

Separate Appendix for “The Location of Sales Offices
and the Attraction of Cities”

Thomas J. Holmes*

September 9, 2004

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

This manuscript provides supporting material cited in my paper:

Holmes, Thomas J. “Sales Offices and the Attraction of Cities.” University of Minnesota manuscript, September 2004.

1 Public Data on Ten Example Companies

This section discusses how I collected, from public sources, information about the establishment locations of the ten familiar companies listed in Table 1 of the paper.

ReferenceUSA, a web-based information source published by InfoUSA, was the initial source of a list of establishments by company and addresses. This directory attempts to be a comprehensive list of all business locations. A problem that plagues this data is that manufacturers sales offices are commonly misclassified as manufacturing plants. (It is worth noting that Dun and Bradstreet data also has this problem. Dun and Bradstreet’s count of the number of manufacturing plants in the U.S. is on the order of twice the count of the Census Bureau.)

To make progress here, I supplemented the ReferenceUSA information with information from Harris InfoSource’s Selectory Online. Unlike the ReferenceUSA, Harris does not aim to be a comprehensive list of all businesses from all industries. Instead it places a special focus on manufacturing and explicitly tries to distinguish between true manufacturing plants and sales offices.

My main strategy was to use Harris to create a list of manufacturing plants and administrative facilities. (This list of was supplemented at times by company information from the web, and by information from Nexis-Lexis). With a list of manufacturing plants in hand, I turned to the ReferenceUSA information and I reviewed the industry classification of each establishment. In the review process, information about establishment name was helpful as establishments might have names like, “XEROX CORPORATION - ATLANTA SALES OFFICE.” For examining names, looking at yellow pages listings was useful. ReferenceUSA is based on yellow pages listings but names are sometimes altered (e.g. the previously mentioned name might be shortened to “Xerox Corporation”) so examining the original listings is useful. In two cases (Cisco Systems and Rockwell International), information about sales

office locations was specifically provided at company web sites.

Three considerations were used to select the example companies: (1) I restricted attention to large companies. (2) I restricted attention to companies for which I was able to obtain data that I viewed was relatively clean. (3) I attempted to obtain a broad mix of example industries.

2 Logit Model with Non-MSA included

The empirical section of the paper presents several several tables of location quotients by city size class and firm size class based on a logit model that adds dummy variables for industries. (Panels B and C of Table 5) As mentioned in the paper, the procedure used in the paper included rural areas, but to be consistent with the other location quotients contained in the paper, the rural population was left out in the calculation of the LQs for Table 5.) Here I show what the location quotients look like with the non-MSA areas included. I am using the same logit estimates, the difference is in the denominator of the location quotients. Here, the denominator is the population share of the total U.S. population, instead of just total metropolitan population, as in the paper. Adding the non-MSA increases the location quotients in the largest cities because their share of sales office activity doesn't go down much (since MSA sales office activity is negligible) while the share of population in the denominator decreases.

Panels B and C in Table N1 correspond to Panels B and C in Table 6 of the paper. Panel C uses number of establishments of the firm to categorize firm size. The results using establishments counts as a size measure are similar to the results using sales.

3 Notes for Theoretical Section

3.1 Notes for Proposition 1

The proof of Proposition 1 claims that straightforward calculations show that $\overline{LQ}'_J(\frac{\phi}{q}) > 0$ when evaluated at $\frac{\phi}{q} = 0$. These notes provide these straightforward calculations. From

the formulas derived in the paper, the probability that location J services location itself is

$$\begin{aligned} p_{J,J} &= \frac{n_J e^{\gamma n_J - \frac{\phi}{q n_J}}}{\sum_{j \neq J} n_j e^{\gamma n_j - \tau} + n_J e^{\gamma n_J - \frac{\phi}{q n_J}}} \\ &= \frac{n_J e^{\gamma n_J}}{\sum_{j=1}^{J-1} n_j e^{\gamma n_j - \tau + \frac{\phi}{q n_J}} + n_J e^{\gamma n_J}} \end{aligned}$$

and the probability it services location j is

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

Differentiating,

$$\begin{aligned} \frac{dp_{J,J}}{d\left[\frac{\phi}{q}\right]} &= -\frac{e^{\gamma n_J}}{\left[\sum_{j=1}^{J-1} n_j e^{\gamma n_j - \tau + \frac{\phi}{q n_J}} + n_J e^{\gamma n_J}\right]^2} \sum_{j=1}^{J-1} n_j e^{\gamma n_j - \tau + \frac{\phi}{q n_J}} \\ &= -\frac{e^{\gamma n_J}}{\left[\sum_{j=1}^{J-1} n_j e^{\gamma n_j - \tau} + n_J e^{\gamma n_J}\right]^2} \sum_{j=1}^{J-1} n_j e^{\gamma n_j - \tau}, \text{ at } \frac{\phi}{q} = 0 \\ \frac{dp_{J,j}}{d\left[\frac{\phi}{q}\right]} &= \frac{n_J e^{\gamma n_J}}{\left[\sum_{k \neq j} n_k e^{\gamma n_k} + n_j e^{\gamma n_j + \tau - \frac{\phi}{q n_j}}\right]^2} e^{\gamma n_j + \tau - \frac{\phi}{q n_j}} \\ &= \frac{n_J e^{\gamma n_J}}{\left[\sum_{k \neq j} n_k e^{\gamma n_k} + n_j e^{\gamma n_j + \tau}\right]^2} e^{\gamma n_j + \tau} \end{aligned}$$

Now look at:

$$\begin{aligned} \overline{LQ}_J &= p_J + \sum_{j=1}^{J-1} \frac{n_j}{n_J} p_j \\ \frac{d\overline{LQ}_J}{d\left[\frac{\phi}{q}\right]} \left(\frac{\phi}{q} = 0\right) &= -\frac{e^{\gamma n_J}}{\left[\sum_{k=1}^{J-1} n_k e^{\gamma n_k - \tau} + n_J e^{\gamma n_J}\right]^2} \sum_{j=1}^{J-1} n_j e^{\gamma n_j - \tau} \\ &\quad + \sum_{j=1}^{J-1} \frac{n_j e^{\gamma n_J}}{\left[\sum_{k \neq j} n_k e^{\gamma n_k} + n_j e^{\gamma n_j + \tau}\right]^2} e^{\gamma n_j + \tau} \end{aligned}$$

To show this is positive, it is sufficient to show that for all $j < J$,

$$\frac{n_j e^{\gamma n_J}}{\left[\sum_{k \neq j} n_k e^{\gamma n_k} + n_j e^{\gamma n_j + \tau}\right]^2} e^{\gamma n_j + \tau} > \frac{e^{\gamma n_J} n_j e^{\gamma n_j - \tau}}{\left[\sum_{k=1}^{J-1} n_k e^{\gamma n_k - \tau} + n_J e^{\gamma n_J}\right]^2}$$

or

$$\frac{e^\tau}{\left[\sum_{k \neq j} n_k e^{\gamma n_k} + n_j e^{\gamma n_j + \tau}\right]^2} > \frac{e^{-\tau}}{\left[\sum_{k=1}^{J-1} n_k e^{\gamma n_k - \tau} + n_J e^{\gamma n_J}\right]^2}$$

or

$$> \frac{\left[\sum_{k=1}^{J-1} n_k e^{\gamma n_k} + n_J e^{\gamma n_J + \tau}\right] \left[\sum_{k=1}^{J-1} n_k e^{\gamma n_k - \tau} + n_J e^{\gamma n_J}\right]}{\left[\sum_{k \neq j} n_k e^{\gamma n_k - \tau} + n_j e^{\gamma n_j}\right] \left[\sum_{k \neq j} n_k e^{\gamma n_k} + n_j e^{\gamma n_j + \tau}\right]}$$

or, noting the cancellations,

$$n_j^2 e^{2\gamma n_j - \tau} + n_j^2 e^{2\gamma n_J + \tau} + 2n_j n_J e^{(n_j + n_J)\gamma} > n_j^2 e^{2\gamma n_j + \tau} + n_j^2 e^{2\gamma n_J - \tau} + 2n_j n_J e^{(n_j + n_J)\gamma}$$

or

$$n_j^2 e^{2\gamma n_j - \tau} + n_j^2 e^{2\gamma n_J + \tau} > n_j^2 e^{2\gamma n_j + \tau} + n_j^2 e^{2\gamma n_J - \tau}$$

or

$$n_j^2 e^{2\gamma n_J} (e^\tau - e^{-\tau}) > n_j^2 e^{2\gamma n_j} (e^\tau - e^{-\tau})$$

which holds since $n_J > n_j$.

3.2 Notes for Proposition 2

Proposition 2 uses the following lemma which I prove here.

Lemma. Take a vector (n_1, n_2, \dots, n_J) satisfying $0 \leq n_j$, all j , $\sum_{j=1}^J n_j = 1$, and

$$n_1 \leq n_2 \leq \dots \leq n_J < \frac{1}{2}.$$

Then

$$H \equiv \sum_{k=1}^J n_k^2 + \alpha^2 + 2 \left[\sum_{k=1}^J n_k^2 \right]^2 - 3 \sum_{k=1}^J n_k^3 > 0.$$

Proof. Denote $n_J = \alpha$. I need to show the result holds for any $\alpha < \frac{1}{2}$. I prove the result in two steps.

Step 1.

Fix $\alpha < \frac{1}{2}$. For any $x \leq (1 - \alpha)^2$, consider the problem of

$$\begin{aligned} & \max_{J, (n_1, n_2, \dots, n_{J-1})} \sum_{k=1}^{J-1} n_k^3 \\ & : \text{ subject to } \sum_{k=1}^{J-1} n_k^2 \leq x \\ & : n_k \in [0, \alpha]. \end{aligned}$$

I claim that the solution to this problem has the following corner solution. Let y be the number of α units in $1 - \alpha$ so that

$$\begin{aligned} y\alpha & \leq 1 - \alpha \\ (y + 1)\alpha & > 1 - \alpha. \end{aligned}$$

The solution is to fill up cities sequentially up to α then start new ones. This filling procedure stops when we have $z \leq y$ full cities (i.e. with population α) and a residual city type with n_r so that

$$z\alpha^2 + n_r^2 = x,$$

So the first constraint in the above problem is binding.

To prove this claim, suppose we have two cities, say j and k , and $n_j > 0$, $n_k > 0$ but $n_j < \alpha$ and $n_k < \alpha$. Assume $n_j < n_k$. Take a small amount w from city j . Then put $v < w$ in the larger city where v solves

$$(n_j - w)^2 + (n_k + v)^2 = (n_j)^2 + (n_k)^2 \tag{1}$$

This new allocation is feasible for the problem. (Observe we have free disposal of population units in the problem so $v < w$ is feasible.). Now since $n_j < n_k$, straightforward calculus shows that

$$(n_j - w)^3 + (n_k + v)^3 > n_j^3 + n_k^3$$

for small w , letting $v(w)$ be the implicit function solving (1). Hence the initial allocation could not solve the above problem.

Step 2. Take as given some J and $(n_1, n_2, \dots, n_{J-1})$ and that $\alpha < \frac{1}{2}$.

Let

$$x = \sum_{k=1}^{J-1} n_k^2$$

I need to show that

$$\begin{aligned} H &= \sum_{k=1}^{J-1} n_k^2 + \alpha^2 + 2 \left[\sum_{k=1}^{J-1} n_k^2 + \alpha^2 \right]^2 - 3 \sum_{k=1}^{J-1} n_k^3 - 3\alpha^3 \\ &= x + \alpha^2 + 2 [x + \alpha^2]^2 - 3 \sum_{k=1}^{J-1} n_k^3 - 3\alpha^3 > 0 \end{aligned}$$

But let $z(x), n_r(x)$ be such that

$$z(x)\alpha^2 + n_r(x)^2 = x$$

where start by filling up cities up to α and stop when get the above. Then from step 1 we know that

$$\begin{aligned} &x + \alpha^2 + 2 [x + \alpha^2]^2 - 3 \sum_{k=1}^{J-1} n_k^3 - 3\alpha^3 \\ &\geq x + \alpha^2 + 2 [x + \alpha^2]^2 - 3 [z(x)\alpha^3 + n_r(x)^3] - 3\alpha^2 \\ &\equiv G. \end{aligned}$$

We have to show that the bottom term is strictly positive for all $x \leq (1 - \alpha)^2$. Rather than start with x and back out z and n_r is easier to vary z and n_r . So then we can write the above as

$$G = z\alpha^2 + n_r^2 + \alpha^2 + 2 [z\alpha^2 + n_r^2 + \alpha^2]^2 - 3z\alpha^3 - 3n_r^3 - 3\alpha^2$$

where

$$z \leq \frac{1 - \alpha}{\alpha}$$

Pick a z satisfying the above and a n_r such that

$$z\alpha + n_r \leq 1 - \alpha$$

Set $\lambda = z + 1$. Then

$$\begin{aligned} G &= z\alpha^2 + n_r^2 + \alpha^2 + 2 [z\alpha^2 + n_r^2 + \alpha^2]^2 - 3z\alpha^3 - 3n_r^3 - 3\alpha^2 \\ &= \lambda\alpha^2 + n_r^2 + 2 [\lambda\alpha^2 + n_r^2]^2 - 3\lambda\alpha^3 - 3n_r^3 \\ &= \lambda\alpha^2 [1 + 2\lambda\alpha^2 - 3\alpha] + n_r^2 [1 + 2n_r^2 - 3n_r] + 4\lambda\alpha^2 n_r^2 \end{aligned}$$

Now clear $n_r \leq \frac{1}{2}$, so the second term is positive. Now $\lambda \geq 1$. The first term is strictly positive for $\alpha < \frac{1}{2}$. The result follows. *Q.E.D.*

Table N1

Estimated Location Quotients from Logit Model with Industry Controls

B. 1997 Data and Size Defined by Sales

Sales of Firm	MSA population (millions)				
	Non- msa	Under .5	.5-2	2-8	8+
Under 25	0.30	0.61	0.96	1.31	1.74
25-50	0.21	0.56	0.94	1.38	1.81
50-100	0.13	0.39	0.93	1.58	1.76
100-250	0.18	0.50	1.08	1.49	1.55
250-1000	0.15	0.40	1.10	1.55	1.55
1000+	0.04	0.38	1.05	1.80	1.37

C. 1992 Data and Size Defined by Sales

Sales of Firm	MSA population (millions)				
	Non- msa	Under .5	.5-2	2-8	8+
Under 25	0.22	0.54	0.97	1.43	1.69
25-50	0.23	0.49	0.97	1.34	1.88
50-100	0.11	0.41	0.99	1.49	1.83
100-250	0.14	0.40	1.12	1.49	1.65
250-1000	0.13	0.35	1.05	1.69	1.49
1000+	0.06	0.34	1.03	1.80	1.42

D. 1997 Data and Size Defined by Establishment Counts

Number of Establishments	MSA population (millions)				
	Non- msa	Under .5	.5-2	2-8	8+
1	0.10	0.30	1.06	1.73	1.46
2-5	0.10	0.69	0.82	1.20	2.27
6-20	0.06	0.37	1.22	1.46	1.70
21+	0.05	0.36	1.04	1.91	1.19

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