The Location of Sales Offices and the Attraction of Cities

Thomas J. Holmes*†

September 24, 2004

Abstract

This paper examines how manufacturers locate sales offices across cities. Sales office costs are assumed to have four components: a fixed cost, a "frictional" cost for out-of-town sales, a cost-reducing "knowledge spillover" related to city size, and an idiosyncratic "match quality" for each firm-city pair. A simple theoretical model is developed and is estimated using data from the Census of Wholesale Trade. The factors emphasized in home-market effect literature, namely fixed costs and frictional costs, are found to play an important role in location decisions. Match quality also matters. The results for knowledge spillovers are mixed.

*University of Minnesota, Federal Reserve Bank of Minneapolis, and NBER.

†I acknowledge support from NSF grant SES-9906087 for funding this project. I thank the Center for Economic Studies at the U.S. Census Bureau for making it possible for me to access the Census data. I thank the editor and the referees for comments that led to a major revision. I thank Zvi Eckstein, Gautam Gowrisankaran, Joe Harrington, Shawn Klimek, Sam Kortum, Ariel Pakes, Arnie Resnik, and Jim Schmitz for detailed comments or help. I thank Adam Copeland and Junichi Suzuki for research assistance. I am grateful for helpful comments from numerous seminar participants. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis, the Federal Reserve System, or the National Science Foundation.

1 Introduction

Sales offices are the home bases of company representatives who call on customers to mediate sales. Sales offices are highly concentrated in large cities. This paper asks why this is so. Two important explanations in the literature for the attractiveness of large cities are the "home market effect," and "knowledge spillovers." For the sales office sector, it is reasonable to suppose that both effects matter. In many contexts, it is hard to distinguish these factors from each other. The model developed and estimated here with micro data allows these two factors to be separately identified. The key to identification is the behavior of large firms with multiple sales offices and how this behavior varies with firm size.

The model has five main ingredients. First, a frictional cost is incurred when a manufacturer mediates sales to a city without having a local sales office in the city. This is intended to capture the travel cost (including the time cost) for a representative to make a sales call on an out-of-town customer. The home-market effect literature has emphasized the transportation cost of moving goods. Surely this must be dwarfed by the cost of moving people. A salesperson based in Chicago calling on a client in New York for a one hour meeting incurs a time cost of at least a full working day.

Second, there is a fixed cost to setting up an office at any location. As such, the firm has an incentive to limit the number of locations where it opens an office. On account of the frictional cost, it is advantageous to allocate the limited number of office to the largest cities; i.e., those with the largest "home market." This interaction between scale economies and transportation cost in the model captures the essence of the home-market effect. See Fujita, Krugman, and Venables (1999) and Fujita and Thisse (2002) for recent textbook treatments of this literature.

Third, there is a systematic relationship between productivity and city size. A large literature emphasizes the role of cities in facilitating the diffusion of "knowledge spillovers."[1] There is reason to believe this factor matters for sales offices. A sales representative's job is to match the needs of customers with the products of the firm; information is the essence of the

[1] Recent work includes includes Eaton and Eckstein (1997), Black and Henderson (1999), Glaeser (1999), and Lucas (2001).

job. A salesperson needs to know the market, not just the product line carried by his or her firm, but also the products offered by competing as well as complementary firms. In a large city, this kind of information likely spills over from contacts with others. Beyond knowledge spillovers, there may be other advantages to larger cities (higher quality workers) or even disadvantages (higher rents). All these other various factors—apart from the home-market factor—are netted out in the model in a single parameter called the knowledge-spillover term.

Fourth, there are random, firm-specific factors that make some cities a good "match" for a firm and other cities a bad match. Without this heterogeneity, all firms of a given size act the same way. In particular, all firms with one office would place it in the largest city. This is inconsistent with the data. With the match component in the model, firms sometimes choose smaller locations over larger locations, trading off the benefits of a larger home market for the benefit of a better match. The matching factor is a force of dispersion in the model.

Fifth, firms vary exogenously in size. There are manufacturers like Kraft Foods that are large, and others like Tom's Widgets that are small, for reasons outside the model.

In this environment, large and small firms face fundamentally different problems. Large firms have the scale economies to open a vast network of many offices; small firms perhaps have only one. Differences in the number of offices lead to differences in the geographic distribution of sales office activity. To understand the nature of these differences, consider first what happens when there are no matching considerations in the model. Then if a firm has only one office, it goes in the largest city; if it has a second office it goes in the second largest city, and so on. Here, increases in firm size shift the distribution of sales office activity away from the largest cities. Next consider what happens when match considerations are important. A small firm can't do much about reducing frictional costs. Since it will have just a few offices to work with and since even the largest cities are only a small portion of the national market, a small firm incurs frictional costs on the vast majority of its customers, regardless of what it does. In contrast, a large firm, with potentially dozens of offices, is in the position to substantially reduce frictional costs. As a result, larger firms end up putting more weight on frictional costs, and less weight on matching costs, as compared to smaller firms. This makes larger firms less dispersed, and in particular, less heavily concentrated in the smallest cities, as compared to smaller firms.

The empirical portion of the paper examines establishment-level data on manufacturers sales offices from the Census of Wholesale Trade. The analysis determines how the geographic distribution of sales office activity varies with firm size. The main finding is that as firm size increases, the distribution of activity shifts away from the smallest cities. Moreover, excluding the smallest firms, an increase in firm size also shifts the distribution away from the largest cities. Thus the distribution shifts away from the extremes of city size towards the cities in between. This is an implication of the model with the home-market effect and matching ingredients. A version of the model with only the knowledge-spillover ingredient has no particular implication for how location behavior varies with firm size. If knowledge spillovers make salespeople more productive, then presumably this benefits large firms and small firms alike.

The structural model is estimated by fitting moments for how sales vary by city size and firm size and how the number of offices varies by firm size. The parsimonious model fits the moments reasonable well. In the estimated model, the home-market effect ingredients—frictional costs and fixed costs—as well as the matching ingredient all play substantial roles as location factors. The evidence for knowledge spillovers is mixed. In one estimated model, the knowledge parameter is essentially zero; in another the parameter plays some role, albeit a small one.

While the theoretical literature on the home market effect is substantial, there has been relatively little empirical work that quantifies its importance as a location factor. Previous work, including Rosenthal and Strange (2001), consists of cross-industry studies that attempt to find proxies of variables like scale economies, transportation costs, and knowledge spillovers and then relate differences in geographic concentration of industries to differences in these proxies. A challenge with this approach is that it is difficult to come up with reliable proxies. This paper pursues an alternative approach that makes no attempt to directly measure these variables. Rather, the approach is to infer the parameters from the revealed choice behavior of firms of difference sizes.[2]

[2] Another empirical line of work on the home-market effect by Davis and Weinstein (1996, 1999) examines

The rest of the paper proceeds as follows. Section 2 provides background information about sales offices and show sales office activity concentrates in large cities. Section 3 develops the theory. Section 4 presents the data analysis. Section 5 concludes.

2 Sales Offices and Cities

This section explains what sales offices are and shows they are concentrated in large cities to a remarkable degree. Part 1 of this section presents the Census definition of sales offices and provides summary statistics about sales offices from the Census. Part 2 provides context by discussing the sales office operations of several large and familiar companies. Part 3 uses Census data to document the relationship between sales office location and city size.

2.1 What are Sales Offices?

The Census Bureau defines *manufacturers sales offices and branches* as wholesaling establishments that sell "products manufactured or mined in the United States by their parent company."[3] These are distinct from *merchant wholesalers* who sell goods manufactured by some other firm, handling the goods and taking title in the process. A third distinct category is *agents and brokers* who sell products made by others, but don't handle or take title to the merchandise. These three categories together make up Sector 42, "Wholesale Trade," in the *North American Classification System* (NAICS). Within the category *manufacturers sales offices and branches*, an *office* is an establishment that carries zero inventory and does not handle the goods. These establishments tend to be located in suites in office buildings. A *branch* carries some inventory and includes distribution centers.

In the 1997 Economic Census (U.S. Bureau of the Census (2001)), there were 29,305 establishments in the category of manufacturers sales office and branches. These accounted for $1.3 trillion dollars in sales, $46 billion in payroll, and just under one million employees. Of these totals, offices accounted for 61 percent of the sales and branches the rest. In

how production structure varies with market size. Their approach is quite different from mine. The key to identification here—the role of large firms—plays no part in their analysis.

[3] U.S. Bureau of the Census (2000, p. 5).

what follows, I use the abbreviated term *sales offices* to be short for the entire category of *manufacturers sales offices and sales branches.*

2.2 The Sales Office Operations of Some Familiar Companies

A clearer understanding of what sales offices are and how they fit into the operations of a firm can be gained by discussing the operations of some example companies. This subsection provides such a discussion based on data obtained from publicly available sources, such as business directories, phone books, and the internet.[4] We will see that sales offices of these example firms are highly concentrated in large cities, the same pattern that will be later shown to hold in the broader Census data.

To begin, consider the case of Kraft Foods. The company describes itself as "the largest branded food and beverage company headquartered in the United States." It manufactures and distributes brands such as Nabisco, Oscar Mayer meats, Maxwell House coffee, and Kraft dairy products. The company's facilities can be classified into three groups: *manufacturing plants*, of which there are 72, *administrative facilities,* such as corporate headquarters, divisional headquarters, and research facilities, of which there are 9, and sales offices, of which there are 249.

A good idea of what these hundreds of sales offices do can be obtained from the job description of "sales representative" posted at the company's web site: "As a Sales Representative, you will be responsible for distributing, selling, promoting, and merchandising Kraft Foods products...You'll be responsible for executing company promotions, for meeting inventory needs, and for monitoring the competitive activity within your region...Most of our sales people represent our entire product line."

The locations of the Kraft's facilities are mapped in Figure 1. Sales offices are marked with circles, manufacturing plants with triangles, and administrative facilities with squares. The map also illustrates the distribution of population, with one small gray dot for each 1,000 in people. The map shows that the 256 sales offices form a national network that

[4] The business directories used included ReferenceUSA (a product of InfoUSA) and the Harris InfoSource's Selectory Online. This information along with other web-based information was processed by hand. The data is available on request. See Holmes (2004) for further dicussion about its compilation.

concentrates in the various centers of population. In contrast, the manufacturing facilities (which mostly process food) are regional, concentrating in the Midwest, California, and New York agricultural areas.

Table 1 shows the relationship between sales office location and city size for Kraft and nine other familiar companies. The city size category with the largest cities—over eight million in population—includes three cities: New York, Los Angeles, and Chicago. (Cities are defined to be Metropolitan Statistical Areas (MSAs) or Consolidated Metropolitan Areas (CMAs); see the appendix for details.) The next category, two to eight million, has 19 cities, and includes cities like Portland and Tampa at the bottom of the population scale and Washington D.C. and Philadelphia at the top. There are 57 cities in the .5-2 million size class and 194 in the under .5 million class (the smallest city, Enid, Oklahoma has a population of 57,000). For each company and each city size class, the table reports the fraction of cities that have a sales office for the company. For all the companies, the probability of having a sales office in the smallest cities is quite low. For virtually all the firms, the probability of an office in the largest cities is one; i.e., virtually all have New York, Los Angeles, and Chicago sales offices. There is a strict monotone relationship between probability of an office and city size for all the firms.

Table 1 makes clear that larger cities are more likely to have sales offices in an *absolute* sense. What about a *relative* sense? New York is 350 times larger in population than Enid, OK so even if sales offices were allocated across cities randomly in proportion to population, New York could be expected to end up with more. To control for such relative population differences, consider the following simple statistical model that is in the spirit of the Ellison and Glaeser (1997) dartboard analysis. Suppose that for a particular company i, the probability city j *does not* have a sales office is given by

$$prob_{ij}(\text{no office}) = \lambda_i^{n_j^\alpha} \tag{1}$$

for $\lambda_i \in (0,1)$ and $\alpha \geq 1$, and n_j the population of city j. Assume this random event is drawn i.i.d. across cities. Observe that λ_i is the probability of no office for firm i in a city of unit size, $n = 1$. Think about a unit size city as getting one potential draw for an office and and a n-size city getting n i.i.d. draws. Then if $\alpha = 1$, the probability a city of size

n would *not* get an office is λ_i^n, the probability of missing on all n draws. If $\alpha > 1$, cities scale up in their attractiveness more than in proportion to population. The last column of Table 1 reports the maximum likelihood estimates of α estimated separately for each company. (The separate estimate of λ_i for each company is not reported.) The estimate of α is greater than one for each company and in most cases it is substantially greater than one. If α is constrained to be the same for all companies and the model is estimated jointly, the estimate is $\hat{\alpha} = 1.42$ with a standard error of .05. The interpretation is that when city size increases 100 percent, its "attractiveness" to sales offices increases at rate of 142 percent.

One issue that could be raised at this point is that of the joint location of sales offices and other facilities. One might speculate that if sales offices tend to co-locate with other facilities of the firm and if the other facilities of the firm tend to be located in larger cities, then the connection between sales office location and city size may be spurious. In the appendix I address this issue by extending the statistical model to allow sales office locations to depend upon the locations of a firm's other facilities. I draw the following conclusions. First, controlling for the locations of these other facilities makes no difference for the estimate of α. Second, the connection between sales office locations and manufacturing plants is negligible. This is not surprising since manufacturing activity is so different from sales office activity. Third, there is a positive connection in the co-location of sales office locations and administrative locations that is statistically significant. This is not surprising given the plausibility of some overlap between these white collar jobs. However, the importance of co-location in explaining the location of sales offices is not large. The number of sales offices across the ten companies, at 799, is an order of magnitude greater then the 69 administrative facilities. This limits the extent of possible geographic connection. The overwhelming majority of sales offices for this set of companies (85 percent) are located in MSAs in which the parent company has no administrative facilities. This point can been graphically in Figure 1, where it is clear that the co-location of sales offices with other facilities is negligible.

2.3 The Census Data

I now turn to the data on the 29,305 sales offices in the 1997 Census of Wholesale Trade. The data contains information about sales, employment, payroll, operating expenses, and inventories, among other things. Table 2 shows how these measures of sales office activity vary with city size.[5] The table uses the same city-size groupings as in Table 1 and it also adds a column for activity in nonmetropolitan areas. The top part of the table reports per capita measures. For these, total activity (e.g. total sales) across all sales offices in the geographic grouping is divided by the total population of the geographic grouping. The bottom half of Table 2 reports the same group of measures as in the top of Table 2, but only renormalized. I take the ratio of the per capita measure in city size class to the per capita measure in the United States as a whole. This ratio is commonly called a *Location Quotient* (LQ). When activity is proportional to population, the location quotient is one everywhere. When activity increases more than proportionately with population, it is less than one for small cities and greater than one for large cities.

Table 2 reports that the total sales of offices located in nonmetropolitan areas is $620 per person living in nonmetropolitan areas. Per capita sales increase to $2,700 for small cities (under a half million in population) and to $4,920 for cities in the one-half to two million category. Sales rise all the way up to $7,980 and $6,860 for the two largest city-size categories, more than a tenfold increase compared to nonmetropolitan areas. The sales location quotient for nonmetro areas is only 0.11. The LQ increases with city size all the way up to 1.57 and 1.34 for the largest two size classes. The other measures of sales-office activity reveal a similar pattern. Payroll per capita and operating expenses per capita both increase by a factor of ten, going from the smallest to the largest city-size categories. Employment and inventories also increase, by a factor of five rather than ten.

The significant concentration of office activity in large cities is a recurrent feature in earlier Census years. Table 3 takes the cross section of MSAs and reports the results of a simple regression of the log of sales on the log of MSA population.[6] The population elasticity

[5] This table was constructed with published Census data. Other tables in this paper were constructed with data that can only be accessed at one of the seven regional Research Data Centers operated by the Census. The appendix discusses this further.

[6] MSA definitions change from Census year to Census year. With one exception, the regessions in Table 3 fix MSA definitions equal to their 1987 county-equivalant definitions as set forth by the Census. The exception is that the regression with 1997 data with controls for city characteristics uses the 1997 MSA definitions to maintain comparability with the later analysis.

(the slope of the regression line) for sales ranges from 1.63 in 1987 to 1.71 in 1997. This is roughly comparable with the elasticity of 1.42 obtained with the count data on the ten example firms. Table 3 also reports a regression of differences in log sales on differences in log population between 1997 and 1982. The estimate from this "fixed-city effect, fixed-time effect" regression is 1.80. Thus the pattern that relative sales-office activity increases with size holds within cities over time as well as across cities.

Table 3 also reports the population elasticity when additional city characteristics are included in the regression. These characteristics include a measure of education level of the workforce, a measure of airport access, and a measure of manufacturing activity (see the appendix for details), all of which we would expect to be associated with higher sales-office activity in a city. The additional variables do play some role in the regression, raising the R^2 for 1997 from 0.75 to 0.82, and they lower the population elasticity from 1.71 to 1.56. While lower, the population elasticity remains quite high. I conclude that no matter how I cut the data, sales offices are heavily concentrated in large cities.

3 The Theory

The first part of this section describes the environment. The second part characterizes the solution to the firm's problem and determines how firm behavior varies with firm size.

3.1 The Environment

There are J cities ordered by population size, $n_1 < n_2 < ... < n_J$, and total population is normalized to unity $\sum_{j=1}^{J} n_j = 1$. There are I firms, ordered by total sales, $q_1 < q_2 < ... < q_I$, and each firm's sales are distributed across cities in proportion to population. Hence sales of firm i in city j are $q_{ij} = q_i n_j$.

The cost structure of the sales operation has three components: selling costs, frictional costs, and fixed costs. The fixed cost ϕ is paid per sales office. The frictional cost $\tau \geq 0$ is paid per unit sold outside the city where the servicing sales office is located. Each of these costs is the same for all firms and all cities. The selling cost c_{ij} is paid per unit and varies with the firm i and the city j where the sales office is located. Assume that

$$c_{ij} = \overline{c} - \gamma n_j + \varepsilon_{ij}. \tag{2}$$

Aside from the constant $\overline{c}$, this has two components. The component $-\gamma n_j$ allows selling costs to vary in a systematic way with city size. If $\gamma > 0$, then larger cities have a cost advantage. External knowledge spillovers of the type emphasized by Lucas are one example. If $\gamma < 0$ were to hold, larger cities have a cost disadvantage, perhaps because rents are higher. In the analysis I assume that $\gamma \geq 0$ and I refer to γ as the *knowledge spillover* parameter. The final term $\varepsilon_{i,j}$ is an idiosyncratic component of cost that varies across firms and cities as I now explain.

The $\varepsilon_{i,j}$ term captures the idea that for various reasons outside of the model, a particular firm might find a particular city to be a good "match" for locating a sales office. For example, perhaps a firm is looking for salespeople with a unique set of skills. If a particular city happens to be the home of such a unique individual, ceteris paribus, this city is a good place to set up an office. Or perhaps a firm has its headquarters in a particular city and there is some complementarity in locating a sales office near headquarters. It is intuitive that in a larger city, there is a better chance that a particular firm would be able to find a unique talent and a better chance that the firm would have other administrative facilities in the area, since the firm would be drawing from a larger pool. To capture this notion, the value of the match term $\varepsilon_{i,j}$ is assumed to be the minimum of Nn_j draws from a distribution $F(x)$ for some scaling parameter N. With this cost structure, a city that is twice as large gets twice as many idiosyncratic draws x, and has twice the chance of finding a rare firm-specific talent.

It is convenient to assume that the random $\tilde{x}$ are drawn from the double exponential distribution used in the logit model,

$$pr(\tilde{x} \geq x) = 1 - F(x) = e^{-e^x}. \tag{3}$$

In this case, the distribution of the first-order statistic remains double exponential, i.e.,

$$pr(\varepsilon_{i,j} \geq x) = e^{-n_j N e^x}. \tag{4}$$

I normalize $N = 1$.

3.2 The Firm's Problem and Solution

Firm i takes as given its total sales q_i, the distribution of its sales $(n_1, n_2, ...n_J)$ across cities, and its vector of match draws $(\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iJ})$. It chooses a set of office locations and an allocation of servicing activity across offices. The firm's objective is to minimize the sum of selling costs plus frictional costs plus fixed costs.

In a solution to the firm's problem, let $B_i \subseteq \{1, 2, ...J\}$ denote the set of locations where firm i places an office. The variable intermediation cost to firm i for servicing sales at location k from an office $j \neq k$ is $c_{ij} + \tau$. This cost does not depend on k, the location being serviced. Hence the firm has one *export location* j_i^*, the office with the lowest selling cost, i.e.,

$$j_i^* = \arg\min_{j \in B_i} \{c_{ij}\} = \arg\min_{j \in B_i} \{\bar{c} - \gamma n_j + \varepsilon_{ij}\}.$$

Suppose a location $j \neq j_i^*$ also has an office. Then the following condition must hold,

$$\phi + c_{ij} q_i n_j \leq (c_{ij^*} + \tau)\, q_i n_j\,. \tag{5}$$

The left-hand side of (5) is the cost of a local office at city j; it equals the fixed cost plus the selling cost at j times sales $q_i n_j$ at j. The right-hand side is the cost of servicing j from the export location; the fixed cost of ϕ is avoided and the variable cost is the export location's selling cost plus the the frictional cost.

Let *sales office activity* of firm i in city j be denoted s_{ij} and define it as total servicing activity undertaken at city j. (This is zero if there is no office at j.) The location quotient for firm i at location j is that location's share of national sales office activity divided by that location's share of population,

$$LQ_{ij} = \frac{\frac{s_{ij}}{\sum_{k=1}^{J} s_{ik}}}{n_j}. \tag{6}$$

Let $\overline{LQ}_{ij}$ be defined as the expected location quotient for firm i at location j, where the expectation is taken over the match vector $(\varepsilon_{i1}, \varepsilon_{i,2}, ...\varepsilon_{i,J})$.

I begin my analysis of the firm's problem by looking at extreme values of firm size for which complete solutions can be obtained.

3.2.1 Limiting Case 1: The Very Small Firm

Suppose for a firm of size q that

$$qn_J\tau < \phi. \tag{7}$$

Under this assumption, the firm is small enough that the maximum possible savings in frictional cost from opening an office (the fictional cost of servicing the largest city) is less then the fixed cost. If condition (7) holds, there is single office in the optimal configuration. Denote this the case of the *very small firm.*

The firm's objective is to minimize average total cost. When the firm has a single office and puts the office at city j, the average total cost is

$$\begin{aligned} ATC_j &= c_j + (1-n_j)\,\tau + \frac{\phi}{q} \\ &= (\bar{c} - \gamma n_j + \varepsilon_j) + (1-n_j)\tau + \frac{\phi}{q} \\ &= \left[\bar{c} + \tau + \frac{\phi}{q}\right] - (\gamma + \tau)\,n_j + \varepsilon_j. \end{aligned} \tag{8}$$

The first term in the first line is the selling cost per unit. The second term is frictional costs. These are incurred on all sales, except for the local sales n_j of the office in j. The third term is average fixed cost. The second line substitutes in equation (2) for c_j. Rearranging terms, the third line expresses average total cost as a constant, a term that depends upon n_j and a random term. The firm picks the location that minimizes ATC_j.

Equation (8) for average total cost highlights a fundamental identification problem faced in this paper. The sum of γ and τ enters multiplicatively with city size n_j. Higher values of τ and γ increase the relative advantage of large cities in the same way. (τ also affects the constant but that has no effect on behavior.) If firms always chose only one location, there would be no way to separately identify γ from τ. But large firms tend to open more than one office, and, as we will see, this opens up a window for identification.

3.2.2 Limiting Case 2: The Very Large Firm

Suppose for a firm of size q that

$$qn_1\tau > \phi. \tag{9}$$

Under this assumption, the firm is large enough that the minimum possible savings in frictional cost from opening an office (the fictional cost of servicing the smallest city) exceeds the fixed cost. Call this the case of the *very large firm*. In this case, it is always optimal to open an office at the location with the lowest selling cost and use this as the export location j^*,

$$j^* = \arg \min_{j \in \{1,2,...J\}} \{c_j\} . \tag{10}$$

Dividing condition (5) through by sales qn_j at j, there is an office at a location $j \neq j^*$ if and only if the average total cost is lower,

$$\frac{\phi}{qn_j} + c_j \leq c_{j^*} + \tau.$$

Substituting in $c_k = \overline{c} - \gamma n_k + \varepsilon_k$, and rearranging, the condition for an office at j is

$$\frac{\phi}{qn_j} + \gamma \left(n_{j^*} - n_j\right) \leq \varepsilon_{j^*} - \varepsilon_j + \tau. \tag{11}$$

Notice that γ is multiplied by the difference in city size, but τ is not. This is unlike the small firm's problem in (8) where only the sum $\gamma + \tau$ matters . The crucial difference between the two problems is that, to open at j, the large firm's problem involves moving *only* sales at j, while the small firm's problem involves moving the servicing location for total sales across all locations. Consequently, γ and τ can be separately identified by the behavior of large firms.

Using (10) and (11) and the assumption (3) of a double exponential distribution, it is straightforward to derive analytic expressions for the probability that an office at location j services location k. (A related probability is derived in Eaton and Kortum (2002).) A formula for the location quotient $\overline{LQ}_j$ at j can then be derived. The formula is reported in the appendix and is used to prove the following comparative statics result.

Proposition 1. For large enough firm size q, the expected location quotient $\overline{LQ}_J(q)$ in the largest city strictly decreases in q.

Proof. See appendix.

To understand the intuition, take a fixed draw $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_J)$ of the match components and consider how firm behavior varies as a function of size q. For large q, the firm always opens an export office at the lowest cost location j^*. Whether or not it opens a local office in another city j depends upon condition (11). The condition is more likely to be satisfied the larger is q and the larger is the city size n_j. As firm size increases, it becomes cost efficient to open up offices in smaller cities, shifting the distribution of sales office activity away from the largest city.

3.2.3 General Firm Size

If the firm is neither very small nor vary large, an analytic expression for the expected location quotient appears unattainable. It is straightforward, however, to solve numerical examples on a computer In this subsection, I present a particular numerical example and discuss the generality of the findings.

In the example, the city population distribution equals the empirical size distribution of MSAs for 1997 used in the empirical analysis. The knowledge spillover parameter γ is set to zero. The frictional cost τ is set equal to

$$\tau = E\left[\sum_{j=1}^{J} n_j \varepsilon_j\right] - E\left[\min_j \{\varepsilon_1, \varepsilon_2, ..\varepsilon_J\}\right]. \tag{12}$$

This is the difference in expected selling cost per unit between servicing all sales with a local office and servicing all sales with the best match location. This is a measure of the importance of the matching. With τ equal to this, there is a sense that the example contains frictional cost and matching in equal measure, at least for a large firm for which it might be feasible to open an office in every city. Figure 1 plots $\overline{LQ}_J(q)$ and $\overline{LQ}_1(q)$, the expected location quotients in the largest and smallest cities, as a function of firm size q.

Note first that $\overline{LQ}_J(q)$ decreases in q for large enough q. This has to be true from Proposition 1. But observe that $\overline{LQ}_J(q)$ first increases before decreasing in q. For the very small firm with only one office, the primary job of the office is necessarily export activity. This is true even if the office is located in the largest city since the population share n_J of the largest city (New York in this case) is only 9.4 percent. The savings in frictional cost from locating in the largest city applies to 9.4 percent of sales; any saving on matching cost applies to all sales. Hence, given the value (12) of frictional cost chosen, minimization of selling cost is the primary location factor for the very small firm, and the location of the single office is usually not the largest city. If firm size increases to give the firm sufficient scale economies to open a second office, it is relatively likely to placed in the largest city, if the first office is not already there. This follows because the second office services only local sales, not exports, so frictional cost reduction is a relatively more important location factor. These "second" offices in city J account for why $\overline{LQ}_J(q)$ is initially increasing.

If τ is extremely large or if γ is extremely large, then the argument just made doesn't apply. In either case, the very small firm will put its single office in the largest city, so concentration there can only go down.. But if neither parameter is large, I find in other example simulations that $\overline{LQ}_J(q)$ is first flat in q (the region where there is a single office), and then increases in q (where second offices are placed in the largest city) before it eventually decreases.

Note next that $\overline{LQ}_1(q)$ sharply decreases in firm size q. This is surprising since this happens on a range of q where $\overline{LQ}_J(q)$ also decreases. In this range, expected sales office activity is being shifted from the largest and smallest cities towards the cities in between. To get a rough understanding of the forces underlying result, consider a very small firm with one office. If it were ever to open an office in the smallest city, virtually all of sales activity would be export activity, since the population of the city itself is negligible (it is $n_1 = .0003$). If the firm's size were to increase enabling it to add local offices in other cities, the (relative) export activity of the office in the smallest city would decline, pulling down its location quotient. An potential offsetting positive effect is that, in the event that some other city is the export location, an increase in firm size might lead the firm to open a local office, just to service customers in that city. However, this latter positive effect isn't relevant for the smallest city because its negligible size means the savings in frictional cost from a local office are swamped by the fixed cost and the bad matches usually found in tiny cities. This effect of replacing imports with a local office *does* matter for medium-sized cities, accounting for why the location quotients of intermediate-sized cities increases in firm size, while the location quotients in the smallest and largest cities decline.

If τ is extremely large or if γ is extremely large, then the argument just made doesn't apply. In either case, the very small firm will not put its single office in the smallest city, so concentration there can only go up.. But if neither parameter is large, the pattern for $\overline{LQ}_1(q)$ is like that illustrated in Figure 1. My formal result is

Proposition 2. Assume $\gamma = 0$, $n_J \leq .5$ and $n_1 = 0$. For small enough τ, there is a range of q where $\overline{LQ}_1(q)$ strictly decreases.

Proof. See appendix.

4 The Empirical Analysis

This section takes the theory to the data. Part one shows that the qualitative patterns found in the model economy are also found in the data. Part two shows that the theory is a quantitative success—the model is estimated and it fits the data reasonably well.

4.1 Location and Firm Size

The data are from the 1997 Census of Wholesale trade discussed in Section 2. I make adjustments to the data to make it consistent with the theoretical model. In the model, a firm has at most a single office at a given location. In the data, there are firms with multiple offices in the same MSA. I handle this by aggregating the establishments of the same firm within the same MSA into a single office. The fixed cost in the model is best interpreted as the cost to open the first sales office in the city (with additional offices having zero marginal cost). I also exclude sales offices outside of MSAs; these account for a negligible amount of sales office activity (2.5 percent of total sales). Excluding offices outside of MSAs reduces the total number of establishments from 29,305 to 26,629 and then aggregating establishments of the same firm in the same MSA results in 19,711 offices for 3,786 firms.

Table 4 presents summary statistics and cell counts by firm-size class. Firm size is defined by aggregating sales across MSAs. The size classes range from "Under 25 million" in sales at the bottom to "Over one billion" at the top. The last column of the table shows that the average number of offices increases sharply with firm size, starting at 1.7 offices in the bottom category and rising to 28.4 offices in the top category. This is an immediate implication of the theory and this relationship will be incorporated in the estimation below.

Table 5 presents location quotients by firm-size class and city-size class. The MSAs are grouped into the same size classes as in Section 2. To construct Panel A, I used the raw data from the 1997 Census to calculate the location quotient for each firm in each city size class according to (6) and then took unweighted means across firms by city size class and firm size class. Panels B and C uses an alterative procedure to summarize the relationship that includes controls for industry. The procedure treats each million dollars of sales as an observation and classifies it by firm size and city size categories. It estimates a multinomial logit model for the allocation of the sales units across the city size classes conditioned upon firm size class and industry.[7] The industry definitions are at the four-digit NAICS level for 1997 and the three-digit SIC level for 1992.[8] The estimated sales shares by city size and firm size are evaluated at the means of the industry dummy variables and are converted into location quotients by dividing through by population shares. Panel B shows the results for 1997. Panel C reports the results from applying the same procedure with the earlier 1992 Census data and 1992 MSA definitions..

The are four notable features of the location quotients in Table 5. First, fixing firm-size class, location quotients increase in city size, i.e., when moving from left to right along a row. Second, looking at the column for the largest cities (the right-most column), for large enough firm size, the location quotient decreases in firm size; i.e., far enough down the column it is decreasing. This is the implication of Proposition 1. Third, there is some evidence that for the largest city class, the location quotient first increases before decreasing. This is the same pattern as in Figure 2 and it holds in the model if γ and τ are not too large. In panel A and C, there are sizeable increases between the first and second firm-size categories: 1.27 to 1.47 and 1.42 to 1.58. However, the increase in panel B is negligible.[9] Fourth, looking at the column with the smallest cities (the left-most), the location quotient decreases in firm size. This is an implication of the model if γ and τ are not too large. For example, in Panel

[7] The logit procedure also incorporated non-MSA sales activity, which, as mentioned above, is negligible. Non-MSA locations were included as a fifth city-size class. In Panels B and C, the LQs for non-MSA areas are not reported and the LQs that are reported do not include non-MSA population when calculating population shares. I do this to be consistent with Panel A and with the later results. Holmes (2004) reports panels B and C with the non-MSA locations included and the results are similar.

[8] There are 18 different such industries for each classification. A major switch from the SIC to the NAICS classification occured between the 1992 and 1996 Censuses.

[9] In Holmes (2004), I also report the results of a logit model with 1997 data where firm size is defined by number of offices. For firms with one office the LQ in the largest city size class is 1.46 and this increases to 2.27 for firms with 2-5 offices.

A, the location quotient falls monotonically from .76 for the smallest firm-size class down to .32 for the largest firm-size class.

It should be emphasized that the features just described are characteristics of the complete census of the universe of firms, not of a sample. As can be seen in Table 4, in each firm size class there are thousands of offices and hundreds of firms or more underlying the statistics that are reported. Note also that the quantitative results for the years 1992 and 1997 in panels B and C are very similar, despite the substantial entry and exit and reallocation of sales across firms that occurs between a five year census period.

It is worth nothing that the pattern for the largest city size classes and larger firms can also be seen in the earlier Table 1 from the discussion of the ten example large companies. The companies are sorted by total number of sales offices. The table suggests a hierarchy where offices are placed in the largest cities first.

4.2 Structural Estimation

The model is estimated with population size and the other observable city characteristics that made a difference in the regression analysis of Table 3. In particular, the selling cost of firm i in city j depends upon three additional characteristics,

$$c_{ij} = \bar{c} - \gamma n_j - \eta_1 z_{1,j} - \eta_2 z_{2,j} - \eta_3 z_{3,j} + \varepsilon_{ij}, \qquad (13)$$

where characteristic $z_{1,j}$ is the education level in city j, $z_{2,j}$ is airport access, and $z_{3,j}$ is the level of manufacturing activity. The constant term $\bar{c}$ is normalized to be zero since changing it does not affect any choices. The parameters to be estimated are the frictional cost τ, the systematic city size effect γ, the fixed cost ϕ, and the coefficients η_1, η_2, η_3 on the city-specific characteristics. Let $\theta = (\phi, \tau, \gamma, \eta_1, \eta_2, \eta_3)$ denote the parameter vector.

It simplifies computation to discretize firm size. I use the six sales size categories from Table 4 and assume that the sales q_h of each firm in a given size category $h \in \{1, 2, ..., 6\}$ are equal to the mean sales of firms in the size category.

The estimation procedure is simulated method of moments. There are three sets of moments. The first set consists of the 6*4=24 location quotients by firm-size class and city size-class in panel A of Table 5. Let $\overline{LQ}_k^h(\theta)$ be the expected location quotient in city-size class k and firm-size class h, given the parameter vector θ. The second set, six moments, are the number of offices per firm by firm-size class, from the last column of Table 4. Let $\overline{O}^h(\theta)$ be the expected number of offices. For scaling, the actual number of offices is divided by $\overline{O}^h(\theta)$, so the expectation of the normalized variable is one. The third set, three moments, are correlations between the three city characteristics (e.g. airport access, etc.) in city j and the residuals (the realized location quotient in the city aggregated across all firm size levels less the predicted level).[10] The expected values of these three correlations are zero. There are a total of 33 moments. The estimate $\hat{\theta}$ minimizes the sum of the squared deviation of the moments from their expectations.

Because analytic expressions for the expectations $\overline{LQ}_k^h(\theta)$ and $\overline{O}^h(\theta)$ are not available, I use simulation to approximate them. Specifically, I draw the match vector $(\varepsilon_1, \varepsilon_2, ...\varepsilon_J)$ 100,000 times, solve the firm's problem for each ε and each size class h, and then average over the 100,000 draws.

Before discussing the parameter estimates, it is useful to first discuss the magnitude of the cost variation due to the match term ε_{ij}. Just as in a standard logit model, a normalization is required, since when the ε_{ij} are scaled up by a constant factor and τ, γ, and ϕ all change in the same proportion, all decisions remain the same. Hence, the normalization $N = 1$ in (4). Table 6 presents statistics about the distribution of ε_{ij}, under the normalization. The variance *within* a city is constant at 1.6, independent of city size. This is a little less than half of the overall variance of 3.8, meaning the variation *across* cities is substantial. The expected value of ε_{ij} falls from 5.1 for the city at the first (population weighed) quartile of the city size distribution and falls to 2.9 for the city at the third quartile, a differential of 2.2. The expected value of the lowest match component for a firm, $\min_j\{\varepsilon_{ij}\}$, is $-.6$. The expected difference between the mean of ε_{ij} and this minimum is 4.6. (This difference is the expression on right-hand side of (12) and is the value of τ used to construct figure 2.)

Table 7 reports the parameter estimates for two specifications.[11] Model 1 zeros out city

[10] Sales data aggregated across firm size classes is published for the majority of cities. (This public data is used to construct Table 2.) For this third set of moments, I use the cities with public data, so that, along with panel A of Table 5, all of the data moments are public, so the estimation procedure can be replicated outside the Census.

[11] Non-differentiability of the objective function precluded the use of a gradient-type method for optimiza-

characteristics other than population (i.e., $\eta_1 = \eta_2 = \eta_3 = 0$) and excludes the corresponding moments. Model 2 is the full specification.

The estimate of τ is approximately 3.8 in both models and is precisely estimated. To interpret this magnitude, it is useful to compare it to the magnitude of matching cost. There are two interesting points to be made about this comparison.

First, the relative importance of frictional costs versus matching depends upon city size. For the smallest city, the expected difference between its match draw and the best match draw is $7.7 - (-.6) = 8.3$, much higher than $\tau = 3.8$. So the savings in frictional cost from opening in the smallest city would be more than offset by the higher cost of a (typically poor) bad match in the smallest city. In contrast, for the largest city, the same difference is $1.8 - (-.6) = 2.4$, which is much less than τ. Matches are typically good in the biggest city, so there is less of a tradeoff between enjoying a good match and saving on frictional costs.

Second, the relative importance of frictional costs and matching also depends upon firm size. Suppose a firm is so large that fixed costs are irrelevant and consider two strategies. Strategy one is to service sales in each location out of a local office. Strategy two is to open a single office, putting it in the best match location (the lowest ε_{ij}) and using this to service sales in all locations. Strategy one raises expected match cost by 4.6, but reduces frictional cost by anywhere from $(1 - n_J)\tau = 3.5$ to $(1 - n_1)\tau \approx 3.8$, depending on where the single office with strategy two is located. On net, strategy two is preferred to strategy one. But for this large firm, the tradeoff between better matching and savings on frictional cost is a close call.

Next suppose a firm is so small that it chooses a single office. The analog of strategy one above is to locate the office in city J to minimize frictional costs. Strategy two, as before, is to place a single office in the city with the lowest matching cost. Strategy one raises expected matching costs by $1.8 - (-.6) = 2.4$, and reduces frictional costs by at most $(n_J - n_1)\tau = .36$. Hence, the savings on frictional cost is small compared with the increase in matching costs, since even the largest city is only a small portion of the total population. Thus reducing frictional costs is relatively unimportant for a very small firm. This is the

tion. Instead, I used a simplex-type method called the *amoeba method*. A bootstrap procedure was used to approximate standard errors.

fundamental reason why, in the model, smaller firms are more concentrated in the smallest cities.

The estimate of the fixed cost ϕ is 1.5 in both models and is precisely estimated. To interpret the magnitude, it is useful to relate it to the population shares of the smallest and largest cities, $n_1 = 0.0003$ and $n_J = 0.094$. The smallest firm type has sales of $q_1 = 7.5$, as reported in Table 4.. If the smallest firm were to locate an office in the largest city, the savings in out-of-town costs would be $\tau q_1 n_J = 3.8 \times 7.5 \times 0.094 = 2.7$. This exceeds the fixed cost of $\phi = 1.5$. This doesn't mean a small firm will always open an office in the largest city, because it has to consider selling costs as well. The population of the smallest city is so tiny that, for the small firm, the fixed cost obviously swamps savings in the frictional cost. For the largest firm type, the savings in frictional cost from locating in the smallest city is $\tau q_6 n_1 = 5.6$, more than three times as large as the fixed cost of opening the office. If these were the only considerations, the largest firms would always open an office in the smallest city. But again, firms also take selling costs into consideration.

The estimates of the knowledge spillover parameter γ have standard errors in the range of .7 to 1.4 which are high compared to the standard errors on the τ and ϕ estimates. The point estimates vary from .15 in Model 1 to 5.9 in Model 2. Consider the magnitude of the high estimate. The cost difference due to knowledge spillovers between the largest and smallest cities is $(n_J - n_1)\gamma \approx .1 \times 5.9 = .6$. This is swamped by the expected difference in match quality, which from Table 7 equals $7.7 - 1.8 = 5.9$.

In Model 2, the coefficient estimates for the airport and manufacturing city characteristics variables are positive as expected. The coefficient on the education variable is negative, but the magnitude is close to zero. Table A1 in the appendix shows how the city characteristics vary across city-size classes. A notable feature of that table is that the manufacturing activity measure tends to decrease with city size. Using the coefficient estimate on this characteristic, the change in average selling cost between the top city size class (8 million plus) and the bottom (under half million) is .28.[12] The change attributable to the knowledge spillover component is -.44. Thus the manufacturing activity component to some extent

[12]The change attributable to changes in the education level is .08 and to changes in the airport variable is -.37.

offsets the knowledge spillover component and this may account for why the estimate of γ increases when the additional city characteristics are included. In any case, the average cost differences from spillovers and other characteristics across city size classes are quite small relative to the differences in average matching quality.

The estimated model economy fits the data well, considering it is highly stylized and has only a few parameters. Panels B and C of Table 8 report the moments of the model economy data for Model 1 and Model 2. For comparison purposes, panel A repeats the moments reported earlier from the actual data. The estimated models in B and C are qualitatively like the data and the numerical values are in many instances fairly close. Each row is increasing from left to right. The first column is decreasing from top to bottom; the second is U-shaped, the third is increasing top to bottom; the last has an inverted U-shape.

5 Conclusion

This paper develops a model in which firms can choose multiple locations for sales offices. There are two main theoretical results. First, the concentration of sales office activity in the *largest* cities decreases in firm size for large enough firms. Second, if the frictional cost and knowledge spillover parameters are not too large, then the concentration of sales office activity in the *smallest* cities also decreases in firm size.

Using micro data on sales offices from the Census of Wholesale Trade, I find that the implications of the theory are salient features of the data. After first normalizing the match quality parameters, I estimate the remaining parameters of the model. While there are multiple estimates for the knowledge spillover parameter, even with the high-end estimate, the parameter is unimportant relative to matching considerations. The estimates of the fixed cost and frictional cost parameters are precise. The fixed cost matters for the smallest firms in the data but is of negligible importance for the largest firms. The savings in frictional cost from having a local office is approximately equal to the savings in matching costs from using the best match location rather than the average.

The analysis has a number of limitations. First, the model assumes that all firms, large and small, sell to a national market and that the distribution of sales across cities is the same for all. In reality, some firms are regional. Second, the model assumes that all firms are constrained to mediate sales through an internal sales operation. In reality, firms are sometimes able to outsource this activity through independent agents. Third, the modeling of frictional costs is crude. In a richer model, the frictional cost would increase with distance rather than be flat. I expect that the tensions at work in the simpler model considered here would continue to hold in a more general model incorporating these features.

The paper has implications beyond the sales office sector. Sales office activity is but one example of the white collar, information-oriented work, that is highly concentrated in large cities. (See Holmes and Stevens (forthcoming) for a recent accounting.) The frictional costs found to exist for the sales office sector are likely to matter for this information-oriented work more generally. The case for extending the results is easiest to make for the remaining components of the wholesaling sector besides manufacturers sales offices, namely, merchant wholesalers and agents and brokers, because the activities are very similar; e.g., the making of sales calls. It is more of a stretch to apply the results to other sectors like finance and business services that are different from sales offices in many ways. But like sales offices, these other sectors rely heavily on face-to-face contact. The potential frictional costs of providing these face-to-face contacts may be just as important for these sectors as it is for sales offices.

The problem of a manufacturer allocating sales offices across cities is analogous to the problem of a multi-national firm allocating local affiliates across countries (Helpman, Melitz and Yeaple (2004)). Also related is how a firm's export decision depends upon the size of the destination country (Eaton, Kortum, and Kramarz (2004).) This trade literature has not so far incorporated matching considerations. Since matching seems to matter for sales offices, perhaps it also matters in these trade contexts as well.

Appendix

A. Notes for Section 2

Definition of Cities

I use MSAs as defined by the 1997 Economic Census and the population figures are from the geographical file that accompanies this data. In cases where MSAs are combined into a *consolidated* MSA (CMSA), I use the consolidated entity. For example, New York is an aggregation of 15 *primary* MSAs (PMSAs), including, for example, the Newark, New Jersey, PMSA, the Danbury, Connecticut, PMSA, as well as, of course, the New York, New York, PMSA. With these aggregations, there are 273 different metropolitan areas in the data.

Sales Offices and Other Facilities

The issue of the joint location of sales offices and other facilities was raised in the text. To address this issue, I extend the earlier statistical model to allow for the sales office location probability to depend upon the presence other facilities. Suppose the λ parameter in (1) takes the functional form

$$\lambda_{ij} = \frac{\exp(\theta_i - \beta^M y_{ij}^M - \beta^A y_{ij}^A)}{\exp(\theta - \beta^M y_{ij}^M - \beta^A y_{ij}^A) + 1},$$

where y_{ij}^M and y_{ij}^A are the number of manufacturing and administrative establishments that firm i has in city j. If $\beta^M = 0$ and $\beta^A = 0$, things reduce to what we had before. If $\beta^A > 0$, then, holding everything else including city size fixed, an increase in the number of local administrative facilities increases the likelihood a city will have a sales office (since the probability $\lambda_{ij}^{\alpha^{n_j}}$ of *not* getting one decreases).

Allowing for this more general structure has virtually no effect on the point estimate of α (the estimate is 1.43 rather than 1.42). The estimate of β^M is -.34, with a standard error of .20, making it barely statistically significant. The estimate of β^A is .94, with a standard error of .26. To make sense of the magnitudes, I evaluate the estimated probability of an office at the mean value of city size n and the mean value of θ_i, starting with $y^M = y^A = 0$. Adding one manufacturing plant reduces the probability of an office from .248 to .193. Adding one administrative facility increases the probability from .248 to .440. Thus, the probability of having a sales office in a city goes up if the firm also has an administrative office in a city. But again, as emphasized in the text, the number of administrative facilities is small compared to the number of offices and 85 percent of all sales offices are in cities without an administrative office.

Controls for Additional City Characteristics

The variables used in the regression with additional controls are defined as follows. The education measure is the fraction of workers 25 years and older with a bachelor's, graduate, or professional degree in the MSA in 1990. The source is the U.S. Bureau of the Census (1996). The airport variable is domestic enplanements in 1999 per person. The source is the U.S. Bureau of Transportation Statistics (2000). The manufacturing intensity measure is sales of manufacturing plants per person. The source is the 1997 Economic Census (U.S. Bureau of the Census (2001)). Table A1 shows a cross tabulation of these three variables.

Accessing the Census Data

Three kinds of Census data were used in the project. First, raw micro data were examined in the Census Bureau's Center for Economics Studies site in Suitland, Maryland. Second, moments from the micro data were constructed that were released by Census official after a disclosure review. These moments were subsequently analyzed outside the Census. Third, the project used data from published tabulations. Table 2 uses only published tabulations. Tables 3 through 6 use the raw micro data. Table 7, the structural estimates, takes as input the disclosed moments (Table 5) as well published tabulations of sales by MSA (used for estimating the coefficients on city characteristics). This data and the programs to estimate the model are available from the author.

B. Notes for Section 2

Proof of Proposition 1 (sketch)

The first step in the proof is to derive an analytic formula for $\overline{LQ}_j$ in the large firm case (9). Let $p_{j,k}$ be the probability location j services location k. Rewriting (11), location j supplies itself if and only if

$$\frac{\phi}{qn_j} + \varepsilon_j - \gamma n_j \leq \min_{\ell \neq j}(\varepsilon_\ell - \gamma n_\ell + \tau).$$

Using the standard logit arguments, the probability of this event is

$$p_{j,j} = \frac{n_j e^{\gamma n_j - \frac{\phi}{q n_j}}}{\sum_{\ell \neq j} n_\ell e^{\gamma n_\ell - \tau} + n_j e^{\gamma n_j - \frac{\phi}{q n_j}}}$$

Analogously, the probability location j services k is

$$p_{j,k} = \frac{n_j e^{\gamma n_j}}{\sum_{\ell \neq k,} n_\ell e^{\gamma n_\ell} + n_k e^{\gamma n_k + \tau - \frac{\phi}{q n_k}}}.$$

The expected location quotient at j is

$$\overline{LQ}_j = p_{j,j} \frac{n_j}{n_j} + \sum_{k \neq j} p_{j,k} \frac{n_k}{n_j}. \tag{14}$$

The second step in the proof is to write $\overline{LQ}_J(\frac{\phi}{q})$ as a function of $\frac{\phi}{q}$. Straightforward calculations show that $\overline{LQ}'_J(\frac{\phi}{q}) > 0$ when evaluated at $\frac{\phi}{q} = 0$ which proves the claim.

Proof of Proposition 2 (Sketch)

Using the analysis of the small firm case, the limiting location quotient for small enough j is

$$\lim_{q \to 0} \overline{LQ}_j = \frac{1}{n_j} p_j = \frac{1}{n_j} \frac{n_j e^{-\tau n_j}}{\sum_{k=1}^{J} n_k e^{-\tau n_k}} = \frac{e^{-\tau n_j}}{\sum_{k=1}^{J} n_k e^{-\tau n_k}},$$

where p_j is the probability city j is the location of the single office and $\gamma = 0$ is imposed. For a very large firm, taking q to infinity, the limiting location quotient from formula (14) is

$$\lim_{q \to \infty} \overline{LQ}_j = \frac{n_j}{n_j + (1 - n_j)\, e^{-\tau}} + \sum_{k \neq j}^{J} \frac{n_k}{n_k e^{\tau} + (1 - n_k)}. \tag{15}$$

Using $n_1 = 0$, and taking inverses of these limits, it is sufficient to show that when $n_J \leq .5$,

$$G(\tau) \equiv \frac{1}{\sum_{k=2}^{J} \frac{n_k}{n_k e^{\tau} + (1 - n_k)}} - \sum_{k=2}^{J} n_k e^{-\tau n_k}$$

is strictly positive for small τ. Straightforward calculations show $G(0) = 0$ and $G'(0) = 0$. Furthermore, $G''(0) > 0$ if and only if

$$H \equiv \sum_{k=2}^{J} n_k^2 + 2 \left[\sum_{k=2}^{J} n_k^2 \right]^2 - 3 \sum_{k=2}^{J} n_k^3$$

is strictly positive. In the separate notes available on the web, Holmes (2004), it is shown that $n_J < .5$ and $n_j \leq n_{j+1}$ implies $H > 0$, which proves the claim. *Q.E.D.*

References

Black, Duncan, and Vernon Henderson (1999), "A Theory of Urban Growth," *Journal of Political Economy* 107 (April): 252–84.

Bresnahan, Timothy F., and Peter C. Reiss (1991), "Entry and Competition in Concentrated Markets," *Journal of Political Economy* 99 (October): 977–1009.

Davis, Donald R., and David E. Weinstein (1996), "Does Economic Geography Matter for International Specialization?" Working Paper 5706, National Bureau of Economic Research.

__________ (1999), "Economic Geography and Regional Production Structure: An Empirical Investigation," *European Economic Review* 43(2): 379–407..

Eaton, Jonathan, and Zvi Eckstein (1997), "Cities and Growth: Theory and Evidence From France and Japan," *Regional Science and Urban Economics* 27: 443–74.

Eaton, Jonathan, and Samuel S. Kortum (2002), "Technology, Geography, and Trade," *Econometrica,* 70(5), 2002, 1741-1779.

Eaton, Jonathan, Samuel Kortum and Francis Kramarz (2004), "An Anatomy of International Trade: Evidence from French Firms," working paper.

Fujita, Masahisa; Paul Krugman; and Anthony J. Venables (1999), *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, Mass.: MIT Press.

Fujita, M., and Thisse, J.F. (2002) *Economics of Agglomeration: Cities, Industrial Location and Regional Growth.* Cambridge: Cambridge University Press.

Ellison, Glenn and Edward L. Glaeser (1997), "Geographic Concentration in U.S. Manufacturing Industies: A Dartboard Approach," *Journal of Political Economy* 105: 889-927..

Glaeser, Edward L. (1999), "Learning in Cities," *Journal of Urban Economics* 46 (September): 254–77.

Glaeser, Edward L., and David C. Mare (2001), "Cities and Skills," *Journal of Labor Economics* 19 (April): 316–42.

Henderson, J. Vernon (1986), "Efficiency of Resource Use and City Size," *Journal of Urban Economics* 19: 47–70.

Helpman, Elhanan, Marc J. Melitz and Stephen R. Yeaple, "Export versus FDI with Heterogeneous Firms," *American Economic Review*, 2004, vol. 94, 300-316

Holmes, Thomas .J. (2004) Notes for "Sales Offices and the Attraction of Cities." Available at http://www.econ.umn.edu/~holmes/papers/separate_appendix_sales_offices.pdf, 2004.

Holmes, Thomas J. and John J. Stevens, (2004) "Spatial Distribution of Economic Activities in North America," forthcoming *Handbook of Regional and Urban Economics*

Krugman, Paul (1991), "Increasing Returns and Economic Geography," *Journal of Political Economy* 99 (June): 483–99.

Lucas, Robert E., Jr. (2001), "Externalities and Cities," *Review of Economic Dynamics* 4 (April): 245–74.

Rosenthal, Stuart S., and William C. Strange (2001), "The Determinants of Agglomeration," *Journal of Urban Economics* 50 (September): 191–229.

U.S. Bureau of the Census (1996), *USA Counties 1996,* CD-ROM, Washington, D.C.: U.S. Department of Commerce.

__________ (2000), *1997 Wholesale Trade Geographic Area Series,* Washington, D.C.: U.S. Department of Commerce.

__________ (2001), *1997 Economic Census,* CD-ROM, Washington, D.C.: U.S. Department of Commerce.

U.S. Bureau of Transportation Statistics (2000), *Public Use Airports,* Washington, D.C.: Department of Transportation.

Table 1
Sales Office Location Patterns for Selected Companies

Company	Industry	Number of Sales Offices	Fraction of MSAs with Sales Office By MSA population groupings (millions)				Estimate of α (std. err.)
			Under .5 (194 MSAs)	.5-2 (57 MSAs)	2-8 (19 MSAs)	Over 8 (3 MSAs)	
Philip Morris	Cigarettes	17	.01	.11	.37	.67	1.16 (.19)
Merck	Pharmaceuticals	29	.02	.14	.53	1.00	1.41 (.21)
Clorox	Household products	29	.01	.14	.58	1.00	1.61 (.22)
Kimberly Clark	Paper	33	.02	.11	.58	.67	1.08 (.14)
Eli Lilly	Pharmaceuticals	60	.02	.42	.79	1.00	1.36 (.15)
Rockwell	Industrial Automation	78	.08	.60	.89	1.00	1.15 (.11)
Sun Microsystems	Computers	84	.03	.56	1.00	1.00	2.38 (.30)
Cisco Systems	Networking Equipment	105	.11	.70	1.00	1.00	1.79 (.20)
Xerox	Photocopiers	138	.10	.81	1.00	1.00	2.21 (.25)
Kraft	Food Products	265	.16	.72	1.00	1.00	1.26 (.15)

Table 2
Sales Office Intensity Measures
By City Size

	MSA population (millions)				
	Non-msa	Under .5 (194 MSAs)	.5-2 (57 MSAs)	2-8 (19 MSAs)	8+ (3 MSAs)
Per capita measures					
Sales ($1,000 per person)	0.62	2.70	4.92	7.98	6.86
Employment (per 1,000 in population)	0.99	2.71	3.86	4.86	5.23
Payroll ($1,000 per person)	0.03	0.11	0.18	0.27	0.28
Operating Expenses ($1,000 per person)	0.07	0.22	0.36	0.54	0.61
Inventories ($1,000 per person)	0.05	0.14	0.16	0.22	0.24
Location Quotients					
Sales	0.11	0.29	0.96	1.57	1.34
Employment	0.23	0.41	1.03	1.30	1.39
Payroll	0.15	0.32	0.95	1.47	1.48
Operating Expenses	0.17	0.33	0.93	1.40	1.57
Inventories	0.27	0.49	0.97	1.31	1.42

Source: Author’s calculations with publicly available data from the 1997 Census of Wholesale Trade.

Table 3
MSA-level Regressions
Log Sales on Log Population

	Slope (std. err.)	R^2
Cross-section regressions (year)		
1982	1.64 (.05)	.78
1987	1.63 (.06)	.76
1992	1.68 (.06)	.75
1997	1.71 (.06)	.75
1997, with controls for: - education, - airport access - manufacturing activity	1.56 (.06)	.82
Fixed-effect regression (1982-1997)	1.80 (.34)	.09

Source: Author's calculations with confidential micro data from the Census of Wholesale Trade, 1982-1997

Table 4
Summary Statistics by Sales Size of Firm
Mean Sales Size and Cell Counts
By Sale Size Category

Sales of Firm	Mean Sales ($ Millions)	Number of Firms	Number of Offices	Offices Per Firm
Under 25	7.5	2,097	3,551	1.7
25-50	35.8	426	1,403	3.3
50-100	70.7	364	1,772	4.9
100-250	159.1	368	2,791	7.6
250-1000	479.0	335	4,628	13.8
1000+	4,856.9	196	5,566	28.4
All Firms	324.3	3,786	19,711	5.2

Source: Author's calculations from confidential micro data from the 1997 Census of Wholesale Trade. Non-MSA establishments excluded

Table 5
Location Quotients
by Sales Size of the Firm and MSA Size

A. Raw Data, 1997 Census

	MSA population (millions)			
Sales of Firm	Under .5	.5-2	2-8	8+
Under 25	0.76	0.91	1.04	1.27
25-50	0.62	0.78	1.09	1.47
50-100	0.52	0.80	1.13	1.48
100-250	0.55	0.89	1.16	1.28
250-1000	0.44	0.89	1.22	1.29
1000+	0.32	0.85	1.33	1.28

B. Estimates of a Logit Model of Sales Distribution, 1997 Census

	MSA population (millions)			
Sales of Firm	Under .5	.5-2	2-8	8+
Under 25	0.52	0.82	1.12	1.48
25-50	0.47	0.78	1.15	1.51
50-100	0.32	0.76	1.29	1.44
100-250	0.42	0.90	1.23	1.29
250-1000	0.33	0.91	1.28	1.27
1000+	0.30	0.84	1.45	1.10

C. Estimates of a Logit Model of Sales Distribution, 1992 Census

	MSA population (millions)			
Sales of Firm	Under .5	.5-2	2-8	8+
Under 25	0.45	0.82	1.20	1.42
25-50	0.41	0.81	1.13	1.58
50-100	0.34	0.81	1.22	1.51
100-250	0.33	0.92	1.23	1.36
250-1000	0.29	0.86	1.39	1.23
1000+	0.28	0.84	1.47	1.15

Source: Author's calculations with confidential micro data from the 1997 and 1992 Census of Wholesale Trade

Table 6
The Distribution of the Match Component
Scaling set at N = 1

City Index j	Population Share n_j	Cumulative Population Share	Mean of ε_j	Variance of ε_j
All Cities			4.0	3.8
Individual Cities				
1	0.0003	0.0003	7.7	1.6
215	0.0035	0.2509	5.1	1.6
269	0.0314	0.7592	2.9	1.6
273	0.0935	1.0000	1.8	1.6

The cumulative population share is the population share of city *j* plus that of all cities smaller than *j*. The mean and variance calculations use the population weights.

Table 7
Structural Parameter Estimates

Parameter	Model 1: No Additional City Characteristics	Model 2: Additional City Characteristics
τ	3.81 (.07)	3.75 (.07)
ϕ	1.52 (.10)	1.53 (.09)
γ	.15 (.69)	5.93 (1.35)
$\lambda_{college}$	-	-.02 (.01)
$\lambda_{airports}$	-	.22 (.02)
$\lambda_{manufacturing}$	-	.05 (.01)
Number of MSAs	273	273
Number of Firms	3786	3786

Source: Author's estimates with moments from the 1997 Census of Wholesale Trade that have been cleared for release by the Census Bureau and are available from the author.

Table 8
Comparison of Model 1 with 1997 Census Data

Panel A. The 1997 Census Data

Sales of Firm	Offices per firm	Sales Location Quotient by MSA population groupings (millions)			
		Under .5	.5-2	2-8	8+
Under 25	1.7	0.76	0.91	1.04	1.27
25-50	3.3	0.62	0.78	1.09	1.47
50-100	4.9	0.52	0.80	1.13	1.48
100-250	7.6	0.55	0.89	1.16	1.28
250-1000	13.8	0.44	0.89	1.22	1.29
1000+	28.4	0.32	0.85	1.33	1.28

Panel B. Model 1

Sales of Firm	Offices per firm	Sales Location Quotient by MSA population groupings (millions)			
		Under .5	.5-2	2-8	8+
Under 25	1.3	0.88	0.90	0.97	1.28
25-50	3.9	0.80	0.81	0.96	1.48
50-100	6.2	0.75	0.77	1.01	1.49
100-250	9.8	0.71	0.75	1.07	1.46
250-1000	15.6	0.67	0.77	1.10	1.43
1000+	25.4	0.67	0.80	1.10	1.40

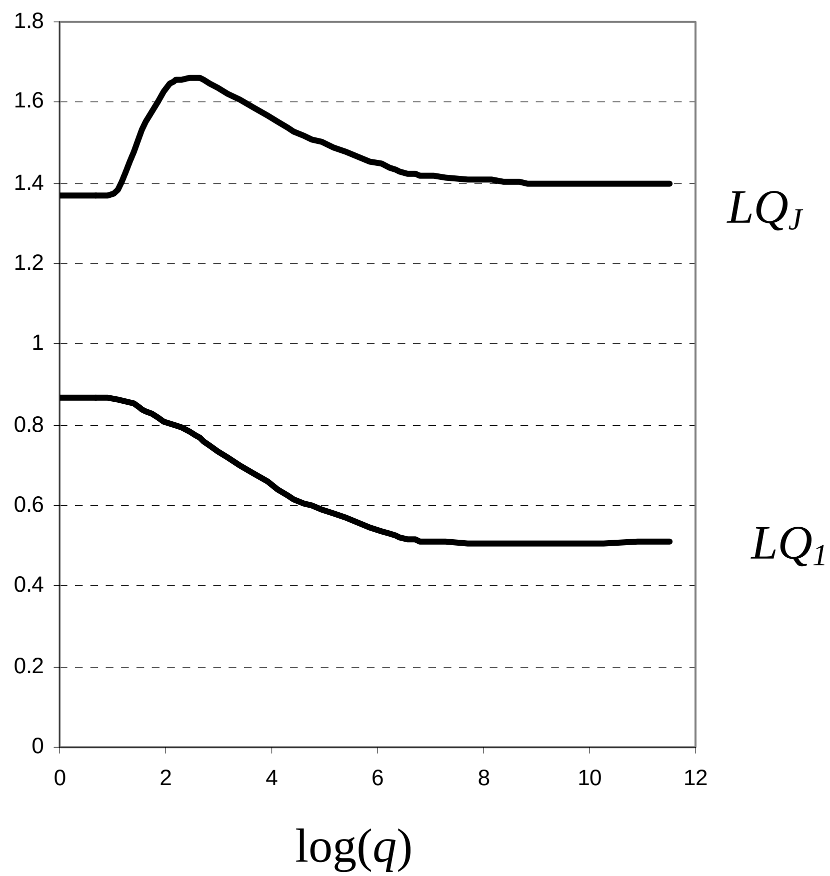
Panel C. Model 2

Sales of Firm	Offices per firm	Sales Location Quotient by MSA population groupings (millions)			
		Under .5	.5-2	2-8	8+
Under 25	1.3	0.65	0.89	1.16	1.19
25-50	3.9	0.59	0.80	1.12	1.41
50-100	6.4	0.56	0.76	1.17	1.42
100-250	10.0	0.53	0.74	1.22	1.40
250-1000	15.1	0.50	0.76	1.24	1.38
1000+	22.4	0.50	0.78	1.23	1.36

Table A1
Distribution of City Characteristics
by City Size

	MSA population (millions)				
	Non-msa	Under .5	.5-2	2-8	8+
Education Level (percent of population 25 years and older with 4 years of college)	13.26	18.56	20.44	24.64	23.80
Airport Activity (domestic enplanements per person in 1997)	-	.92	2.76	4.06	2.64
Manufacturing Activity (sales of manufacturing plants, $1,000 per person in 1997)	14.54	16.65	15.36	14.27	10.86

Figure 2
Location Quotients in City 1 and City *J* as a Function of Firm Size
Numerical Example



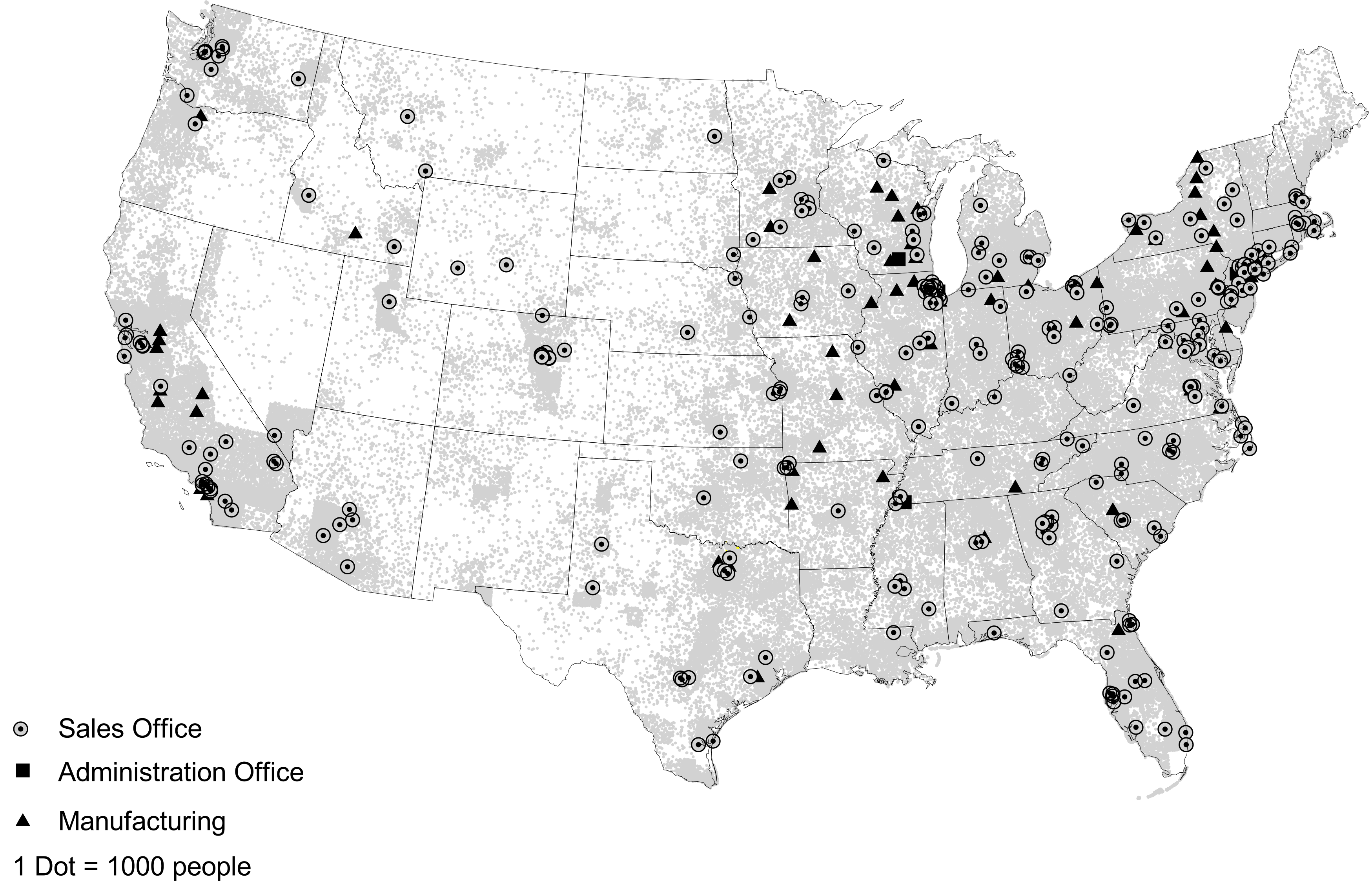


Fig. 1. Map of facilities (Kraft Foods)