurant business where large cities will have restaurants that specialize in a wide range of cuisines–scale economies mean that specialized retail can only be supported in places large enough to have a critical mass of consumers'. Waldfogel writes: 'some products are produced and consumed locally, so that provision requires not only a large group favoring the product but a large number nearby'. While this demand-side argument is intuitive it is still not clear what defines a critical mass and how entry might be affected by markets with different populations and land areas. Implicitly, a critical mass is a concept about population density but is density sufficient for making predictions about entry or does the effect of density depend upon the geographic area of the market? To address this and other issues I will use a simple model with the objective of showing how the two dimensions of demand aggregation—population and population density—can affect variety even when all cities have identical characteristics. I base this model on Salop's circular city model of monopolistic competition but rather than normalizing population and land to one I keep these as separate parameters. I add cuisine-specific entry thresholds in the spirit of Bresnahan and Reiss (1991)[7] to define the minimum conditions that would allow the first firm of a cuisine to enter the market. I then map these entry thresholds in population and land space to show how product variety will be affected by different market sizes. I should emphasize that there are other models of demand aggregation[8] and that in relating entry thresholds to market size my model turns out to closely overlap models from Central Place Theory, especially that of Hsu (2012). In this article I present the model not as a substantial theoretical contribution but rather as a simple framework to clearly guide the empirical analysis and illustrate the parallels to Central Place Theory that I observe in the data.

I will be discussing a market in which firms are differentiated by both product type and location, and must choose a price given a market with free entry and positive transportation costs. If firms differentiate by both type and location then consumers may value different configurations of firm locations and types equally. For example, a consumer may be indifferent between their preferred variety at one distance and a lesser-preferred variety at a shorter distance. This trade-off between distance and type could result in multiple equilibria or imply the nonexistence of an equlibrium.[9] In order to make the model tractable I make the strong assumption that firms of different types don't compete or in this context, that restaurants of different cuisines don't compete. While this is clearly unrealistic, I believe that the intuition gained from this model still holds for several reasons. First, the idea that there must be a minimum number of

7 Bresnahan and Reiss (1991) develop a model of entry thresholds in which a market must satisfy specific conditions in order to permit entry of each successive firm in an industry (from monopolist to perfect competitor).

8 There are of course many models with increasing returns in the spirit of Krugman (1991) that predict greater variety in larger markets. These models usually feature transportation costs that affect trade across cities or regions, while I am interested in the effect of within city transportation costs on variety. More recent work, such as Ottaviano et al. (2002) and Behrens and Robert-Nicoud (2014), incorporates both increasing returns and within city transportation costs but the models have a broader focus, and are thus more complex, than what I present here.

9 Salop specifies a symmetric, zero profit equilibrium where firms are equally spaced and all earn zero profit. If there are multiple types and asymmetric competition, meaning that a firm of one type competes more strongly with firms of the same type than with firms of other types, then this equilibrium often can't exist. For example, there can not be an even number of firms of one type and an odd number of firms of another type since some firms would have different neighbors and face more competition, and thus profit cannot be zero for all firms.

consumers in a small enough space to allow entry of a given type must hold with and without inter-type competition.[10] Second, the assumption of no competition across types can be viewed as an extreme form of the assumption that firms of the same type compete more closely with each other than with other types. In this article I am interested in simply showing that there must be a minimum number of consumers in a small enough space to permit a firm of that type to exist, and I do not try to predict the count of restaurants in each cuisine for each city[11] nor within city location patterns.[12] Narrowing my aims to this objective allows me to make the strong assumption that firms of different types don't compete.

An alternative supply-side argument would suggest that specific types are only found in cities with restaurant owners (restaurateurs) of that type. However, the supply-side explanation alone is not very convincing: if there is significant demand in a city for a specific type then a restaurateur of that type should move to the city. I therefore focus on providing evidence for this critical mass argument but discuss this alternative explanation in the empirical section.

2.1. Entry threshold for a single firm

Following Salop's circular city model (Salop, 1979), there is a total population N of consumers located uniformly around a circle with perimeter L. Each consumer must decide whether to purchase a good from a firm or consume their reserve good while firms also locate around the perimeter of the circle and can enter the market freely. Consumers are utility maximizers and receive utility u_1 from the firms' product and utility u_0 from the reserve good, which I normalize to zero without loss of generality. There is positive transportation cost per unit distance, τ, and thus a consumer located at

10 Consider a market where there are two types of firms, consumers like both types, but all consumers prefer one type to the other. When two asymmetrically differentiated firms compete the greater-preferred firm always has an advantage over the lesser-preferred firm, and thus demand conditions must be quite favorable for the lesser-preferred firm to exist. I will show that even without competition across types, a market must have a minimum population, which varies with land area, in order to sustain any given type. Therefore allowing competition between firms simply makes it even more difficult for a lesser-preferred firm to exist.

11 Mazzeo (2002) develops a model of oligopoly where firms make entry and product quality decisions simultaneously and estimates the distribution of motels across quality types for small exits along US highways. Unfortunately this model is less relevant to the monopolistically competitive restaurant industry where simultaneous entry of thousands of differentiated firms seems a very strong assumption. Further, while Mazzeo looks at markets with a small number of motels across three quality types, the markets in my dataset have thousands of restaurants with up to 82 cuisines, making estimation intractable.

12 In their theoretical paper Irmen and Thisse (1998) show that when duopolists compete in multiple dimensions they will choose to maximally differentiate in one dimension and minimally differentiate in all other dimensions. It is tempting to try and apply this result to the context of restaurant locations but for several reasons the industry is not a good fit. As noted above, cities have thousands of restaurants and Tabuchi (2009) shows that once there are three or more firms this max–min result no longer holds. Further, restaurant consumers are not uniformly distributed in location or characteristics space, consumer tastes themselves do not fit easily into a continuous space (how different is Italian from Japanese?), and restaurants enter markets sequentially over long periods of time making current location patterns path dependent. Studying location choice with product differentiation in the restaurant industry is a fascinating topic but beyond the scope of this article.

l will purchase from a firm located at l_i with product price p_i only if the net utility of this transaction is higher than the consumer's reserve utility.

$$\max\left[u_1 - \tau|l_i - l| - p_i, 0\right] \tag{2.1}$$

This article is concerned with entry and thus I focus on the case of a single firm deciding whether to enter the market, a potential monopolist's entry problem (under my assumptions the model's implications are unchanged with multiple firms—see Appendix A). A consumer will only purchase from this firm if the price and transportation costs are low enough. Define d as the distance that would make a consumer indifferent between the firm's good and their reserve utility:

$$d = \frac{u_1 - p}{\tau} \tag{2.2}$$

From the firm's perspective, the geographic extent of their market (g) is the sum of the distances to the indifferent consumer on either side of the firm: $g = 2 * d$. We then have:

$$p = u_1 - \frac{\tau g}{2} \tag{2.3}$$

Demand for the firm's product is the geographic extent multiplied by the population density, $D = \frac{N}{L}$. The firm must pay a fixed cost F to enter and then produces the good with constant marginal cost c, making profit:

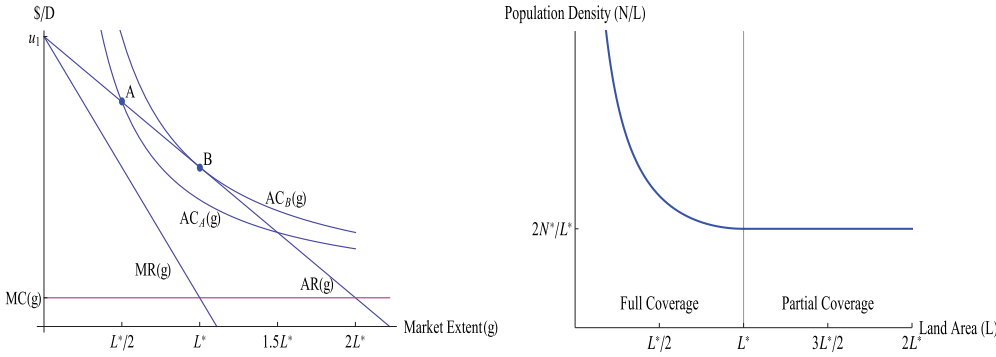
$$\Pi = \left(u_1 - \frac{\tau g}{2} - c\right)Dg - F \tag{2.4}$$

The monopolist will choose the geographic extent that maximizes profit, which I will refer to as L^*:

$$g^* = \frac{u_1 - c}{\tau} \equiv L^* \tag{2.5}$$

If the total market area is less than the profit-maximizing geographic extent, $L < L^*$, the monopolist will sell to the whole market, $g = L$. By writing average revenue, marginal revenue and marginal cost as a function of g I can show the monopolist's problem of choosing the geographic extent graphically in Figure 2.[13] The quantity sold is $q = Dg$, and thus density is just a scaling factor for these three functions that does not affect the monopolist's choice of g. Given this, I scale the vertical axis by density, making the units $\frac{\$}{D}$, in order to show the problem for any density. In the left hand panel of Figure 2 the horizontal axis shows the monopolist's choice of g. If the market area is large enough, $L \geq L^*$, the monopolist chooses $g = L^*$ where the marginal revenue curve intersects marginal cost. If $L < L^*$ the monopolist is constrained and chooses $g = L$; this case is analogous to a monopolist choosing quantity with a capacity constraint. I define 'full coverage' as the situation where the single firm sells to every consumer in the market, or $g = L$, and 'partial coverage' as when the firm sells to a subset of consumers, $g = L^* < L$. The two cases converge at $L = L^*$ since the monopolist chooses $g = L^*$. Given the monopolist's choice of g, the necessary condition for entry is that profit is

13 Average revenue is $AR(g) = D\left(u_1 - \frac{\tau g}{2}\right)$, marginal revenue is $MR(g) = D(u_1 - \tau g)$, and marginal cost is $MC(g) = Dc$.



(a) Monopolist Choice of Geographic Extent (b) Minimum Density Given Land Area

Figure 2. Geographic extent and minimum density.

weakly greater than zero. I consider the boundary case where profit is equal to zero—the minimum condition for the first firm to enter—and write average revenue and average cost in terms of g:

$$AR(g) = AC(g) : D\left(u_1 - \frac{\tau g}{2}\right) = Dc + \frac{F}{g} \tag{2.6}$$

$$D_{min}(g) = \frac{F}{g\left(u_1 - c - \frac{\tau g}{2}\right)} \tag{2.7}$$

Equation (2.7) yields a unique density for every g: this is the minimum density in the market that would allow the first entrant. Point A in the left panel of Figure 2 shows an average cost curve drawn for the level of density that would make the monopolist's profit equal to zero for geographic extent $L^*/2$.[14] Point B shows the average cost curve for $g = L^*$, and is tangent to average revenue since marginal revenue is equal to marginal cost at L^*.

Incorporating the monopolist's choice of g given the city's land area L, I plot the population density that makes profit equal to zero against L in the right hand panel of Figure 2. It can be seen that the required density decreases monotonically until reaching a minimum at $L = L^*$. If the monopolist were to choose a market extent $g > L^*$ the required density would actually increase since marginal profit is negative. However, for all $L > L^*$ the monopolist chooses $g = L^*$ and thus the required population density is constant for markets with area greater than L^*. Figure 2 also shows that the minimum density rises without bound as land area shrinks to zero and so it can be helpful to express Equation (2.7) in terms of population by multiplying both sides by land area. By setting $g = L$ (full coverage) and $L = 0$, I can solve for the absolute minimum population

14 Average cost is plotted on the $\$/D$ scale and thus the curve plotted is average cost divided by density: $c + \frac{F}{Dg}$.

required to allow entry when consumers incur no transportation costs (no land), denoted as N^*:

$$N^* \equiv \frac{F}{u_1 - c} = \frac{F}{\tau L^*} \tag{2.8}$$

The constant density required for markets with land area of L^* or greater is thus $D = \frac{2N^*}{L^*}$. Using the firm's optimal choice of g given a market's land area, plugging N^* and L^* into Equation (2.7), and rearranging in terms of N, I can write an expression giving the minimum population required for entry as a function of land:

$$N_{min}(L) = \begin{cases} \dfrac{2N^*L^*}{2L^* - L} & \text{if } L < L^*, \text{'full coverage'} \\ \dfrac{2N^*L}{L^*} & \text{if } L \geq L^*, \text{'partial coverage'} \end{cases} \tag{2.9}$$

To better understand the intuition of Equation (2.9) it can be helpful to think about increasing the land area of a city from zero and mapping how the minimum population required for entry changes. When land area is equal to zero, $L = 0$, consumers have no transportation costs and the monopolist can charge a price equal to the full difference between the consumer's utility from the monopolist's good and the reserve good. The revenue at this price will just cover the monopolist's fixed cost at the minimum population level N^*. As land area increases from zero consumers bear transportation costs, requiring the monopolist to lower the price, and thus the minimum population required increases. The required rate of population increase is proportionally less than the increase in land area and thus minimum population density declines. At land area $L = L^*$ the monopolist has reached the profit maximizing market extent and thus will not further lower the price, preferring to sell only to those within this geographic extent. This implies that further increases in land area must be accompanied by proportional increases in population so that the population within the monopolist's fixed market extent is always the same, or that population density is always $D = \frac{2N^*}{L^*}$.

2.2. Population, land and variety

To investigate variety I will assume that consumers have heterogeneous tastes and will only consume their preferred variety. I define the proportion of consumers who like variety v as δ_v. There are V different varieties in the market and consumers have a taste for only one variety, making $\sum_{i=1}^{V} \delta_i = 1$. I further assume that consumers are located uniformly throughout the perimeter of the circle so that if I were to randomly select any segment the percentage of consumers favoring cuisine v is always equal to δ_v.[15] A firm of type v only sells to consumers of type v, who have total mass equal to $N * \delta_v$. All firms, regardless of type, have the same fixed cost F and marginal cost c and thus require the same minimum consumer population for each value of land. If the market has N

15 An alternative and equivalent assumption would be to assume all consumers are identical and consume δ_v amount of each cuisine. Note that this alternative assumption could be considered one form of a 'taste for variety'. If one city offers more variety than another then the representative consumer will consume more restaurant meals in the more diverse city.

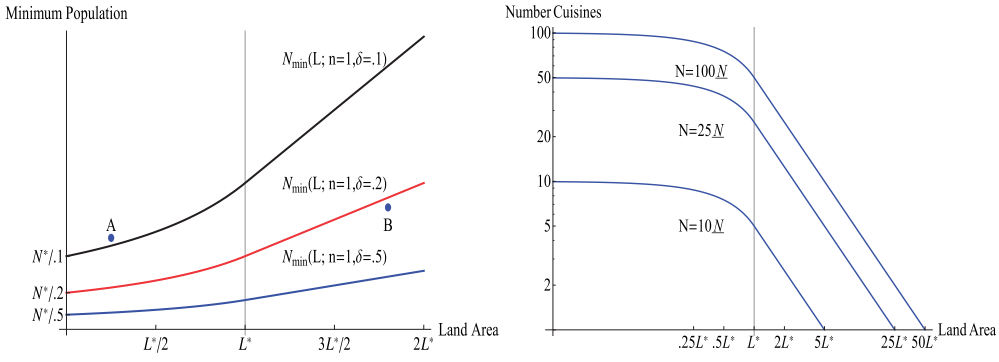
consumers but only δ_v percent can potentially consume product type v then the minimum conditions for a firm of type v to enter the market become:

$$N_{min}(L;\delta_v) = \begin{cases} \frac{1}{\delta_v} * \frac{2N^*L^*}{2L^* - L} & \text{if } L < L^*, \text{ 'full coverage'} \\ \frac{1}{\delta_v} * \frac{2N^*L}{L^*} & \text{if } L \geq L^*, \text{ 'partial coverage'} \end{cases} \tag{2.10}$$

In this way Equation (2.10) defines a variety-specific population threshold for each value of land area; if a city's total population is below the threshold given its land area then it cannot support that variety. I can rank varieties by the percentage of consumers who prefer the variety, δ_v. In order for a firm to enter a market and sell a low δ_v variety the market must have a relatively higher population and relatively smaller land area. Further, the minimum population required for less-preferred varieties increases faster in the amount of land in a city:

$$\frac{\partial N_{min}(L;\delta_v)}{\partial L} = \begin{cases} \frac{2N^*}{\delta_v L^*} * \frac{L^*}{(2L^* - L)^2} & \text{if } L < L^*, \text{ 'full coverage'} \\ \frac{2N^*}{\delta_v L^*} & \text{if } L \geq L^*, \text{ 'partial coverage'} \end{cases} \tag{2.11}$$

The left panel of Figure 3 provides an example of mapping varieties to markets in land–population space. In the figure, market A has three varieties while market B only has one variety, despite having a larger population. The number of varieties is increasing to the north and west as cities cross the thresholds for different varieties. Holding land constant, more populous markets will have more varieties and holding population constant, smaller geographic markets will have more varieties. Throughout this article I will interpret the effect of changing land area, holding constant population, as the effect of changing population density, independent of population size.



(a) Multiple Types (b) Zipf Tastes (log-log scale)

Figure 3. Population and land thresholds, count of varieties.

The left panel of Figure 3 also shows a hierarchical relationship between the number and composition of varieties found in a market. Specifically, if market i has more varieties than market j then market i will have all of the varieties found in market j; if a market has more varieties than another it is because it can support the less-preferred (smaller δ_v) varieties. More formally, since the threshold lines (frontiers) cannot cross—$N_{min}(L; \delta_i) > N_{min}(L; \delta_j)$ if $\delta_i < \delta_j$—I can define each frontier uniquely by its population intercept, $\frac{N^*}{\delta_v}$. Each market can then be defined by the intercept of the frontier that would pass through the market's location in land–population space, I_m. Since each intercept can be interpreted as $\frac{N^*}{\delta_v}$, this v is the variety preferred by the smallest percentage of people, or rarest taste, that market m can support. For example, in Figure 3 market B is on a frontier between the first and second varieties and has an intercept between $N^*/.5$ and $N^*/.2$. Generally, for markets m and l, if $I_m > I_l \Rightarrow \#Varieties_m \geq \#Varieties_l$ where the weak operator stems from the fact that with discrete varieties a higher frontier may still not be high enough for the next variety threshold. I can use this relationship to make ordinal statements about how population and land affect the number of varieties in a market.

$$\#Varieties_m \sim I_m = N_m\left(1 - \frac{L_m}{2L^*}\right)\mathbf{1}(L_m < L^*) + \frac{N_m L^*}{2L_m}\mathbf{1}(L_m \geq L^*) \tag{2.12}$$

From Equation (2.12) it can be seen that the count of varieties is increasing in population and decreasing in land area.[16] These implications are not reliant upon a specific distribution of tastes but do require that tastes are similar enough across cities to yield the hierarchical relationship linking the number and composition of varieties found in a market. In order to better illustrate this model's predictions for the effect of population and land on variety and tie the theory closer to the empirical work I now show these effects when tastes follow the Zipf distribution. The Zipf distribution is analytically convenient because it offers a simple form for proportional tastes that increase in rarity; however, my empirical work makes no distributional assumptions about tastes. Below I show the Zipf probability mass function (pmf) $f(v; s, V)$ where v is the rank of a discrete variety, V is the total number of varieties, and s is a shape parameter, with $s > 1$:

$$f(v; s, V) = \frac{1/v^s}{\sum_{v=1}^{V}(1/v^s)} = \frac{1}{v^s H_{V,s}}, \quad \text{where} \quad H_{V,s} \equiv \sum_{v=1}^{V}(1/v^s) \tag{2.13}$$

The pmf $f(v; s, V)$ has the same interpretation as δ_v, representing the percentage of people who like a variety of rank v, with smaller percentages for higher v varieties. I can find the threshold population for a variety of rank v by plugging the above expression into Equation (2.10) for δ_v. I then invert this expression to get the highest rank cuisine a city can support, v_m, which is also the maximum number of cuisines in the city:

$$v_m \leq \left[\frac{N_m(2L^* - L_m)}{2H_{V,s}N^*L^*}\right]^{(1/s)} * \mathbf{1}(L_m < L^*) + \left[\frac{N_m L^*}{2H_{V,s}N^*L_m}\right]^{(1/s)} * \mathbf{1}(L_m \geq L^*) \tag{2.14}$$

16 The comparative statics are $\frac{\partial \#Varieties}{\partial N} \sim \left(1 - \frac{L}{2L^*}\right)\mathbf{1}(L < L^*) + \frac{L^*}{2L}\mathbf{1}(L \geq L^*) > 0$ and $\frac{\partial \#Varieties}{\partial L} \sim \frac{-N}{2L^*}\mathbf{1}(L < L^*) + \frac{-NL^*}{2L^2}\mathbf{1}(L \geq L^*) < 0$.

In the empirical work I will estimate log–log specifications and so taking logs and assuming the above holds with equality gives:

$$ln(v_m) = \begin{cases} (1/s) * \left(ln(N_m) + ln(2L^* - L_m) - ln(2L^*) - ln(\underline{\mathrm{N}})\right) & \text{if } L_m < L^* \\ \\ (1/s) * \left(ln(N_m) - ln(L_m) + ln(L^*) - ln(2\underline{\mathrm{N}})\right) & \text{if } L_m \geq L^* \end{cases} \tag{2.15}$$

I define $\underline{\mathrm{N}} \equiv (H_{V,s}N^*)$, analogous to N^*, as the minimum population required to have any varieties when land area is zero. The comparative statics (elasticities) for the effect of log population and log land on log variety are:

$$\frac{\partial ln(v_m)}{\partial ln(N_m)} = \frac{1}{s} \tag{2.16}$$

$$\frac{\partial ln(v_m)}{\partial ln(L_m)} = \left(\frac{-L_m}{2L^* - L_m} * \frac{1}{s}\right)\mathbf{1}(L_m < L^*) + \left(\frac{-1}{s}\right)\mathbf{1}(L_m \geq L^*) \tag{2.17}$$

The first equation shows that the effect of log population on log variety, or the elasticity, is constant and not a function of land area. In the second equation the partial coverage term ($L_m \geq L^*$) is larger in absolute value than the first term. This implies that the log number of varieties decreases faster in log land when land area is larger than the threshold L^*, or that the elasticity of variety with respect to land area increases with land area. An alternative and equivalent interpretation is that increases in population density, holding constant population level, increase variety more for cities with greater surface area. These effects are easily seen in the right hand panel of Figure 3, where I plot equation 2.15 in log–log scale for three population levels, setting $s = 1$ (or arbitrarily close to 1 since $s > 1$). For any land area a proportional increase in population (a unit increase in log population) has the same effect on log variety, shown in Figure 3 by the parallel lines for the three population levels. The steepening downward sloping curves show that an increase in log land has a small effect for small land areas which becomes much larger, and constant, for land areas greater than L^*. Lastly, the flattening of each line as land area decreases to zero shows that every city is constrained by its population to a maximum count of cuisines, here $N_m/\underline{\mathrm{N}}$, irrespective of density.

To summarize, the theory has the following implications:

1. The existence of a variety in a market can be determined by an entry threshold in land–population space where the minimum population threshold is both higher and increases faster with land area for less-preferred varieties
2. For a given variety, more populous markets and smaller geographic markets are more likely to have the variety
3. If proportional tastes are the same across markets there will be a hierarchical relationship between the number of varieties and the composition of those varieties and this hierarchy can be predicted by population and land area
4. The elasticity of variety with respect to land area increases with land area, meaning the negative effect of land area on variety is greater for larger land area markets.

3 Describing city product variety

3.1. Data collection

In order to study the relationship between market characteristics and product variety, one needs a data source that satisfies several requirements. First, there must be a consistent categorization of firms into varieties. Second, the data set must be exhaustive; if there is only data on select firms in the market then it is not possible to compare counts of variety across markets. Third, the data set must have precise geographic information on locations so that firms can be matched with the appropriate data on market characteristics. Finally, the data set must cover a sufficient number of markets to allow comparisons across markets. These requirements rule out the use of restaurant guides (not exhaustive, only cover a few markets) and information directories or yellow pages (inconsistent categorization of types). The online city guide Citysearch.com has exhaustive listings and covers many US cities.[17] Additionally, while the same restaurant may be listed under multiple cuisine headers, the actual entry for that restaurant lists one consistent, unique cuisine. Furthermore, the entry for the restaurant also lists an exact street address. This allows me to avoid the difficulty of reconciling census city boundary definitions to the boundary definitions used by the site in order to find the appropriate city demographic (Census) information. In the spring of 2007 and the summer of 2008 I used a software package and custom programming to download all the restaurant listings for the largest cities listed on the site.

I attempted to collect data for the 100 most populous US Census places and ended up finding 88 Citysearch cities matching these Census places. However, the Citysearch definition of cities often extended far beyond the Census place geography. For example, the listing for 'Los Angeles' included restaurants in the Census place boundaries for 'Beverly Hills', 'Long Beach', 'Santa Ana', 'Santa Monica' and many smaller cities. I therefore ended up collecting many Census places within the vicinity of large cities. Since I will be using demographic data to explain characteristics of a city's restaurant industry it is important that my data for each city is fairly complete. In order to gauge how close the Citysearch data was to the complete set of restaurants for every Census place I matched the count of Citysearch restaurants to the count found in the 2007 Economic Census, combining the categories 'Full Service Restaurants' (NAICS 7221) and 'Limited Service Eating Places' (NAICS 7222). While close to the Citysearch types of restaurants, the Economic Census includes some restaurants not always covered by Citysearch data, such as fixed location refreshment stands, and therefore I expect that the Economic Census will record more restaurants for every Census place than Citysearch. I define the 'match ratio' for each Census place as $MatchRatio = \#CityseArchRestaurants/\#EconomicCensusRestaurants$ and keep all Census places with a match ratio between 0.7 and 1.1, leaving 726 Census Places.[18]

17 When I collected my data this website appeared to be the most popular of its type; today Yelp.com has many of the same features and seems to be more prominent.

18 I dropped New Orleans because of damage from Hurricane Katrina and Anchorage, Alaska because the Census place geographic definition is very different from other cities (it includes an enormous swath of virtually uninhabited land). I also dropped Industry, California, which is almost entirely industrial (only 777 residents but many firms) and thus a very large outlier in restaurants per capita. Washington, DC is not in my data set because the DC site was hosted by the Washington Post and used a completely different page format, leading to difficulties in data collection.

For each restaurant I have a cuisine designation, such as 'Italian' or 'Ethiopian'. In Table D1 (Appendix D), I list all the cuisines and the count of cities in each land quartile with that cuisine (some of the empirical work will also use land quartiles, discussed in Section 4.1). The rarest cuisines are Armenian and Austrian, found in just two cities, while there are a number of cuisines found in every or nearly every city (Chinese, Deli, Fast Food, Italian, Mexican, Pizza). This is the 'primary' cuisine listed for the restaurant; there may be German restaurants that serve Austrian food and use 'Austrian' as an additional cuisine but only one city has restaurants whose primary cuisine is Austrian.[19]

Table 1 shows summary statistics for the 726 cities by land quartile. Both the number of restaurants and number of cuisine are larger for geographically bigger cities. The demographic characteristics, with 'Young' and 'Old' categorizations following Berry and Waldfogel (2010), are quite similar across land quartiles. I define the statistic ethnic HHI, analogous to the Herfindahl–Hirschman Index of market competition, as the sum of the squared shares of each ethnicity in a city and use this as a measure of ethnic diversity. The theoretical range of this measure is from one (a single ethnic group) to zero (infinite number of equal-sized groups) and I find that larger cities have somewhat lower values, indicating greater ethnic diversity.

3.2. Descriptive evidence

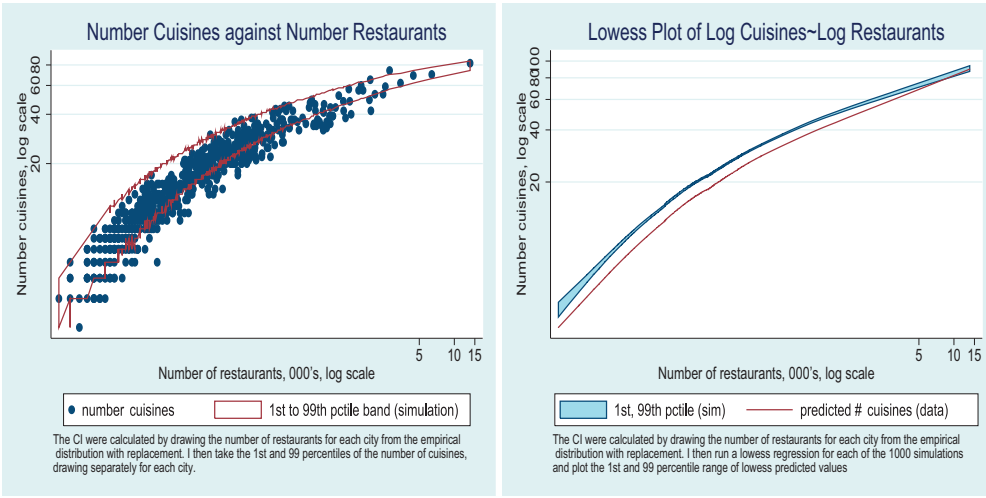
Before testing the model's predictions I first provide some descriptive evidence of how restaurant diversity differs across cities. In Figure 4 the dots in the left panel represent the number of cuisines plotted against the number of restaurants for all 726 Census Places, where both axes are in logs. As the number of restaurants increases the number of cuisines first rises rapidly and then becomes roughly linear (in logs). In the preceding section I proposed a theory suggesting an individual city's characteristics determined the number of its cuisines through demand aggregation. However, a simpler theory could be that the cuisine of any restaurant is just a random draw from an exogenous nationwide distribution of restaurants to cuisines. For example, if the cuisine of a restaurant is determined entirely by the cuisine-specific skills of the restaurateur, and some skills are rarer than others, then cities with many restaurants are more likely to have a rare cuisine and cities with few restaurants will have just common cuisines. This is equivalent to assuming that every restaurant exists in its own independent market, unaffected by city level characteristics, and thus aggregating a city's n restaurants and looking at features of the cuisines should be no different from randomly drawing n restaurants from many different cities. Therefore, as a benchmark for comparison, I will look at the pattern that would result from randomly assigning cuisines to restaurants based on the observed aggregate distribution of restaurants to cuisines across all cities.

To create this benchmark pattern I draw restaurants for each city from the nationwide pool with replacement and then *for each city* show the 1st and 99th

19 Earlier versions of this article also used cuisine-price pairs, such as 'Italian $$', as an additional measure of variety. The patterns I found with this measure of variety were very similar to using the primary cuisine. However, price information was missing for about 60% of the restaurants and using the price may have been mixing horizontal and vertical differentiation. I therefore only use the primary cuisine when measuring variety.

Table 1. Summary statistics by land quartile

	Land quartile 1 ($n = 182$)				Land quartile 2 (n = 181)				Land Quartile 3 (n = 182)				Land Quartile 4 (n = 181)			
	mean	SD	min	max	mean	SD	min	max	mean	sd	min	max	mean	sd	min	max
# Restaurants	539.4	1250.7	8.0	13,664.0	93.4	72.2	6.0	380.0	53.8	52.7	5.0	359.0	27.4	29.7	4.0	192.0
# Cuisines	29.4	14.1	3.0	82.0	19.0	8.0	3.0	45.0	14.9	7.8	3.0	43.0	10.5	6.7	2.0	38.0
Pop 2007-08 (000's)	336.5	755.3	7.2	8328.5	54.9	35.8	6.7	239.2	30.4	21.0	4.6	107.1	16.3	11.0	3.1	75.7
Land Area (sq km)	232.5	298.6	61.5	1962.4	44.2	9.1	30.1	61.3	21.8	4.3	15.0	30.0	9.7	3.2	2.6	14.9
Density (000's/sq km)	1.3	1.2	0.1	10.6	1.3	0.8	0.2	6.2	1.4	0.9	0.3	6.4	1.8	1.3	0.3	12.1
MSA Pop 2000 (mns)	4.5	4.5	0.1	21.2	5.5	5.1	0.2	21.2	5.6	5.0	0.3	21.2	5.4	5.1	0.3	21.2
Average HH Size	2.6	0.3	2.0	4.1	2.6	0.3	2.0	3.7	2.6	0.3	1.8	3.6	2.6	0.5	1.7	4.4
Med HH Inc (000's)	49.2	15.4	24.5	111.8	56.5	19.4	26.8	139.9	52.0	16.4	25.2	146.5	50.3	17.6	17.7	134.3
Ethnic HHI (x100)	75.3	15.3	24.0	97.5	77.3	17.9	17.6	100.0	78.7	18.6	25.7	99.5	79.0	18.6	26.2	99.8
%Old (> 64)	10.4	4.1	3.3	34.5	11.7	4.7	3.1	30.2	12.4	5.4	2.8	37.2	13.7	6.5	3.7	43.2
%Young (< 35)	51.7	5.9	28.0	66.2	48.9	6.3	32.5	67.6	49.1	7.3	27.4	81.2	48.0	8.2	21.2	69.4
%College grad	36.2	12.8	7.1	70.6	39.0	16.1	11.1	78.5	36.6	14.5	10.3	75.2	33.8	17.3	4.3	81.5

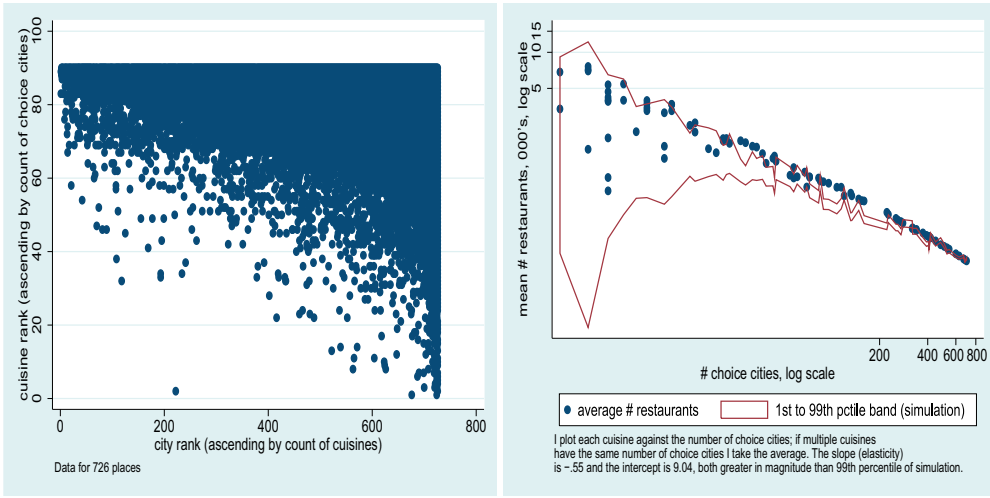


(a) Percentile Bands (b) Lowess Plot

Figure 4. Number of cuisines versus number of restaurants.

percentiles of the number of cuisines from 1000 simulations. The confidence bands shown in the left-hand panel of Figure 4 for each city do not come from the same replication (simulation run) but rather represent the city-specific bands across all runs, similar to looking at a distribution of maxima. The figure shows that for many cities the observed number of cuisines is significantly below even the 1st percentile of the simulations. Results from simulations without replacement are even further from the observed data. To look at the likelihood of finding so few cuisines across many cities in the same replication I run locally weighted regressions (lowess) of log cuisines against log restaurants for each replication and then show the 1st and 99th percentile bands from the predicted values for each city separately. I compare this to the predicted values from a lowess regression on the observed data. While the percentile bands for two cities could still come from different replications, the predicted values for any city are affected by all the other cities in that run and thus this technique is a conservative approximation to looking at the whole cuisine-restaurant distribution across many replications. The confidence bands are quite narrow and the observed data is significantly below the 1st percentile for every city except the very largest cities, of which there are few. Therefore this random draw benchmark does not fit the observed pattern of cuisines across cities.

In addition to the count of cuisines across cities, there is also a systematic, hierarchical pattern to the specific cuisines found in each city and the number of restaurants. Mori et al. (2008), or MNS, investigate the distribution of industries across Japanese cities and find a strong hierarchical pattern that results in several empirical regularities. Hsu (2012) finds a similar pattern in US data and proposes a model of Central Place Theory consistent with these regularities. In Hsu's model (2012), scale economies in production lead to a central place hierarchy but he also notes that heterogeneity in demand would lead to the same hierarchical structure of goods across cities. While in this article I focus on nontradable consumer products and also incorporate spatial concentration in addition to population size, the underlying



(a) Hierarchy diagram

(b) Number average size plot

Figure 5. Central Place Theory patterns of cuisines across cities.

mechanism of my model—demand aggregation—is consistent with Hsu's model. I therefore use several of the techniques from MNS (2008) to describe the pattern of cuisines across cities and show that this pattern is quite similar to the Central Place patterns found in the distribution of industries across cities.

In the left panel of Figure 5, I show a hierarchy diagram from MNS for cuisines. For each cuisine I count the number of cities that have that cuisine, 'choice cities', and for each city I count the number of cuisines in the city. On the vertical axis I plot the rank of cuisines by decreasing choice cities (rank 1 is the least common cuisine) and on the horizontal axis I plot the rank of cities by increasing number of cuisines (the rank 1 city has the fewest cuisines).[20] Each observation is a city–cuisine pair and it can immediately be seen that cities with more cuisines also have the rarer cuisines. Further, if a city has a rare cuisine that city also tends to have all of the more common cuisines. This implies that simply knowing the count of cuisines allows prediction of the specific cuisines found in the city.

MNS propose a hierarchy statistic that captures this pattern and then compare the statistic to draws from a uniform multinomial. When I do this (see Appendix B) I find that the pattern of cuisines across cities is far more hierarchical than that from a uniform distribution across cuisines. However, their test implies that comparison distributions must be at the cuisine level, meaning that in simulations for each city I draw n cuisines, rather than drawing n restaurants and using this draw to determine the

20 While MNS plot the actual choice city count and number of industries I use the rank because there are cuisines with the same count of cities, and cities with the same count of cuisines, and thus many observations would be hidden behind single points. Here I break ties arbitrarily but use this only as a graphical method; in the formal testing of hierarchy I will follow the test from MNS exactly, which allows for ties.

number of cuisines.[21] In order to compare the hierarchical pattern in the data to simulations from the empirical distribution of restaurants I use a closely related technique from the same article, the Number Average Size plot.

For each cuisine I calculate the average number of restaurants found in the corresponding choice cities. In right panel of Figure 5, I plot the 90 cuisines with the choice cities on the horizontal axis and the average number of restaurants on the vertical axis (dots). As an example, the point (222, 462) corresponds to the Indian cuisine and indicates there are 222 cities with an Indian restaurant and that the average number of restaurants (across all cuisines) for those cities is 462. In the plot the relationship between the average number of restaurants and choice cities is linear in logs with an elasticity of −0.55, indicating that relatively rare cuisines, those with few choice cities, are found only in cities with large numbers of restaurants.

I compare this pattern to the 1st and 99th percentiles from simulations in which I draw from the empirical distribution of restaurants with replacement, as done in the earlier exercise. Again, the bands are calculated at the cuisine level and thus represent the percentiles across all simulations for each cuisine. The plots shows that the average number of restaurants in the choice cities of each cuisine is often higher than the 99th percentile from the simulations, indicating that rare cuisines are found in fewer and larger cities than would be predicted. While the simulation results are also linear in logs, that pattern is not nearly as steep with both the intercept and the elasticity from the observed data greater than the 99th percentile of the simulations. The observed pattern of cuisines across cities is significantly more hierarchical than the benchmark.

The results from this section follow the model's prediction that rarer cuisines are only found in cities with more cuisines. There is a strong hierarchical pattern to the distribution of cuisines across cities that is statistically different from the distribution of cuisines across restaurants. This pattern is consistent with the proposed model of entry thresholds and quite similar to the Central Place Theory regularities found in the distribution of industries across cities from other papers.

4 Empirics

In this section I evaluate the implications of the model and attempt to estimate the causal effects of population and land area, separately, on city-level variety. I do not attempt to structurally estimate the model, which relied on simplifying assumptions I wish to relax, but rather employ an empirical specification that incorporates the main features of the model. First I present cross-city ordinary least squares (OLS) results and then show a set of results instrumenting for population and land area. I conclude the section with a series of robustness exercises in which I discuss additional issues related to sorting and the spatial clustering of ethnic populations.

21 The problem with drawing at the restaurant level is that even for strongly hierarchical distributions, cities with many restaurants will receive many cuisines. This leads to many cuisines being found in every city, which differs from the observed pattern, but still results in a high-hierarchy statistic.

4.1. Variety across cities

The model in the theory section showed how demand aggregation affects variety when cities are identical in characteristics and suggested that the effect of land area on variety depended on whether the market was fully or partially covered. In order to incorporate differences in the characteristics of cities and allow for nonlinearity in the effect of land I take a reduced form approach based on the log–log specification from Equation (2.15). In Appendix C, I provide a more detailed derivation of this reduced form specification but essentially I make two main changes to the theoretical Equation (2.15). First, I include a vector X_m of demographic percentages and population characteristics (e.g. median income, percent college educated). Second, I replace the kinked function of land area with the simple log of land area, $ln(L_m)$, but allow for nonlinearity by running regressions separately by land quartile or with a quadratic term in land area. This yields the following specification:

$$\ln(\#Cuisines_m) = \gamma_0 + \gamma_1 ln(N_m) + \gamma_2 ln(L_m) + X_m'\beta + \varepsilon_m \tag{4.1}$$

In Equation (4.1) the m subscripts stand for market (city), N_m is population with γ_1 predicted to be positive, L_m is land with γ_2 predicted to be negative, and X_m is the vector of demographic controls shown in summary Table 1. In addition, I include 45 ethnicity control variables, calculated as the percentage of the city's population born in a given country (e.g. percentage Argentine) from the 2000 US Census. In order to allow for the possibility that residents of a given city travel to other cities within the same MSA, I include MSA fixed effects and cluster at the MSA level; 23 cities are not in MSAs and are dropped. Running the specifications without fixed effects leads to very similar coefficients and smaller standard errors.

In the first column of Table 2 I run the basic specification, finding that a 10% increase in population is associated with a 4% increase in the count of cuisines and that land area has no significant effect. In the second column I include squared log land area and find a significant positive coefficient on the linear term and a significant negative coefficient on the quadratic term. With log land area ranging from 14.8 to 21.4 over the 703 cities this implies a slightly positive effect for the smallest cities, then a fairly flat effect that turns negative at the 60th percentile of land area (log land of 17.6), and a fairly large negative effect for the biggest cities. This pattern is also seen in the quartile specifications (columns 3 through 6) with insignificant coefficients for the bottom three quartiles and then a large negative and significant coefficient for the largest land quartile. This is consistent with the theory which showed a small, although always negative, effect for land area that became much larger for land areas $L^* \leq L$. The coefficient on the largest quartile specification can be (noncausally) interpreted as a 10% decrease in land area, which effectively increases density without changing population, is associated with an 2.3% increase in the number of cuisines. The coefficient on population also increases with larger land quartiles, which is not an implication of the model. One possibility for this finding is that the discreteness, or lumpiness, of cuisine counts makes it easier to measure changes across larger land quartile cities, which also tend to have larger populations. This effect was present in the theoretical Zipf simulations in Figure 3, where a change in log population for cities with small populations and small land areas resulted in a very small change in the count of cuisines. For example, in the right panel of Figure 3 a 0.1 unit increase in log population for the small population city (bottom curve) at $L = .5L^*$ results in 0.788

Table 2. Number of cuisines

	(1) All	(2) All	(3) LQ4	(4) LQ3	(5) LQ2	(6) LQ1
Pop 2007-8 (logs)	0.395***	0.421***	0.199*	0.388***	0.508***	0.511***
	(0.038)	(0.037)	(0.099)	(0.082)	(0.113)	(0.045)
Land sq mtrs (logs)	0.020	1.337***	0.241*	−0.068	−0.069	−0.229***
	(0.039)	(0.218)	(0.137)	(0.154)	(0.167)	(0.056)
Squared log land		−0.038***				
		(0.006)				
Average HH size	−0.573***	−0.582***	−0.765**	−0.582*	−0.002	−0.657**
	(0.106)	(0.108)	(0.292)	(0.345)	(0.221)	(0.282)
Median HH income (logs)	0.085	0.027	0.267	−0.170	−0.952**	−0.004
	(0.147)	(0.136)	(0.425)	(0.354)	(0.420)	(0.373)
%Old (> 64)	−0.351	−0.553	−0.649	−0.610	−2.212	−1.708
	(0.794)	(0.776)	(1.640)	(1.278)	(1.426)	(2.294)
%Young (<35)	−0.443	−0.572	0.404	−1.395	−2.508*	−1.181
	(0.554)	(0.515)	(1.108)	(1.041)	(1.273)	(1.276)
%College grad	0.753***	0.792***	0.695	1.369**	1.720**	0.830
	(0.270)	(0.245)	(0.920)	(0.634)	(0.669)	(0.612)
Observations	703	703	177	172	175	179
R^2	0.854	0.863	0.846	0.921	0.901	0.948
MSA FE	YES	YES	YES	YES	YES	YES
#Clusters	74	74	44	48	46	56

Dependent variable is log count of cuisines.
Standard errors in parentheses, clustered at MSA level. *$p<0.1$ **$p<0.05$ ***$p<0.01$
Each regression is run with MSA fixed effects and 45 ethnicity control variables which match cuisine varieties.
Note: 23 Census Places are not in MSAs and were dropped from estimation.

new cuisines while a 0.1 unit increase to the large population city (top curve) at $L = 1.5L^*$ results in 3.5 new cuisines. These are the same changes in log cuisines (0.1 units since $s = 1$) but given that there are no fractional cuisines, empirically it may be easier to observe changes for larger cities.

I now turn to estimating the causal effect of population and land area on product variety. Urban economics models of city growth imply that changes in both population and land area can result from the same fundamental factors. For example, in the monocentric city model a change in the attractiveness of a city would result in population growth and an expansion of the urban fringe. If these fundamental factors are also correlated with restaurant variety then both population and land area could be endogenous. Therefore I will instrument for both population and land area to try and estimate the causal effects on variety.

Estimating the causal effect of population and land area on product variety shares identification issues with the problem of estimating the effect of population on wages and much of the following discussion is informed by Combes et al. (2011), or CDG. One potential concern is reverse causality: perhaps restaurant diversity actually increases population, similar to high wages attracting population in CDG. A related issue is that some of the factors that increase population, or population density, may also affect restaurant diversity. For example, a city with pro-density zoning laws and

land use constraints may also have a culture favorable to restaurant diversity. To address this problem I use a number of different historical and geographic variables that can predict current population and land area but that may not be associated with present-day factors affecting restaurant diversity. For the first set of instruments I follow Abel et al. (2012) who used the population of the city's county in the year 1900 as an instrument for current log density. Since I wish to estimate both population and land area separately I instrument for current population with the historic county population and use the land area of the county in 1900 as an instrument for current land area.[22] The intuition for the relevance of these instruments is that the positive correlation between historic population and current population levels may reflect the persistence generated by agglomeration while the positive correlation between historic land size and current land size could result from geography or some persistence in political boundaries. The rationale for using historic populations as an instrument is that this persistence in population is unrelated to current productivity (see the discussion in Duranton and Puga (2014), Section 5); I make a similar assumption that the persistence of both population and land area is not associated with unobservables affecting restaurant variety. While this exogeneity assumption is commonly applied to population it also serves an important role for land area. More productive cities, which could have greater restaurant variety, may expand their boundaries faster. Rusk (2006) suggests a relationship between a city's fiscal health and ability to annex neighboring land. Using historic land area would exclude the additional land from more recent productivity shocks.

In addition to county measures from 1900, I also include the 1950 Census Place population and 1950 Census Place land area from the 1952 City and County Data Books (U.S. Department of Commerce, 1952). While it is an advantage that these instruments are at the same spatial unit as my data, I am able to match fewer of my cities since City and County Data Books only have data for cities with 25,000 people in 1950. As an additional instrument I include the share of unavailable land in the principal city of each MSA from Saiz (2010).[23] The share of unavailable land is based exclusively on measures of geography (presence of water bodies, slope of land) and thus cities in areas with less developable land may be naturally constrained to higher densities. Lastly, following Glaeser and Gyourko (2005) I use the average daily temperature in January, averaged over the 30-year period from 1971 to 2000, from the 2007 County and City Data Book (U.S. Department of Commerce, 2007). The authors explain that warmer cities have experienced greater population growth since 1970 and suggest that the value of weather as an urban amenity has increased.

For each of these instruments I am able to match a different subset of the cities in my dataset and so I present a number of different regressions and re-estimate the OLS specification with the corresponding subset for most specifications. For specifications with only county-level instruments I restrict the dataset to just the cities with the largest

22 I match Census Places to counties and obtained historical population and land area data from the National Historical Geographic Information System (NHGIS, 2011).

23 Saiz calculates his measure for principal cities of MSAs with greater than 500,000 people and thus the data is at the PMSA, MSA or NECMA level. I first match Census Places to PMSAs and then for Census Places I cannot match directly to a PMSA I assign the value of the PMSA for the shared MSA. For example, Anaheim, CA cannot be matched directly to a PMSA in the Saiz data so I assign it the value of the Los Angeles PMSA since both Anaheim and Los Angeles are in the LA-Riverside MSA. For this work the MableCORR system was very helpful (MableGeocorr, 2010).

Table 3. Number of cuisines: IV specifications

	(1) IV	(2) OLS	(3) IV	(4) OLS_LQ1	(5) IV_LQ1	(6) OLS_LQ1	(7) IV_LQ1	(8) IV_LQ1
Pop 2007-8 (logs)	0.479***	0.377***	0.364***	0.465***	0.455***	0.389***	0.435***	0.497***
	(0.114)	(0.048)	(0.076)	(0.055)	(0.054)	(0.045)	(0.064)	(0.103)
Land sq mtrs (logs)	−0.185	1.475***	1.814***	−0.171***	−0.164**	−0.121**	−0.207**	−0.252*
	(0.150)	(0.382)	(0.509)	(0.052)	(0.079)	(0.047)	(0.087)	(0.137)
Squared log land		−0.041***	−0.050***					
		(0.010)	(0.014)					
Average HH Size	−0.381**	−0.410**	−0.438**	−0.399*	−0.410**	−0.023	0.013	−0.341**
	(0.174)	(0.186)	(0.170)	(0.201)	(0.174)	(0.225)	(0.185)	(0.151)
Median HH Income (logs)	−0.212	0.066	0.069	0.079	0.082	−0.442**	−0.502***	−0.095
	(0.213)	(0.190)	(0.176)	(0.150)	(0.176)	(0.165)	(0.180)	(0.170)
%Old (> 64)	−2.942**	−1.553	−1.427	−0.261	−0.201	−0.772	−2.665	−2.037
	(1.435)	(1.997)	(2.331)	(1.371)	(1.566)	(2.198)	(2.946)	(1.672)
%Young (< 35)	−1.692*	0.057	0.300	0.522	0.576	−0.575	−1.306	−0.090
	(0.966)	(1.101)	(1.350)	(0.997)	(1.017)	(1.177)	(1.450)	(0.829)
%College grad	0.869***	1.048***	1.052***	0.976***	0.971***	2.012***	2.223***	0.915***
	(0.272)	(0.321)	(0.305)	(0.281)	(0.262)	(0.553)	(0.517)	(0.313)
Ethnic HHI	−0.532*	−0.451***	−0.492**	−0.111	−0.147	−0.123	0.062	−0.133
	(0.271)	(0.154)	(0.196)	(0.237)	(0.209)	(0.248)	(0.265)	(0.277)
Observations	203	105	105	91	91	60	60	137
R^2	0.810	0.886	0.884	0.891	0.891	0.899	0.887	0.833
Instrument(s)	1,2		3,4,5		1,2		2,3	1,6,7
K-P Wald F	18.06		9.33		8.30		15.79	8.93
Stock Yogo 10% Size	7.03				7.03		7.03	13.43
Over-id *p*-value								0.19
#Clusters	71	47	47	49	49	41	41	47

Dependent variable is log count of cuisines. Standard errors in parentheses, clustered at MSA level. *$p<0.1$ **$p<0.05$ ***$p<0.01$

IV specifications instrument for log population, log land, and squared log land.

Instruments are (1) county population 1900, (2) county land 1900, (3) census place population 1950, (4) place land 1950, (5) squared land 1950 (6) average Jan temp (7) Saiz unavailable land. No specifications include MSA fixed effects.

population in the 1900 counties. There is not enough variation in my instruments to include MSA fixed effects but I do cluster standard errors at this level. I also do not have enough variation to include the earlier 45 ethnicity control variables so I instead proxy for ethnic diversity with the ethnic HHI described in section 3.1. I also include the rest of the demographic variables from Table 2 as controls with no causal interpretation. In the first column of Table 3 I show the results from instrumenting with 1900 county population and land area for 203 cities across all land quartiles (OLS results for this subset are similar to the first column of Table 2 and are omitted for space). The coefficient on population is somewhat larger than the 0.39 from Table 2 and the effect of land area is considerably larger and negative, although still insignificant. County instruments were too weak to estimate the quadratic specification but I was able to estimate this with the 1950 place-level population, log land area, and squared log land area. In column 2, I show the quadratic OLS specification for this subset, which has similar coefficients to the full sample specification from column 2 of Table 2. The IV

estimates in column 3 are slightly smaller for population and larger for land and land squared, implying that the effect of land on variety becomes negative at a larger land area and is then stronger for the largest cities. Specifications 4–8 are estimated for the largest land quartile, with 4 and 5 using county-level instruments, 6 and 7 using a mix of county and place-level instruments, and specification 8 using county and geographic/climate instruments. While the coefficients vary somewhat across samples and instruments, they are all fairly close and consistent with the OLS estimates from Table 2. In specifications 4 and 5 the IV estimates are very close to the OLS estimates while in 6 and 7 both population and land elasticities are larger in the instrumented specification. The high correlation between population and land area (0.59 in levels) makes it difficult to find more than two instruments without running into weak instrument issues. In specification 8 I use the Saiz unavailability measure, average January temperature, and 1900 county population with the instruments passing the over-identification test (OLS omitted for space). The estimates from this specification are larger but still broadly consistent with the other estimates. While it is reassuring that a different set of instruments yields roughly similar results, the K-P Wald F-statistic is fairly small for three instruments, potentially suggesting a weak instrument issue. Therefore excluding this last specification and summarizing the results suggests that the elasticity of restaurant variety with respect to city population is between 0.35 and 0.49 and the (negative) elasticity with respect to city land area, for the top quartile cities only, is between 0.16 and 0.21.

4.2. Hierarchy and cuisine level estimation

A hierarchical pattern of cuisines across cities means that cities increase their count of cuisines with rarer cuisines. This in turn implies that the average cuisine in a city with many cuisines will be found in fewer cities than the average cuisine in a low-cuisine count city. In the descriptive section of this article I documented this hierarchical relationship between cuisine count and rarity and in Table 3 I estimated the effect of population and land area on cuisine count. I now provide evidence for the effect of population and land area in generating this hierarchy by estimating the effect of these variables on cuisine rarity. For each cuisine v I count the number of cities that have that cuisine, referred to earlier as 'choice cities'. I then average this count across all the cuisines in a given city (V_m) to calculate a city-level measure of cuisine rarity:

$$R_m = (1/V_m) \sum_{v \in V_m} ChoiceCities_v \tag{4.2}$$

Lower values of R_m indicate the average cuisine in a city is less common. I estimate the same cross-city specification, Equation (4.1), but replace the log of the number of cuisines with the log of average cuisine rarity, $ln(R_m)$. The results, shown in Table 4 are consistent with those from the cuisine count estimation. In the first column of Table 4 I estimate the specification using OLS for the full set of cities, including ethnicity controls and MSA fixed effects (all specifications are clustered at the MSA level). I find that a 1% increase in population is associated with a 0.136% decrease in the average cuisine's number of choice cities—meaning an increase in rarity—and find no effect for land area. In column 3 I show the IV results, using the two county-level instruments, with column 2 showing the OLS results for the same sample. I find a stronger effect for population and again no effect for land area. Columns 4 and 5 show the OLS and IV

Table 4. Cuisine rarity

	(1) OLS	(2) OLS	(3) IV	(4) OLS_LQ1	(5) IV_LQ1
Pop 2007-8 (logs)	−0.136***	−0.148***	−0.226***	−0.185***	−0.266***
	(0.009)	(0.015)	(0.049)	(0.026)	(0.032)
Land sq mtrs (logs)	0.000	0.000	0.102	0.031	0.131***
	(0.009)	(0.016)	(0.067)	(0.026)	(0.047)
Average HH Size	0.160***	0.228***	0.183***	0.204***	0.150*
	(0.044)	(0.057)	(0.058)	(0.076)	(0.080)
Median HH Income (logs)	0.086	0.067	0.142	0.131	0.195
	(0.054)	(0.083)	(0.114)	(0.086)	(0.147)
%Old (> 64)	0.407	0.636*	1.228*	0.534	1.472
	(0.333)	(0.360)	(0.634)	(0.669)	(1.182)
%Young (< 35)	0.322	0.285	0.677	0.224	0.728
	(0.214)	(0.345)	(0.495)	(0.417)	(0.654)
%College grad	−0.377***	−0.244*	−0.293**	−0.621***	−0.633***
	(0.092)	(0.123)	(0.144)	(0.179)	(0.191)
Ethnic HHI		0.391***	0.231**	0.452***	0.217
		(0.097)	(0.109)	(0.120)	(0.143)
Observations	703	203	203	91	91
R^2	0.843	0.866	0.832	0.891	0.855
MSA FE	Yes	No	No	No	No
Ethnicity controls	Yes	No	No	No	No
Instrument(s)			1,2		1,2
K-P Wald F			13.525		12.793
Stock Yogo 10% Size			7.03		7.03
Pop sim p-val	0	0.002	0.006	0	0
Land sim p-val	0.99	0.946	0.236	0.494	0.1
#Clusters	74	71	71	49	49

Dependent variable is a measure of the rarity of a city's cuisines (lower is rarer).
This is constructed as the log of the average count of cities with a cuisine, averaged across each city's cuisines. Standard errors in parentheses, clustered at MSA level. *$p<0.1$ **$p<0.05$ ***$p<0.01$
The simulated p-values come from 1000 permutations under the null of random cuisine assignment–see text.
Instruments are (1) county population 1900, (2) county land 1900.

results for just the top land quartile. Mirroring the cuisine-count results, here I find the strongest effect for population and a significant, positive effect for land area, indicating that increasing the land area (decreasing density) decreases cuisine rarity.

In an earlier section I demonstrated that the observed pattern of cuisines across cities is far more hierarchical than patterns resulting from a simple random assignment of cuisines to cities, a data generating process which would naturally result in cities with more cuisines also having rarer cuisines. However, one might still wonder how much of the results in Table 4 are explained by this mechanical relationship between the number of cuisines and cuisine rarity. In other words, do the estimated effects of population and land area on cuisine rarity really just reflect the fact that cities with a bigger population or smaller land area have more cuisines, making Table 4 a restatement of the cuisine count regressions from Table 3? To address this potential concern I simulate this data generating process by randomly assigning cuisines to cities while maintaining the

original cuisine count. I then re-run the specifications of Table 4 to see whether the coefficients estimated on the simulated data are similar to those from the actual data. I create 1000 simulated samples and re-run all five specifications from Table 4 to obtain distributions of point estimates for the effect of population and land area on cuisine-rarity. I then compare these to the estimates on the actual data and calculate two-sided *p*-values. These *p*-values are shown in the bottom rows of Table 4. I find that the coefficients estimated from the actual data are quite different from the vast majority of those estimated on the simulated data, implying that the results shown in Table 4 are unlikely to stem from a mechanical relationship between cuisine count and rarity. Specifically, using these *p*-values leads to inferences that are mostly unchanged from the original estimates, except for that the coefficient on land area in specification 5 is now only significant at the 10% level.[24]

The theory suggests that for any variety, regardless of tastes, markets with greater populations and smaller land areas aggregate consumers and are thus more likely to support that variety. Therefore another way to evaluate the effect of population and land area on product variety is to run specifications at the cuisine level for the likelihood a given city has a specific cuisine. An advantage of conducting the analysis at this level is that I can more accurately control for ethnicity by matching cuisines with the percentage of people born in the corresponding country.[25] Another advantage is that the additional variation provided by city-cuisine pairs allows me to include fixed effects for MSAs with many Census places in IV regressions. Using the same simplifications from the previous section I can write a limited dependent variable model, where C_{mv} is an indicator variable for whether city m has cuisine v and the η_v are cuisine fixed effects:

$$Pr(C_{mv} = 1) = Pr(\Pi^*_{mv} > 0)$$
$$\Pi^*_{mv} = \gamma_1 ln(N_m) + \gamma_2 ln(L_m) + X_m{}'\beta + \eta_v + \varepsilon_{mv}$$

I estimate the above equation as a linear probability model so that I can easily run IV specifications. I use the county instruments only, which allow for the largest number of cities, and estimate OLS and IV specifications for all cuisines and ethnic cuisines only. The unit of observation is a city-cuisine pair and I run the specification for all cuisines found in three or more cities, clustering errors at the city (Census place) level. I cannot estimate the model with fixed effects for every MSA but I am able to include fixed effects for all MSAs with five or more Census places.

In columns 1 and 2 of Table 5 I compare OLS estimations with and without MSA fixed effects for the sample of 203 cities and 84 cuisines. The coefficients on population and land area are quite close across the two specifications, with the fixed effects specification having slightly larger and more significant coefficients. In the fixed effects specification a 10% increase in population increases the probability of having a given cuisine, averaged across all cuisines, by 1.3% while a 10% increase in land area

24 The distribution from the permutations is asymmetric and so I calculate two-sided *p*-values by doubling the one-sided probability of being less than my population coefficient and the one-sided probability of being greater than my land coefficient. This method is probably overly conservative; one might argue that I should be using a one-side test since I only wish to know whether a mechanical relationship could have resulted in coefficients as large as I have found.

25 In most cases there was an obvious match with a specific nationality but in some cases, such as the cuisines 'Central European' or 'Latin American', I aggregated nationalities or used the Census variables for region of birth.

Table 5. Likelihood of having a given cuisine

	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV
Pop 2007-8 (logs)	0.124***	0.130***	0.119***	0.152***	0.187***	0.184***
	(0.010)	(0.011)	(0.013)	(0.022)	(0.030)	(0.034)
Land sq mtrs (logs)	−0.019*	−0.025**	−0.032***	−0.067**	−0.108**	−0.119**
	(0.010)	(0.011)	(0.012)	(0.031)	(0.042)	(0.047)
Average HH size	−0.078***	−0.081**	−0.081**	−0.075**	−0.078**	−0.081*
	(0.029)	(0.032)	(0.032)	(0.032)	(0.040)	(0.044)
Median HH income (logs)	−0.046	−0.065	−0.063	−0.092	−0.116*	−0.102
	(0.041)	(0.042)	(0.045)	(0.056)	(0.066)	(0.070)
%Old (> 64)	−0.398*	−0.452**	−0.392*	−0.805***	−1.036***	−0.948**
	(0.232)	(0.222)	(0.202)	(0.311)	(0.362)	(0.389)
%Young (< 35)	−0.265	−0.334**	−0.290*	−0.491**	−0.708***	−0.660**
	(0.176)	(0.161)	(0.157)	(0.232)	(0.268)	(0.280)
%College grad	0.179***	0.176***	0.161**	0.201***	0.165*	0.131
	(0.063)	(0.063)	(0.062)	(0.077)	(0.087)	(0.087)
%Corresponding ethnicity			0.234*			0.224
			(0.140)			(0.139)
Observations	17052	17052	10759	17052	17052	10759
R^2	0.554	0.557	0.542	0.550	0.548	0.530
Cuisines	84	84	53	84	84	53
Cities	203	203	203	203	203	203
MSA FE	No	Yes	Yes	No	Yes	Yes
Instrument(s)				1,2	1,2	1,2
K-P Wald F				19.33	9.21	9.20
Stock Yogo 10% Size				7.03	7.03	7.03
#Clusters	203	203	203	203	203	203

Dependent variable is an indicator for cuisine, run on all cuisines found in 3 or more cities.
Standard errors in parentheses, clustered at Census Place level. *$p<0.1$ **$p<0.05$ ***$p<0.01$
Instruments are (1) county population 1900 and (2) county land 1900. Fixed effects for MSAs with 5 or more census places.

decreases this probability by 0.25%. In column 3 I run the specification for 53 ethnic cuisines only, controlling for ethnic percentage, and find similar effects for population and land area. Columns 4 and 5 show the IV results for all cuisines, with and without fixed effects. Adding the MSA fixed effects reduces the first stage statistic but it remains at an acceptable level. The coefficients on both variables are significantly larger than the OLS specifications. In the fixed effect specification a 10% increase in population increases the likelihood of having a cuisine by 1.87% and a 10% increase in land area decreases the probability by 1.1%. Column 6 shows the IV specification for ethnic cuisines only and again the coefficients are similar to those from all cuisines, with the coefficient on land area slightly larger. In the earlier cross-city IV specifications of Table 3 I showed that the overall count of cuisines was larger for cities with greater population and greater population density. The results from Table 5 show that these earlier findings do not stem from some kind of aggregation effect but also hold at the individual cuisine level, even with better controls for ethnicity. When combined with the cuisine rarity results from Table 4 the estimates from this section suggest that cities with

larger populations and smaller land areas have greater variety through the addition of rarer varieties, generating a hierarchical pattern.

4.3. Robustness exercises: clustering of ethnic populations

In previous sections I considered omitted variable bias/reverse causality and used instruments to address this possible endogeneity. However, an additional and separate endogeneity problem arises if particular groups of people associated with restaurant diversity sort into large, dense cities; this is somewhat similar to the issue of skilled workers choosing to live in denser cities, discussed in CDG. For example, if producers of cuisine *v* tend to live in large, dense cities for reasons unrelated to demand for cuisine *v*—new immigrants may be skilled restaurateurs and feel more comfortable in large, dense cities—then this would generate a relationship between population, land and product variety but the channel would on the supply side, and not demand aggregation. There could also be alternative demand-side channels: perhaps new immigrants have a particularly strong taste for their home country's food and thus restaurant variety is being driven by demand from a small subset of the population who sort into large cities and whose effect would not be captured by my ethnicity control variables. In this section I conduct a series of robustness exercises providing evidence against alternative explanations of this type.

Ideally, with a dataset on restaurateurs I could compare similar producers of cuisine *v* in different cities, or the same restaurateur living in different cities over time, and show that a producer of cuisine *v* only creates a restaurant of cuisine *v* when demand in their current city is sufficiently aggregated. Since I don't have this data, I take several approaches. First, if a small subset of the population is driving my earlier results than after controlling for the size of this population I should find no additional effect of population and land area. I will proxy for the presence of both producers and consumers of cuisine *v* with the population of ethnicity *v* and re-run the specification from Table 5. Second, I explore the effect of ethnicity in greater detail by considering the implications of geographically clustered populations with an affinity for a particular variety, as in ethnic neighborhoods. In the context of the model, this could be interpreted as concentrating all of the population of a given taste *v* into one segment of the circle, rather than having them uniformly distributed. Doing so would lower the minimum city population required for entry of that variety. Therefore I expect that increased spatial concentration of a particular ethnicity, controlling for the size of that group, will increase the likelihood a city has the corresponding cuisine. I argue that finding this relationship is more suggestive of demand aggregation than a supply-side story. A restaurant requires far more customers than employees and it is much easier for a few employees to commute to a different neighborhood than to move the customer base. From the perspective of the restaurant owner it would seem that consumer location should weigh more heavily than employee location. Of course both effects can, and probably do, exist simultaneously, but I am positing that demand aggregation is more important. However, this is just an argument and I cannot completely rule out the importance of supply-side stories, such as ethnic clustering providing a more easily accessible supply of workers skilled in the preparation of a particular cuisine.

In order to more precisely segment the population into people whose ethnicity matches a given cuisine, 'ethnic pop', and the rest of the population, 'non-ethnic pop', I use Census 2000 population levels. I use the Moran's I measure of spatial

auto-correlation to capture the general clustering/dispersion of an ethnic population in a city. Specifically, I count the number of people born in each country in each census tract with 2000 Census data and then calculate the correlation between neighboring tracts:

$$I = \left(\frac{N}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}} \right) \left(\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(X_i - \overline{X})(X_j - \overline{X})}{\sum_{i=1}^{N}(X_i - \overline{X})^2} \right) \tag{4.3}$$

In Equation (4.3) X_i and X_j represent the ethnic population in two neighboring tracts where $\overline{X}$ is the average across all tracts. The weights matrix w_{ij} is a matrix of 1's and 0's indicating whether two tracts are neighbors and N is the total number of tracts in a city. I use the 'Queen' adjacency definition—named after the chess piece's movement—so that two tracts are neighbors if they have any shared borders. I use the 'spdep' package in R to calculate this measure for each ethnicity in each city; if a city does not have any people of that ethnicity I cannot calculate the statistic and drop that city-cuisine pair from this analysis. The Moran's I is a measure of correlation and thus ranges from -1 to 1 with an expected value of $\frac{-1}{N-1}$ when there is no spatial autocorrelation. It is important to note that the expected value of Moran's I does not depend on X and thus cities with a greater percentage of people from a given ethnicity do not necessarily have a greater Moran's I, although I find that this is generally true empirically. Many of the cities in the lower land quartiles have few census tracts and so I restrict the analysis to the 182 cities of the top land quartile. Even after limiting to the top land quartile some cities have too few people from a given ethnicity to calculate the spatial distribution.

To look at how spatial concentration is correlated with ethnic cuisines I subtract the average Moran's I of cities without a particular ethnic cuisine from cities with that ethnic cuisine. I do the same for the percentage of the population from that ethnicity (ethnic percentage) and then plot the differences in Figure 6 for 55 ethnic cuisines, using a log scale for ethnic percentage to show all cuisines more clearly. Not surprisingly, the average spatial concentration of ethnic populations in cities with a cuisine is always higher than in cities without this cuisine, as is the ethnic percentage. There also appears to a strong correlation between spatial concentration and ethnic percentage.

In the first column of Table 6 I run the linear probability model on the likelihood of a city having a cuisine, pooled across ethnic cuisines with cuisine fixed effects and fixed effects for MSAs with 5 or more Census Places. Greater ethnic populations are associated with a higher likelihood of having the corresponding cuisine but I still find a large positive correlation for the remaining population and a negative correlation for land area, although this is only significant at the 10% level. In column 2 I instrument for the non-ethnic population and the city's land area with the county level historical instruments. In this specification I find large and significant effects for non-ethnic population and land area, as in earlier specifications. With the MSA fixed effects these instruments may be weak and so in column 3 I run the same specification without MSA fixed effects and find smaller but still significant coefficients. I interpret the significant negative coefficient on land area, controlling for the size of the ethnic population, as additional evidence in support of the theory. As noted earlier, the significant coefficient on population neither supports nor refutes the theory—if people from ethnicity v were

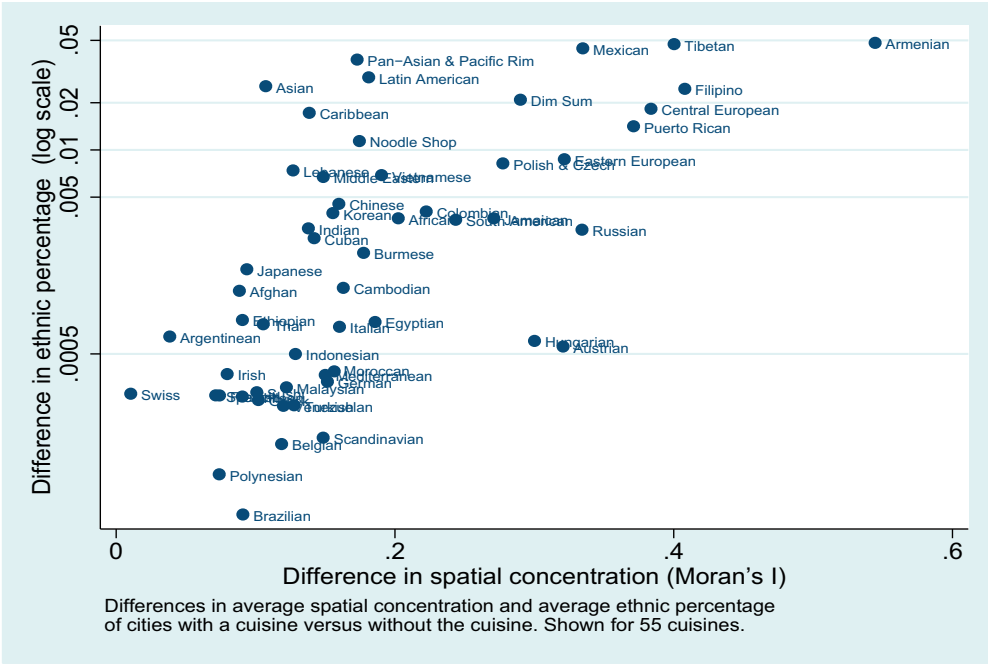


Figure 6. Spatial concentration and ethnic percentage.

the only consumers of cuisine *v* then an insignificant coefficient on the remaining population would not be inconsistent with the model. However, the finding that this remaining population is still important suggests that people of ethnicity *v* are not the only consumers of cuisine *v* and therefore ethnic populations are not completely driving the earlier results.

In column 4 I add in the Moran's I measure of spatial concentration, which reduces the sample to just the top land quartile cities in which I could calculate this measure. It can be difficult to interpret the magnitudes of this coefficient but since Moran's I ranges from −1 to 1 a rough approximation is that a 1% increase in the spatial correlation of the population (non-causally) increases the likelihood by 0.16%, a bit under half of the elasticity of the ethnic population size. In other words, bringing the ethnic population of a city closer together may achieve the same result as increasing the size of that population, consistent with the theory. In column 5 I run the same specification but instrument for the non-ethnic population and land area and find the expected signs, showing that while spatial concentration of ethnicity is certainly important it also cannot explain all of the earlier results. While columns 4 and 5 suggest an effect of spatial concentration I also wish to provide evidence that this effect comes from reduced transportation cost to a restaurant. To show the proximity of restaurants to ethnic concentrations I estimate the probability a particular census *tract* has a restaurant of cuisine *v* on the population of that tract of ethnicity *v*. I take the set of all census tracts with a restaurant for every city and pool all ethnic cuisines, removing cuisines with too few observations to be estimated, and include controls, cuisine fixed effects, and Census place fixed effects. I cluster errors at the tract level since the unit of observation is the tract-cuisine; clustering at the Census place level leads to larger standard errors that are still easily significant at the 1% level. The final column of Table 6 shows that census

Table 6. Likelihood of having a given cuisine with clustered ethnic populations

	Places					Tracts
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	IV	OLS	IV	OLS
Non-ethnic Pop 2000 (logs)	0.072***	0.129***	0.093***	0.075***	0.113***	
	(0.009)	(0.028)	(0.022)	(0.016)	(0.026)	
Land sq mtrs (logs)	−0.015*	−0.115**	−0.056*	−0.012	−0.069**	
	(0.009)	(0.047)	(0.033)	(0.016)	(0.033)	
Ethnic Pop 2000 (logs)	0.026***	0.034***	0.027***	0.038***	0.035**	
	(0.002)	(0.010)	(0.008)	(0.006)	(0.014)	
Average HH Size	−0.092***	−0.097**	−0.089***	−0.050	−0.077	−0.030***
	(0.029)	(0.043)	(0.030)	(0.044)	(0.048)	(0.001)
Median HH Income (logs)	−0.025	−0.059	−0.047	−0.042	−0.060	−0.003**
	(0.041)	(0.071)	(0.056)	(0.078)	(0.100)	(0.001)
%Old (> 64)	−0.333*	−1.026**	−0.591*	0.295	−0.285	−0.021**
	(0.195)	(0.420)	(0.342)	(0.559)	(0.817)	(0.009)
%Young (< 35)	−0.158	−0.548*	−0.269	0.192	0.006	0.035***
	(0.150)	(0.287)	(0.232)	(0.338)	(0.446)	(0.007)
%College grad	0.065	−0.004	0.071	0.379***	0.328**	0.026***
	(0.059)	(0.094)	(0.068)	(0.115)	(0.128)	(0.003)
Moran's I				0.155***	0.119**	
				(0.046)	(0.057)	
Corr. Ethic Pop (000's)						0.024***
						(0.001)
Remaining Pop (000's)						0.002***
						(0.000)
Observations	10759	10759	10759	4523	4523	954620
R^2	0.548	0.529	0.541	0.580	0.575	0.236
Cuisines	53	53	53	53	53	59
Cities	203	203	203	91	91	703
Spatial FE	MSA	MSA	None	None	None	Place
Instrument(s)		1,2	1,2		1,2	
K-P Wald F		7.07	17.09		7.98	
Stock Yogo 10% Size		7.03	7.03		7.03	
#Clusters	203	203	203	91	91	9285

Dependent variable is an indicator for cuisine, run on all ethnic cuisines found in 3 or more cities for Places regressions. Tract level regression is run on all ethnic cuisines. All specifications include cuisine fixed effects. Standard errors are clustered at Census Place level for Places regressions, Tract level for Tracts regression. *$p<0.1$; **$p<0.05$; ***$p<0.01$, MSA fixed effects for MSAs with 5 or more Census places.
Instruments are (1) county population 1900, (2) county land 1900.

tracts with larger populations of ethnicity v are more likely to have a corresponding restaurant of cuisine v.

Summarizing the results from Table 6, I find that while ethnicity, unsurprisingly, is strongly correlated with the likelihood a city has a restaurant of the corresponding cuisine, the effects of population and land area persist. If the supply of restaurateurs specializing in a cuisine is well measured by ethnic population size then the greater variety of larger, denser cities cannot be fully explained by this supply-side channel. In fact, I argue that ethnicity is more likely to affect variety through the demand aggregation channel suggested by the model. The spatial concentration of an ethnic

population, controlling for the size of the population, is strongly correlated with the likelihood a city has the corresponding cuisine. Further, there is another strong correlation between the size of a Census tract's ethnic population and the existence of a matching restaurant in that same tract. While this suggests the importance of ethnic neighborhoods it is again worth emphasizing that the land area effect persists, implying that denser cities do not have greater variety solely through ethnic neighborhoods.

5 Conclusion

In this article I have argued that a significant amount of variety in the restaurant market results from the aggregation of specific tastes from a heterogeneous population. I presented a simple model of demand aggregation where the fixed cost of opening up a restaurant combined with heterogeneous tastes led to cuisine specific entry thresholds. The importance of transportation cost in restaurant consumption implies that the spatial distribution of a population is a key determinant of whether demand passes this threshold. Cities that concentrate a large population into a small area increase the mass of consumers demanding specific varieties and thereby increase the likelihood a firm finds sufficient demand for its product. In the empirical section I found that the pattern of cuisines across US cities follows a fairly regular hierarchical distribution that is consistent with this type of threshold model and echoes findings from Central Place Theory papers studying industrial composition. Varieties that appeal to fewer people are much more likely to be found in bigger and denser cities because spatial aggregation becomes more important when demand is limited. For this reason, larger and denser cities have both greater variety and rarer varieties of restaurants. This finding was consistent across multiple specifications, including when estimated at the cuisine-level and with measures of cuisine rarity, and when using different instrumental variables to estimate the effect of population and population density. The presence of ethnic neighborhoods or other geographic clustering of tastes facilitates the spatial aggregation of demand and can significantly increase the likelihood a market supports a particular variety. Overall, the results in this article argue for a causal link between city structure and product differentiation and provide empirical evidence for the effect of demand aggregation on one of the important consumer benefits of agglomeration.

Acknowledgements

I thank the editor, Kristian Behrens, and two anonymous referees for detailed and careful comments. I also thank Nathaniel Baum-Snow, James Campbell, Tianran Dai, Tom Davidoff, Andrew Foster, Martin Goetz, Juan Carlos Gozzi, Vernon Henderson, Toru Kitagawa, Brian Knight, Sanghoon Lee, Mariano Tappata, Norov Tumennasan, seminar participants at Brown University, the Center for Economic Studies at the Census Bureau, the Sauder School at UBC, the Federal Reserve Bank of Philadelphia, Tsinghua University, City University of Hong Kong, Shanghai University of Finance and Economics, and participants at the CUERE, IIOC and UEA conferences for helpful comments.

References

Abel, J. R., Dey, I., Gabe, T. M. (2012) Productivity and the density of human capital*. *Journal of Regional Science*, 52: 562–586.

Anderson, S. P., de Palma, A., Thisse, J.-F. (1992) *Discrete Choice Theory of Product Differentiation*. Cambridge, MA: MIT Press.
Behrens, K., Robert-Nicoud, F. (2014) Survival of the fittest in cities: urbanisation and inequality. *The Economic Journal*, doi: 10.1111/ecoj.12099.
Berry, S., Waldfogel, J. (2010) Product quality and market size. *The Journal of Industrial Economics*, 58: 1–31.
Brakman, S., Heijdra, B. J. (2004) *The Monopolistic Competition Revolution in Retrospect*. Cambridge (England): Cambridge University Press.
Bresnahan, T. F., Reiss, P. C. (1991) Entry and competition in concentrated markets. *Journal of Political Economy*, 99: 977–1009.
Campbell, J. R., Hopenhayn, H. A. (2005) Market size matters. *Journal of Industrial Economics*, 53: 1–25.
Chen, Y., Rosenthal, S. (2008) Local amenities and life-cycle migration: do people move for jobs or fun? *Journal of Urban Economics*, 64: 519–537.
Christaller, W., Baskin, C. (1966) *Central Places in Southern Germany*. Englewood Cliffs, NJ: Prentice-Hall.
Combes, P., Duranton, G., Gobillon, L. (2011) The identification of agglomeration economies. *Journal of Economic Geography*, 11: 253–266.
Couture, V. (2013) Valuing the consumption benefits of urban density. University of California, Berkeley, Working Paper.
Dixit, A. K., Stiglitz, J. E. (1977) Monopolistic competition and optimum product diversity. *American Economic Review*, 67: 297–308.
Duranton, G., Puga, D. (2014) The growth of cities. *Handbook of Economic Growth*, 2B: 781–853.
Fujita, M., Krugman, P., Venables, A. J. (1999) *The Spatial Economy*. Cambridge, MA: MIT Press.
Glaeser, E. L., Gyourko, J. (2005) Urban decline and durable housing. *Journal of Political Economy*, 113: 345–375.
Glaeser, E. L., Kolko, J., Saiz, A. (2001) Consumer city. *Journal of Economic Geography*, 1: 27–50.
Handbury, J., Weinstein, D. E. (2012) Is new economic geography right? Evidence from price data. University of Pennsylvania, Working Paper.
Hsu, W.-T. (2012) Central place theory and city size distribution*. *The Economic Journal*, 122: 903–932.
Irmen, A., Thisse, J. (1998) Competition in multi-characteristics spaces: hotelling was almost right. *Journal of Economic Theory*, 78: 76–102.
Krugman, P. (1991) Increasing returns and economic geography. *The Journal of Political Economy*, 99: 483–499.
Lee, S. (2010) Ability sorting and consumer city. *Journal of Urban Economics*, 68: 20–33.
Lösch, A. (1967) *The Economics of Location*. New York: John Wiley.
MableGeocorr, (2010) MABLE/Geocorr2K: geographic correspondence engine Discussion paper, Missouri Census Data Center.
Mazzeo, M. (2002) 'Product choice and oligopoly market structure. *Rand Journal of Economics*, 33: 221–242.
Mazzolari, F., Neumark, D. (2012) Immigration and product diversity. *Journal of Population Economics*, 25: 1107–1137.
Mori, T., Nishikimi, K., Smith, T. (2008) The number-average size rule: a new empirical relationship between industrial location and city size. *Journal of Regional Science*, 48: 165–211.
Mori, T., Smith, T. E. (2011) An industrial agglomeration approach to central place and city size regularities*. *Journal of Regional Science*, 51: 694–731.
NHGIS, (2011) National historical geographic information system: version 2.0. Discussion paper, Minnesota Population Center, University of Minnesota.
Ottaviano, G., Tabuchi, T., Thisse, J.-F. (2002) Agglomeration and trade revisited. *International Economic Review*, 409–435.
Rusk, D. (2006) *Annexation and the Fiscal Fate of Cities*. Brookings Institution Metropolitan Policy Program.
Saiz, A. (2010) The geographic determinants of housing supply. *Quarterly Journal of Economics*, 125: 1253–1296.

Salop, S. C. (1979) Monopolistic competition with outside goods. *Bell Journal of Economics*, 10: 141–156.
Syverson, C. (2004) Market structure and productivity: a concrete example. *Journal of Political Economy*, 112: 1181–1222.
Tabuchi, T. (2009) Hotelling's spatial competition reconsidered. *CIRJE F-Series*.
U.S. Department of Commerce, (1952) *County and City Data Book: 1952*. U.S. Department of Commerce, Bureau of the Census, ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.
U.S. Department of Commerce, (2007) *County and City Data Book: 2007*. U.S. Department of Commerce, Bureau of the Census.
Waldfogel, J. (2008) The median voter and the median consumer: local private goods and population composition. *Journal of Urban Economics*, 63: 567–582.

Appendix

A. Multiple firms

With multiple firms there are three possible equilibria, which I label following Salop (1979): an equilibria where firms compete with neighbors for consumers ('competitive'), an equilibria where firms are local monopolists but do not have a large enough market extent to set monopoly prices ('kinked'), and an equilibria where each firm operates as a local monopolist in its market ('monopoly'). The equilibria are determined by the distance from the firm to the indifferent consumer, who may now be indifferent between the reserve good or the good of a competing firm. Let d_m represent the distance to a consumer indifferent between the product and the reserve good, d_c the distance when a consumer is indifferent between two neighboring firms setting equal prices in equilibrium, and n is the number of firms:

$$d_m = \frac{u_1 - p_m}{\tau} \tag{A.1}$$

$$d_c = \frac{L}{2n} \tag{A.2}$$

The land area in the market determines the price a firm sets, which in turn determines the distance to the indifferent consumer. The competitive equilibrium has a symmetric Nash equilibrium with prices as a best response to competing firms while in the monopoly equilibrium the firm chooses a price to maximize profit given the reserve good. The kinked equilibrium price is found by equating $d_m = d_c$ and solving for price.

$$p_c = c + \frac{\tau L}{n} \tag{A.3}$$

$$p_k = u_1 - \frac{\tau L}{2n} \tag{A.4}$$

$$p_m = \frac{u_1 + c}{2} \tag{A.5}$$

When $p_c < p_k$ then we must be in the competitive equilibrium since any firm charging p_k would be undercut. Once land expands to $L = \frac{2n}{3} L^*$ then p_k becomes smaller than p_c and we are in the kinked equilibrium; if the firm tried to charge p_c the consumer would

consume their reserve good. The kinked equilibrium case is very similar to the minimum entry condition for the single firm in that the firm would like to charge the monopolist's price but is constrained in geographic extent, this time by the presence of the other firms. In fact, the kinked equilibrium frontier collapses to the minimum entry condition when $n=1$. Finally, at $L = nL^*$ there will be exactly n local monopolists charging the monopoly price of p_m. For $L > nL^*$ there will be a monopoly equilibrium with n firms but with gaps between the geographic market extents.[26]

$$N_{min}(n, L) = \begin{cases} \dfrac{N^* L^* n^2}{L} & \text{if } L \leq \dfrac{2n}{3} L^*, \text{ 'full coverage, competitive equilibrium'} \\ \dfrac{2N^* L^* n^2}{2nL^* - L} & \text{if } \dfrac{2n}{3} L^* < L \leq nL^*, \text{ 'full coverage, kinked equilibrium'} \\ \dfrac{2N^* L}{L^*} & \text{if } nL^* < L, \text{ 'partial coverage, local monopoly equilibrium'} \end{cases} \tag{A.6}$$

The focus of this article is on the extensive margin—when does a market have a given variety—but I draw the frontiers for 2 and 3 firms to show how the minimum conditions derived hold for markets with multiple firms. For markets with large land areas and populations just below the frontier, such as a market with $L = 3L^*$ and $N = 6N^* - \varepsilon$, the model suggests that a small increase in population would allow the market to go from having zero firms to three firms. What this shows is that while the count of firms can change dramatically with small changes in population or land area, the extensive margin is still governed by the minimum population frontier.

B. Hierarchy test of Mori et al. (2008)

Following MNS, I formally define a hierarchical structure of cuisines across cities as a distribution such that if a cuisine is found in a city with n cuisines it will also be found in all cities with greater than n cuisines. Consider the matrix of cuisine–city indicators, with cuisines sorted by ascending city choice count and cities by ascending numbers of cuisines, as seen in the plot in the left panel of Figure 5. For each cuisine–city pair, MNS define a hierarchy event as a binary variable equal to one only when all cities with greater or *equal* number of cuisines also have that cuisine. Graphically, for each dot a hierarchy event occurs if all other slots in the same row to the right of the dot are filled. As an example, if 725 of the cities have cuisine v and the city missing cuisine v has the fewest overall number of cuisines (furthest to the left) then a hierarchy event occurs for all cuisine–city pairs in cuisine v's row. However, if instead the city missing the cuisine was the city with the most overall cuisines (furthest to the right), then there would not be a hierarchy event for any of the cuisine–city pairs. It is important to note that if two cities have the same number of overall cuisines but one of the cities is missing cuisine v then neither city has a hierarchy event for cuisine v. MNS define the hierarchy share as

26 If there are n firms in the market and $L > nL^*$ then the profit to the firms as monopolists with gaps is higher than the kinked profit and thus charging the price $p_k < p_m$ is not an equilibrium strategy.

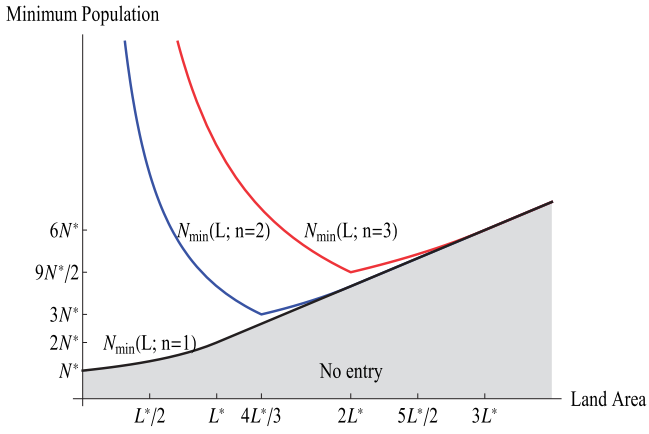


Figure B1. Minimum market conditions with multiple firms.

the percentage of cuisine–city pairs that are hierarchy events, or graphically, the percentage of dots that are hierarchy events.

Having defined the hierarchy share, I then test the null hypothesis that cuisines are randomly assigned to cities, with the number of cuisines in each city fixed and equal to the observed count. In other words, a city with just one type of cuisine should be no more likely to have a Chinese restaurant than an Afghan restaurant. Rather than deriving the distribution of the hierarchy-share statistic, I will run a simulation to estimate the cumulative distribution function of hierarchy share and then accept or reject the null hypothesis of no hierarchy. The procedure is for each city m draw $\#v_m$ varieties from the total set V and then calculate the hierarchy-share. I repeat this procedure 10,000 times to get a distribution of hierarchy-share under the null hypothesis. When I calculate the hierarchy share for all 726 places I find a hierarchy share of 23% that is highly significant (the largest value from 10,000 simulated runs was 2.9%). For comparison, MNS found that the distribution of industries in Japan in 1999 had a hierarchy share of 71%, much larger than the share I find here. Nonetheless, the hierarchical structure of cuisines across cities is still very far from random and is rather surprising considering that cities differ dramatically in ethnic composition, income, family size, education and other factors that may affect the cuisines offered.

C. Derivation of empirical specification from theory

I start with Equation (2.12), which gives an ordinal relationship between population, land area and the number of varieties in a market. To incorporate differences in the characteristics of cities I assume the population of a market N_m can be completely partitioned into a set of groups S: $N_m = \sum_{s=1}^{S} N_{ms}$. These groups could be defined by demographic characteristics, such as ethnicity, income, or education, and group specific affinities for cuisines implies their relative size could affect a city's variety:

$$\#Varieties_m \sim N_m\left(\sum_{s=1}^{S}\frac{N_{ms}}{N_m}\right)\left[\left(1-\frac{L_m}{2L^*}\right)\mathbf{1}(L_m < L^*)+\frac{L^*}{2L_m}\mathbf{1}(L_m \geq L^*)\right] \qquad \text{(C.1)}$$

I then take logs and replacing the log of the percentage sub-groups ($\frac{N_{ms}}{N_m}$) with the vector X_m of demographic percentages and population characteristics (e.g. median income, percent college educated) shown in Table 1. With summary level data I cannot separate individuals into mutually exclusive groups and so I let each characteristic have its own coefficient. After simplifying there are two terms with land area: $ln(2L^* - L_m)\mathbf{1}(L_m < L^*)$ and $-ln(L_m)\mathbf{1}(L_m \geq L^*)$. The form of the first term is dependent upon functional form assumptions from the model, such as linear transportation cost. However, the prediction of a negative effect of log land area, and one that increases in magnitude with larger values of land, is a more general finding and would also hold, for example, under an assumption of quadratic transportation costs. An interesting feature of the model is that the kink point L^* is independent of the variety and the demographic characteristics of the population. This kink point exists because until the firm can reach the desired geographic market extent, the density adjusts with the land area in a way dependent upon transportation cost. When the market's land area is larger than L^* the firm is able to operate at its preferred geographic extent ($L^*/2$ on either side) and thus only needs the market to be above a constant density so that it can always sell to the same number of consumers within this distance. In this way the kink point is a very general feature of these models and implies that different assumptions about transportation cost will still lead to land area having a larger negative effect for geographically bigger markets.[27] While I am not structurally estimating the model, it could still be informative to look for a kink point in the data by allowing L^* to be an unknown structural break. However, the data is too sparse in that many geographically small markets lack many cuisines, and thus a cut-off value in land area perfectly predicts the absence of that cuisine. Therefore I take the approach mentioned in the main text and replace the kinked function with the log of land area, $ln(L_m)$, and allow for non-linearity by running regressions separately by land quartile or with a quadratic term in land area. These simplifications lead to the straight-forward log linear specification used in estimation:

$$\ln(\#Cuisines_m) = \gamma_0 + \gamma_1 ln(N_m) + \gamma_2 ln(L_m) + X_m'\beta + \varepsilon_m \tag{C.2}$$

27 In the quadratic case variety v is found in market m if $N_m\left(\sum_{s=1}^{S}\frac{N_{ms}}{N_m}\right)\left[\left(1-\frac{L_m^2}{3L^{*2}}\right)\mathbf{1}(L_m \leq L^*) + \frac{3L^*}{2L_m}\mathbf{1}(L^* < L_m)\right] \geq N^*$, where $L^* = \sqrt{\frac{4}{3\tau}(u_1 - c)}$ and N^*, which by definition does not depend upon transportation cost, is the same ($N^* = \frac{F}{u_1 - c}$).

D. Count of cities with each cuisine, by land quartile

Table follows.

Table D1. Count of cities with each cuisine, by land quartile

	1	2	3	4	Total
Afghan	9	4	2	2	17
African	20	4	2	1	27
American (New)	119	68	45	21	253
American (Traditional)	173	153	128	91	545
Argentinean	10	1	1	1	13
Armenian	2	0	0	0	2
Asian	128	90	63	41	322
Austrian	1	0	0	1	2
Bagels	65	32	11	12	120
Bakeries	152	112	83	57	404
Barbecue	170	151	121	65	507
Belgian	7	0	0	0	7
Brazilian	21	8	1	4	34
Burmese	2	1	1	0	4
Cajun & Creole	81	30	31	6	148
Californian	70	32	19	13	134
Cambodian	4	0	1	0	5
Caribbean	59	17	14	6	96
Central European	5	2	0	0	7
Cheesesteak	13	1	2	3	19
Chilean	3	0	0	0	3
Chinese	178	175	169	148	670
Chowder	4	2	0	0	6
Coffee Shops & Diners	153	106	72	42	373
Coffeehouse	143	115	81	54	393
Colombian	9	1	0	0	10
Cuban	30	7	2	4	43
Deli	171	163	147	119	600
Desserts	162	137	107	76	482
Dim Sum	2	1	0	1	4
Donuts	37	10	5	9	61
Eastern European	13	0	4	4	21
Eclectic & International	80	40	21	14	155
Egyptian	3	0	0	0	3
English	24	5	5	3	37
Ethiopian	11	0	1	2	14
Family Fare	127	100	67	46	340
Fast Food	178	178	168	160	684
Filipino	6	1	1	1	9
French	111	77	52	32	272
German	58	17	6	4	85
Greek	114	53	43	17	227
Hamburgers	126	73	51	30	280
Health Food	54	6	5	3	68
Hot Dogs	32	12	8	3	55
Total	2940	1985	1540	1096	7561

(continued)

Table D1. Continued

	1	2	3	4	Total
Hungarian	3	1	0	0	4
Ice Cream	154	118	87	58	417
Indian	106	56	39	21	222
Indonesian	10	0	0	0	10
Irish	39	17	8	6	70
Italian	170	147	120	91	528
Jamaican	6	0	1	0	7
Japanese	148	118	82	59	407
Juices & Smoothies	35	12	9	2	58
Korean	62	21	15	10	108
Kosher	15	3	0	3	21
Latin American	31	5	5	3	44
Lebanese	18	2	2	1	23
Malaysian	8	3	3	0	14
Meat-and-Three	4	0	0	0	4
Mediterranean	52	20	7	10	89
Mexican	179	174	153	123	629
Middle Eastern	70	21	16	8	115
Moroccan	9	0	0	1	10
Noodle Shop	21	7	6	5	39
Pan-Asian & Pacific Rim	39	14	11	9	73
Pizza	180	178	178	164	700
Polish & Czech	3	1	0	0	4
Polynesian	6	0	3	0	9
Portuguese	3	0	1	0	4
Puerto Rican	3	0	0	0	3
Russian	5	0	0	0	5
Scandinavian	2	2	0	0	4
Seafood	161	119	104	64	448
Soups	107	74	52	27	260
South American	17	2	2	1	22
Southern & Soul	67	8	4	4	83
Southwestern	33	6	3	3	45
Spanish	37	12	7	2	58
Steakhouse	147	79	59	35	320
Sushi	47	11	9	6	73
Swiss	9	0	0	0	9
Tapas / Small Plates	24	5	1	2	32
Thai	130	91	72	47	340
Tibetan	2	0	1	0	3
Turkish	7	0	0	0	7
Vegan	22	1	3	3	29
Vegetarian	73	31	16	12	132
Venezuelan	4	0	0	0	4
Vietnamese	87	35	26	12	160
Total	2355	1394	1105	792	5646