Agglomeration Externalities and the Dynamics of Firm Location Choices within an Urban Economy*

Jeff Brinkman
Federal Reserve Bank of Philadelphia

Daniele Coen-Pirani
University of Pittsburgh

Holger Sieg
University of Pennsylvania and NBER

July 7, 2011

*We would like to thank Dennis Epple, Vernon Henderson, Thomas Holmes, Matt Kahn, Theodore Papageorgiu, Stephen Redding, Steven Ross, Esteban Rossi-Hansberg and seminar participants at the Federal Reserve Bank of Atlanta, Carnegie Mellon University, a Housing Workshop at Duke University, the University of Iowa, the University of Montreal, the CURE Conference at Princeton, the Regional Science Conference in Denver, the 2011 SED meetings, and the 8th Winter Meeting of German Economists Abroad at the University of Frankfurt for comments. This research is supported by the NSF (SES-0958705).
Abstract

We develop a new dynamic general equilibrium model to explain firm entry, exit, and relocation decisions in an urban economy with multiple locations. We characterize the stationary distribution of firms that arises in equilibrium. The parameters of the model can be estimated using a nested fixed point algorithm by matching the observed distribution of firms by location and the one implied by our model. We implement the estimator using unique data collected by Dun and Bradstreet for the Pittsburgh metropolitan area. Firms located in the central business district are older and larger than firms located outside the urban core. They use more land and labor in the production process. However, they face higher rental rates for office space which implies that they operate with a higher employee per land ratio. Our estimates imply that agglomeration externalities increase the productivity of firms by one to two percent. Economic policies that subsidize firm relocations can potentially have large effects on economic growth and firm concentration in central business districts.
1 Introduction

Cities play an important role in the economy since economic proximity makes for more efficient production and trade. These efficiency gains typically arise because of agglomeration externalities.\footnote{The idea of geographic returns to scale was first introduced by von Thünen (1825). Marshall (1890) suggested that agglomeration effects may exist within industries. Firms may benefit from lower transaction costs or sharing of a common labor pool. Alternatively, efficiencies may arise due to positive diversity externalities and synergies between different industries (Jacobs, 1969).} Firms that operate in locations with high externalities have a competitive advantage over firms that are located in less efficient locations. Since firms will bid for the right to locate in areas with high agglomeration externalities, these locations have higher rental prices for land than locations that are less efficient.\footnote{Krugman (1991) provides theoretical foundations for a two-location model of agglomeration. Ellison and Glaeser (1997) argued that agglomeration externalities are important to understand geographic concentration of manufacturing in the U.S. The literature of agglomeration theory is reviewed in Fujita and Thisse (2002) and Duranton and Puga (2004). Anas and Kim (1996) and Lucas and Rossi-Hansberg (2002) have developed equilibrium models of mono- and poly-centric urban land use with endogenous congestion and job agglomeration. Rossi-Hansberg (2004) studies optimal land use policies in a similar framework.} As a consequence, firms with different productivity levels will sort in equilibrium with high productivity firms locating in areas with high agglomeration effects and high rental prices for land. Low productivity firms are forced to exit the economy or operate in cheaper locations.

As the productivity of a firm changes over time, a firm’s demand for land and labor changes as well. Moreover, productivity shocks create incentives to relocate within a city to exploit a better match with the agglomeration externalities. A firm that may have initially located outside a central business district may find it in its interest to move to a more densely populated central business district in order to grow and capture the full benefits of a persistent positive productivity shock. Similarly, a firm that has experienced a persistent negative shock must downsize and move to the urban fringe where land and labor is cheaper than in the central business district.\footnote{There is some evidence that shows that agglomeration effects are important to understand firm}
The first objective of this paper is then to develop a new dynamic general equilibrium model of firm location choice that can explain the sorting of firms by productivity as well as entry, exit, and relocation decisions of firms in an urban economy with multiple locations.

We consider a model with two distinct locations which can be interpreted as inside and outside of a central business district (CBD). In equilibrium these locations differ in the magnitude of their agglomeration externalities which increase with employment density. Firms are heterogeneous in their productivities. We model firm dynamics and industry equilibrium following Hopenhayn (1992). Firms enter our urban economy with an initial productivity and must pay an entry cost. Productivity then evolves according to a stochastic first order Markov process. Each period firms compete in the product market, must pay a fixed cost of operating, and realize a profit. Entry, exit and relocations are dynamic and based on expectations of future productivity shocks.

We characterize the optimal decision rules for firms in each location as well as dynamics. Henderson, Kunkoro, and Turner (1995) show that agglomeration effects for mature industries are related to Marshall scale economies, while newer industries benefit from diversity akin to Jacobs economies. This work is important because it points to agglomeration as part of a dynamic process. Other research has continued to study the relevance of agglomeration in firm life-cycle dynamics. Duranton and Puga (2001) study the the effect of agglomeration externalities in innovation and the development of production processes, while Dumais, Ellison, and Glaeser (2002) examine the effect of firm dynamics (entry, exit, expansion, and contraction) on the concentration of economic activity.

Deckle and Eaton (1999) find that geographic scale of agglomeration is mostly at the national level, while the financial sector is concentrated in specific metropolitan areas. Other work finds that agglomeration can occur on a much more local scale. In particular, Rosenthal and Strange (2001, 2003) establish the level and type of agglomeration at different geographic scales, and also the measure the attenuation of these externalities within metropolitan areas. Holmes and Stevens (2002) finds evidence of differences in plant scale in areas of high concentration, suggesting production externalities act on individual establishments. A review of empirical evidence of agglomeration economies is found in (Rosenthal and Strange, 2004).

Our work is also related to Ericson and Pakes (1995) who consider the implications of oligopolistic competition on market structure. That framework is more appropriate when there are few competitors in the industry. Pakes and McGuire (1994, 2001) discuss how to solve models with oligopolistic competition. Doraszelski and Pakes (2006) provide a survey of that literature.
those for potential entrants. Low productivity firms exit from the economy, while high productivity firms continue to operate. Relocation choices are driven by the interaction of agglomeration effects and firm productivity shocks. Due to a minimum land requirement in the production function, large firms with higher productivity shocks prefer locations with high agglomeration externalities relative to smaller, less productive firms. As a consequence, a high productivity firm that is located outside the central business district may have strong incentives to relocate to the city center. Low productivity firms leave and move to a location outside the central business district. This process gives rise to a stationary equilibrium in which firms located in the CBD are on average larger and older than firms located outside the CBD.

We then develop an algorithm that can be used to estimate the parameters of our model. We focus on equilibria with entry in both locations since this is a common feature of the data. The parameters of the model are estimated using a nested fixed point algorithm. The inner loop computes the equilibrium for each parameter value, while the outer loop searches over feasible parameter values. Our simulated method of moments estimator matches the observed joint distribution of age, size and land use by location to the one predicted by our model.

We implement the estimator using unique data collected by Dun and Bradstreet

---

6 We abstract from innovation which is discussed in detail in Klette and Kortum (2004).
7 There are some similarities with the literature that studies the sorting pattern of household in urban areas which starts with the classic papers by Alonso (1964), Mills (1967), and Muth (1969).
8 Related to our research is also work by Melitz (2003) who studies the impact of trade on intra-industry relocations. Rossi-Hansberg and Wright (2007) examine the relationship of establishment scale and entry and exit dynamics. Finally, Combes, Duranton, Gobillon, Puga, and Roux (2010) distinguish between selection effects and productivity externalities by estimating productivity distributions across cities.
9 In related work, Davis et al. (2009) develop a growth model in which the total factor productivity of cities depends on the density of economic activity. They estimate the magnitude of this external effect and evaluate its importance for the growth rate of consumption per capita in the U.S. Our paper is thus also related to a growing literature in industrial organization that estimates dynamic models of oligopolistic competition. See, for example, Benkard (2004), Bajari, L. Benkard, and Levin (2007), Aguirregabiria and Mira (2007).
for the Pittsburgh metropolitan area. U.S. cities often act as a hub for services for a larger region. We, therefore, focus on locational choices within the service sector excluding industries in which proximity to the consumer is a key factor for firm location. The data suggest that firms located in the city are older and larger than firms located in the rest of the metro area. As a consequence they use more land and labor in the production process. However, they face higher rental rates for land and office space. Thus, they operate with a higher employee per land ratio. We find that our model explains these observed features of the data reasonably well. Our estimates imply that agglomeration externalities increase the productivity of firms by one to two percent. Economic policies that subsidize firm relocations can potentially have large effects on economic growth and firm concentration in central business districts.

The rest of the paper is organized as follows. Section 2 describes the data set used in our application and characterizes firm sorting within one metropolitan area. Section 3 develops our stochastic, dynamic equilibrium model and discusses its properties. Section 4 describes the estimation of the parameters of our model. Section 5 presents the empirical results and discusses the policy experiments. Section 6 offers some conclusions that can be drawn from the analysis.

## 2 Data

Our empirical application focuses on firm location choices in the City of Pittsburgh and Allegheny County.\(^{10}\) We are interested in characterizing the observed sorting of establishments by age, employment, and facility size.\(^{11}\) We focus on service industries,

---

\(^{10}\)In Appendix A of the paper we show that most other large metropolitan areas in the U.S. show sorting patterns of firms that are similar to the one we find for Pittsburgh. The comparison is based on aggregate Census data while the estimation of our model uses micro level data from Dun and Bradstreet.

\(^{11}\)While we use the terms firm and establishment interchangeably, our unit of analysis in the empirical section is an establishment.
given that there is strong evidence that large U.S. cities have undergone a transformation during the past decades moving from centers of individual manufacturing sectors toward becoming hubs for service industries. Duranton and Puga (2005), for example, show evidence that cities have become more functionally specialized, with larger cities, in particular, emerging as centers for headquarters and business services. They posit that this change is primarily related to industrial structure, and a decrease in remote management costs in particular. Davis and Henderson (2008) provide further evidence that services and headquarters are indeed more concentrated in large cities relative to the entire economy, and that headquarter concentration is linked to availability of diverse services.

We exclude wholesale and retail businesses from our analysis of services since locational decisions of these businesses are primarily driven by proximity to consumers (Hotelling, 1929).\footnote{Following Bresnahan and Reiss (1991), there is a large literature that explains entry and exit into markets with a small number of potential entrants. Holmes (2010) has estimated a dynamic model of market penetration of Walmart.} For similar reasons, we also do not consider businesses in the entertainment sector. Finally, we omit businesses related to agriculture, forestry, mining and fishing for fairly obvious reasons. We thus define the service sector as consisting of businesses that operate in information, finance, real estate, professional services, management, administrative support, education, health care and related sectors.

Figure 1 plots the employment concentration in Allegheny County using data from the U.S. Census. Over 20 percent of employment is concentrated in three zip codes in the center of Pittsburgh which include the downtown central business district and the business district in Oakland which are the two significant dense commercial areas of Pittsburgh.\footnote{The zip codes are 15222, 15219, and 15213.} We treat these locations as the CBD in estimation while all remaining
places of Allegheny County are treated as the alternate location (NCBD).\textsuperscript{14}

We use firm level data from Dun and Bradstreet’s Million Dollar Database to estimate our model.\textsuperscript{15} This database covers establishments in Allegheny county in 2008 and provides detailed information on establishments. The coverage is near universal compared to Census counts of establishments in the county. The database provides data on location, facility size, total employment, industry, and year established.\textsuperscript{16}

We analyze the employment and facility size characteristics for different industries in the Pittsburgh area. Table 1 reports the total employment, the average employment and the facility space per employee for firms in and outside the CBD for selected service industries.

Table 1: Employment and facility size by industry in 2008

<table>
<thead>
<tr>
<th>NAICS</th>
<th>Total Emp.</th>
<th>% Empl</th>
<th>Size CBD</th>
<th>Size NCBD</th>
<th>Facility CBD</th>
<th>Facility NCBD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>16,975</td>
<td>25.15%</td>
<td>13.52</td>
<td>31.16</td>
<td>336.88</td>
<td>214.44</td>
</tr>
<tr>
<td>Finance</td>
<td>42,960</td>
<td>53.51%</td>
<td>8.55</td>
<td>55.66</td>
<td>318.28</td>
<td>193.59</td>
</tr>
<tr>
<td>Real Estate</td>
<td>18,459</td>
<td>17.97%</td>
<td>7.51</td>
<td>12.43</td>
<td>743.36</td>
<td>1190.21</td>
</tr>
<tr>
<td>Professional Services</td>
<td>64,076</td>
<td>32.85%</td>
<td>6.99</td>
<td>13.29</td>
<td>334.83</td>
<td>309.60</td>
</tr>
<tr>
<td>Management</td>
<td>2,062</td>
<td>11.30%</td>
<td>19.46</td>
<td>14.56</td>
<td>272.88</td>
<td>360.52</td>
</tr>
<tr>
<td>Administrative Support</td>
<td>41,830</td>
<td>14.97%</td>
<td>11.01</td>
<td>20.67</td>
<td>240.89</td>
<td>352.14</td>
</tr>
<tr>
<td>Education</td>
<td>52,995</td>
<td>42.69%</td>
<td>30.46</td>
<td>205.66</td>
<td>316.70</td>
<td>121.27</td>
</tr>
<tr>
<td>Health Care</td>
<td>115,048</td>
<td>18.12%</td>
<td>16.53</td>
<td>24.01</td>
<td>293.39</td>
<td>291.46</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>354,405</strong></td>
<td><strong>28.66%</strong></td>
<td><strong>11.78</strong></td>
<td><strong>27.47</strong></td>
<td><strong>326.81</strong></td>
<td><strong>265.19</strong></td>
</tr>
</tbody>
</table>

Note: Size is average employment. Facility is average facility size measured in square foot per employee.

\textsuperscript{14}None of the results reported in this paper rely on this definition of the alternate location. We can, for example, omit those parts of Allegheny county that have little economic development and obtain similar results regarding firm sorting.

\textsuperscript{15}Information on Dun and Bradstreet data is available on-line at http://www.dnbmdd.com/

\textsuperscript{16}While most of the data are complete, the year established field, which is used to determine age of establishments, is only available for 52.5 percent of the observations. However, we find little evidence that the missing data field is systematically correlated with other observable data. We, therefore, treat observations without age of establishment as missing at random.
Figure 1: Allegheny County and the City of Pittsburgh
Table 1 shows that 28.66 percent of all employment in the service industry is located in the CBD or 13.43 percent of all our firms. However, the three zip codes that comprise the city account for a less than one percent of all the land in Allegheny county. We find that finance, education, and professional services are the industries that are most heavily concentrated in the CBD.

Comparing firms that are located inside the CBD with firms that are outside the CBD, we find some important patterns that hold for all service industries. The average employment size of establishments is larger in the central business district. The average establishment in the CBD employs 27 persons while the average firm outside the CBD has only 12 employees. However, rents for office space are higher in the CBD. As a consequence firms located in the CBD only use 265 square foot per employee while firms outside the CBD use 327.

To get some additional insights into the firm sorting process, we need to look at the full distribution of firms by location. Table 2 reports a number of percentiles of the age, facility size, and employment size distribution by location. Moreover, Dun and Bradstreet also report revenue estimates for each firm in the sample. These estimates must be interpreted with caution since they are likely to contain some measurement error. Nevertheless, we can use these estimates to compare output of firms across locations.

Table 2 reveals a number of important facts that characterize sorting of firms across locations. Firms in the CBD not only employ more workers and operate in larger facilities as we have seen above. They are also older and have a higher output per employee. The later fact is consistent with the notion that firms in the CBD may have higher productivity levels than firms located outside the CBD. Table 2 also shows that there are significant differences among firms in the right tail of the distribution. Looking at the 90th and higher percentiles, we find large differences between firms inside and outside of the CBD.
Table 2: Sorting of Firms by Location and Industry in 2008

<table>
<thead>
<tr>
<th>Percentile</th>
<th>CBD</th>
<th>Outside CBD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>age</td>
<td>employ</td>
</tr>
<tr>
<td>10th</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>25th</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>50th</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>75th</td>
<td>34</td>
<td>14</td>
</tr>
<tr>
<td>90th</td>
<td>53</td>
<td>40</td>
</tr>
<tr>
<td>95th</td>
<td>72</td>
<td>78</td>
</tr>
<tr>
<td>99th</td>
<td>119</td>
<td>460</td>
</tr>
</tbody>
</table>

Information Technology

<table>
<thead>
<tr>
<th>Percentile</th>
<th>CBD</th>
<th>Outside CBD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>age</td>
<td>employ</td>
</tr>
<tr>
<td>10th</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>25th</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>50th</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>75th</td>
<td>28</td>
<td>20</td>
</tr>
<tr>
<td>90th</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>95th</td>
<td>90</td>
<td>158</td>
</tr>
<tr>
<td>99th</td>
<td>156</td>
<td>218</td>
</tr>
</tbody>
</table>

Financial Services

<table>
<thead>
<tr>
<th>Percentile</th>
<th>CBD</th>
<th>Outside CBD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>age</td>
<td>employ</td>
</tr>
<tr>
<td>10th</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>25th</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>50th</td>
<td>23</td>
<td>7</td>
</tr>
<tr>
<td>75th</td>
<td>39</td>
<td>17</td>
</tr>
<tr>
<td>90th</td>
<td>70</td>
<td>49</td>
</tr>
<tr>
<td>95th</td>
<td>79</td>
<td>200</td>
</tr>
<tr>
<td>99th</td>
<td>145</td>
<td>2,034</td>
</tr>
</tbody>
</table>
One potential concern of the analysis above is that differences between firms located inside and outside the CBD may be due to aggregation bias. In particular, these differences could just reflect differences of sorting across different industries. In the lower panels of Table 2, we, therefore, report the same statistics for two industries. The middle panel reports the results for the information technology sector which is an "average" service industry in terms of its concentration of employment in the CBD. The lower panel reports the statistics for the finance industry which is the most heavily concentrated industry in our sample. We find that the qualitative differences between firms located inside and outside the CBD are not driven by aggregation across firms in the different service industries. In anything, the differences in the financial service industry are more pronounced than the differences in sample of all service industries.

3 A Dynamic Model of Firm Location within an Urban Area

3.1 Technologies and Markets

We consider a model with two locations, denoted by \( j = 1, 2 \). There is a continuum of firms that produce a single output good and compete in the product market. In each period a firm chooses to stay where it is, relocate to the other location, or shutdown. Firm are heterogeneous and productivity evolves according to a stochastic law of motion.

Assumption 1 In each period a firm is subject to an exogenous probability of exiting. We denote by \( \xi \) the complement probability of a firm surviving into the next period. If the firm survives, it draws a new productivity shock, \( \varphi' \) each time period.
productivity shock evolves over time according to a Markov process with a conditional distribution $F(\varphi'|\varphi)$.

In our parametrized model, we assume that the logarithm of the productivity shock follows an AR(1) process, i.e. $\log(\varphi') = \rho \log(\varphi) + \varepsilon'$, where $\rho$ is the correlation coefficient and $\varepsilon$ is a normally distributed random variable with mean $\mu_\varepsilon$ and variance $\sigma_\varepsilon^2$.

Each firm produces a single output good using labor and land as input factors. The technology that is available to the firms in the economy satisfies the following assumption.

**Assumption 2** The production function of a firm in location $j$ can then be written as:

$$q = f(\varphi, n, l; e_j)$$

where $q$ is output, $n$ is labor, $l$ is land, and $e_j$ is the agglomeration externality in location $j$. The production function satisfies standard regularity conditions.

Rosenthal and Strange (2003) suggest that the externality acts as a multiplier on the production function. We use a standard Cobb-Douglas function with parameters $\alpha$ and $\gamma$ in our computational model, $q = \varphi \; e_j \; n^\alpha (l - \bar{l})^\gamma$. Note that $\bar{l}$ is a minimum amount of land required for production. Since $r_j \; \bar{l}_j$ can also be interpreted as fixed costs, this specification implies that fixed cost vary by location.

The agglomeration externality arise due to a high concentration of firms operating in the same location.

**Assumption 3** The agglomeration externality can be written as

$$e_j = \Theta(L_j, N_j, S_j)$$
where \( N_j \) and \( L_j \) are aggregate measures of labor and land respectively, and \( S_j \) is a measure of the mass of firms in location \( j \). The function \( \Theta \) is such that \( \Theta_L < 0 \), \( \Theta_N > 0 \), and \( \Theta_S > 0 \).

Following Lucas and Rossi-Hansberg (2002), we assume in our computational model that

\[
e_j = \left( \frac{N_j}{L_j - S_j \bar{l}} \right)^\theta
\]

(3)

If \( \theta > 0 \), the externality is an increasing function of a measure of concentration of economic activity in a location \( j \). This measure is represented by the ratio of the total number of workers and the amount of land used in production over and above the minimum land requirement.

The urban economy is part of a larger economic system which determines output prices and wages.\(^{17}\)

**Assumption 4** Output prices, \( p \), and wages, \( w \), are constant and determined exogenously.

Rental prices, \( r_j \), however, are equilibrium outcomes. The supply of land is determined by an inverse land supply function in each location.

**Assumption 5** The inverse land supply function is given by:

\[
r_j = r_j(L_j), \quad j = 1, 2
\]

(4)

The inverse supply function is increasing in the amount of land denoted by \( L_j \).

In the computational analysis, we adopt an iso-elastic functional form: \( r_j = A_j L_j^\delta, \quad j = 1, 2 \), where \( A_j \) and \( \delta \) are parameters. Since rental prices for land must

\(^{17}\)it is straightforward to endogenize wages by adding a local labor market to our model.
be higher in equilibrium in locations with high externalities, the agglomeration externality is, at least, partially capitalized in land rents.

We can break down the decision problem of firms into a static and a dynamic problem. First, consider the static part of the decision problem that a firm has to solve each period. This problem arises because firms compete in the product market each period.

**Assumption 6** *The product market is competitive and firms behave as price takers. Firms make decisions on land and labor usage after they have observed their productivity shock, \( \varphi \), for that period.*

Let \( \pi_j \) denote a firm’s one period profit in location \( j \). The static profit maximization problem can be written as:

\[
\{n, l\} = \arg \max_{\{n, l\}} \pi_j(n, l; \varphi),
\]

where the profit function is given by:

\[
\pi_j(n, l; \varphi) = \frac{pf(\varphi, n, l; e_j)}{e_j} - wn - rl - cf.
\]

The parameter \( cf \) denotes a fixed cost of operation independent of location. Solving this problem we obtain the demand for inputs as a function of \( \varphi \), denoted by \( n_j(\varphi) \) and \( l_j(\varphi) \), as well as an indirect profit function, denoted by \( \pi_j(\varphi) \).

Let \( \mu_j \) denote the measure of firms located in \( j \). The mass of firms located located in \( j \), denoted by \( S_j \), is given by the following expression:

\[
S_j = \int \mu_j(d\varphi)
\]

\[\text{Note that the sub-index } j \text{ summarizes the dependence of the profit and input demand functions on location } j \text{’s rent and externality.}\]
Given the static choices for land and labor use for each firm, we can also calculate the aggregate levels of land and labor:

\[ L_j = \int l_j(\varphi) \mu_j(d\varphi), \]  
\[ N_j = \int n_j(\varphi) \mu_j(d\varphi) \]  

After choosing labor and land inputs, each firm faces the (dynamic) decision of whether to stay in its current location, move to the other location, or shut down. The following Bellman equations formalize the decision problem of a firm that begins the period in location \( j \) with a productivity shock \( \varphi \):

\[ V_1(\varphi) = \pi_1(\varphi) + \beta \xi \max \left\{ 0, \int V_1(\varphi') F(d\varphi' | \varphi), \int V_2(\varphi') F(d\varphi' | \varphi) - c_r(\varphi) \right\} \]  
\[ V_2(\varphi) = \pi_2(\varphi) + \beta \xi \max \left\{ 0, \int V_2(\varphi') F(d\varphi' | \varphi), \int V_1(\varphi') F(d\varphi' | \varphi) - c_r(\varphi) \right\} \]

where \( \beta \) is the discount factor, \( c_r(\varphi) \) is the cost of relocating from one location to another. We explore different specifications in the quantitative analysis. One specification assumes that relocation costs depend on the size of the firm, \( q(\varphi) \).

Solving the dynamic decision problem above implies decision rules of the following form for firms currently in location \( j \):

\[ x_j(\varphi) = \begin{cases} 
0 & \text{if firm exits in next period} \\
1 & \text{if firm chooses location 1 in next period} \\
2 & \text{if firm chooses location 2 in next period}
\end{cases} \]

To close the model, we need to specify the process of entry.

**Assumption 7** Firms can enter into both locations. All prospective entrants are ex-ante identical. Upon entering a new firm incurs a cost \( c_{ej} \) and draws a productivity
shock $\varphi$ from a distribution $\nu(\varphi)$.

Note that we allow the entry cost to vary by location. In our parametrized model, the entrant distribution is assumed to be log-normal with parameters $\mu_{\text{ent}}$ and $\sigma_{\text{ent}}^2$. These assumptions guarantee that the expected discounted profits of a prospective firm are always less or equal than the entry cost:

$$c_{ej} \geq \int V_j(\varphi) \nu(d\varphi), j = 1, 2$$  \hfill (12)

If there is positive entry of firms, then this condition holds with equality.

### 3.2 Equilibrium

We are now in a position to define a stationary equilibrium to our economy.

**Definition 1** A stationary equilibrium for this economy consists of rents, $r_j^*$, masses of entrants, $M_j^*$, stationary distributions of firms, $\mu_j^*(\varphi)$, externalities, $e_j^*$, land demand functions, $l_j^*(\varphi)$, labor demand functions, $n_j^*(\varphi)$, value functions, $V_j^*(\varphi)$, and decision rules, $x_j^*(\varphi)$, for each location $j = 1, 2$, such that:

1. The decision rules (11) for a firm’s location are optimal, in the sense that they maximize the right-hand side of equations (10).

2. The decision rules for labor and land inputs solve the firm’s static problem in (5).

3. The free entry conditions (12) are satisfied in each location, with equality if $M_j^* > 0$.

4. The market for land clears in each location consistent with equation (4).
5. The mass of firms in each location is given by equation (7).

6. The externalities are consistent with (2)

7. The distributions of firms \( \mu_j^* \) are stationary in each location and consistent with firms’ decision rules.

Any stationary equilibrium to our model can be characterized by vector of equilibrium values for rents, mass of entrants, and externalities in each location \( (r_1, r_2, M_1, M_2, e_1, e_2) \). Finding an equilibrium for this model is equivalent to the problem of finding the root of a nonlinear system of equations with six equations. For any vector \( (r_1, r_2, M_1, M_2, e_1, e_2) \), we can

1. solve the firms’ static profit maximization problem and obtain land demand, labor demand, and the indirect profit functions for each location;

2. solve the dynamic programming problem in equations (10) and obtain the optimal decision rules;

3. use the initial mass of entrants in each location and simulate the economy forward until the distribution of firms, \( \mu_j \), converges to a stationary distribution;

4. calculate the aggregates land and labor demands, as well as the land supply in the economy;

5. check whether market clearing conditions and the equations that define the mass of firms and the externalities in each location are satisfied.

If the equilibrium conditions are not satisfied, we update the vector of scalars and repeat the process until all of the conditions for equilibrium are satisfied. If this algorithm converges, we have computed an equilibrium of the model.
Note that the mapping described above is not a simple contraction mapping. As a consequence we cannot apply standard techniques and proof of existence of equilibrium. In Appendix B of the paper, we provide a proof of existence of equilibrium for a simplified version of our model in which firm productivity is constant across time. Moreover, we have computed equilibria for a large number of different specifications of our general model. We, therefore, conclude that equilibria exist for reasonable parameterizations of the model.

The task of computing an equilibrium can be simplified by exploiting some properties of the parametrization used in our computational model. The static first order condition that determines that ratio of land and labor inputs is given by:

\[
\frac{n}{l - l} = \frac{\alpha r_j}{\gamma w}
\]  

(13)

Notice that the ratio in this equation is the same for all firms in the same location \(j\). Aggregating over all firms in such location, we obtain that:

\[
\frac{N_j}{L_j - S_j l} = \frac{\alpha r_j}{\gamma w}
\]  

(14)

Equation (14) then implies an expression linking the externality, \(e_j\) in each location to that location’s rent, \(r_j\). We can, therefore, solve the Bellman equations without knowing the aggregate levels of land and labor. As a consequence we can characterize equilibrium rent values solely based on the free entry conditions expected values functions, which can be written as:

\[
EV_1(r_1, r_2) = \int V_1(\varphi) d\nu(\varphi)
\]  

(15)

\[
EV_2(r_1, r_2) = \int V_2(\varphi) d\nu(\varphi)
\]
The entry condition for location one then defines a mapping \( r_1 = \Gamma_1(r_2) \), i.e. for given \( r_2 \), \( \Gamma_1(r_2) \) is the value of \( r_1 \) such that \( EV_1(r_1, r_2) = c_e \). Similarly, we can define a mapping \( r_1 = \Gamma_2(r_2) \) for location two. These two mappings then effectively define the set of rent pairs \( \{r_1, r_2\} \), such that the two free entry conditions are satisfied with equality.

The non-linearity of the model implies that \( \Gamma_1(r_2) \) and \( \Gamma_2(r_2) \) can intersect multiple times. As a consequence, there may be more than one possible candidate values for equilibria with entry in both locations.\(^{19}\)

Next, define the ratio of entrants in the two locations as \( m = \frac{M_1}{M_2} \) and the distribution of firms standardized by the mass of entrants in locations 2 as,

\[
\hat{\mu}_j = \frac{\mu_j}{M_2}. \tag{16}
\]

The standardized stationary distributions satisfy

\[
\int_0^{\phi'} \hat{\mu}_1(dx) = \xi \int F(\phi'|\varphi) 1 \{x_1(\varphi) = 1\} \hat{\mu}_1(d\varphi) \\
+ \xi \int F(\phi'|\varphi) 1 \{x_2(\varphi) = 1\} \hat{\mu}_2(d\varphi) + m \int_0^{\phi'} \nu(dx)
\]

\[
\int_0^{\phi'} \hat{\mu}_2(dx) = \xi \int F(\phi'|\varphi) 1 \{x_1(\varphi) = 2\} \hat{\mu}_1(d\varphi) \\
+ \xi \int F(\phi'|\varphi) 1 \{x_2(\varphi) = 2\} \hat{\mu}_2(d\varphi) + \int_0^{\phi'} \nu(dx) \tag{17}
\]

where \( 1\{x_j(\varphi) = j\} \) is an indicator function equal to 1 if \( x_j \) equals \( j \) and 0 otherwise. Given a value for \( m \), forward iteration on these two equations yields the equilibrium standardized stationary distributions \( \hat{\mu}_j \), \( j = 1, 2 \).\(^{19}\)

\(^{19}\)In addition to equilibria with entry in both locations, it is also possible to have equilibria in which entry only occurs in one of the two locations.
To find the equilibrium value of $m$, substitute the aggregate demands for land in the two locations into the inverse land supply functions and take their ratios. Given that the inverse elasticity, denoted by $\delta$, is the same in both locations we then obtain:

$$\frac{r_1}{r_2} = \frac{A_1}{A_2} \left[ \int l_1(\varphi) \hat{\mu}_1(d\varphi) \right]^\delta \int l_2(\varphi) \hat{\mu}_2(d\varphi).$$

Let $r_1 = r_m(r_2; m)$ be the value of $r_1$ that clears the relative land markets given $r_2$ and $m$, keeping in mind that both the labor demand functions $l_j(\varphi)$ and the masses of firms $\hat{\mu}_j$ depend on $r_1$ and $r_2$.

We can thus conclude that all rent pairs $\{r_1, r_2\}$ that are consistent with entry in both locations are characterized by the intersection of the two functions $\Gamma^j(r_2)$. In addition, we have characterized the set of rent pairs consistent with land market clearing condition, $r_m(r_2; m)$, corresponding to different values of $m$. By analyzing these functions all together, we can completely characterize the set of triplets $\{r_1, r_2, m\}$ consistent with equilibrium in the economy. Figure 2 illustrates a (locally unique) equilibrium that arises in our model. Note that in this specification of the model firms only relocate from community 2 to community 1 in equilibrium.

Finally, the mass of entrants in location 2, $M_2$, is determined by the market clearing condition for land:

$$\left( \frac{r_2}{A_2} \right)^{\frac{1}{\delta}} = M_2 \int l_2(\varphi) \hat{\mu}_2(d\varphi),$$

Note that $M_2$ can be solved for analytically.

Characterizing additional properties of the equilibrium is difficult. In Appendix B of the paper we consider a simplified version of our model in which the productivity of firms does not change over time. Under this simplifying assumption we can analytically characterize the resulting stationary distribution of firms in equilibrium. For
Figure 2: Graphical Representation of Equilibrium
the general version of the model, additional insights can be gained using numerical methods.

With respect to uniqueness of stationary equilibrium, there are four issues. First, as is common in multi-community models, equilibrium typically exists with communities that are ex post identical. These “non-sorting” equilibria are uninteresting and easily rejected empirically. We analyze sorting equilibria here. Second, the non-convexities in the model associated with community choice preclude use of standard techniques to establish uniqueness of sorting equilibria. Third, entry conditions may not hold with equality which can give rise to equilibria with entry in only one of the two locations. In the computational analysis, we only focus on equilibria with entry in both locations. Last, the endogeneity of the firm productivity distribution in stationary equilibrium may not be unique.

While there are several sources of potential multiplicity, we find in our computational analysis that stationary (sorting) equilibria are robust. When we perturb an equilibrium that we have computed, the algorithm converges back to the original equilibrium. These computational experiments suggest that equilibrium is at least locally unique. We do not have a formal proof of local uniqueness of the sorting equilibrium.

4 Estimation

Let $\theta$ denote the parameter vector of the structural model to be estimated. Given the micro data that we observe for our sample discussed in Section 2, we could, in principle, construct a Maximum Likelihood Estimator for the parameters of the model. The basic idea behind this estimator is the following. We observe firm output, labor input, land use, and the aggregates that determine the agglomeration externality for
a random sample of firms in each location. We can, therefore, express the unobserved productivity, $\omega$, as a function of the observed variables and the unobserved parameters of the model. Moreover, we can characterize this productivity conditional on age as well as relocation, entry, and exit decisions of each firm in the sample. We have seen in Section 3 that the equilibrium of our model defines a nonlinear mapping from the parameter vector $\theta$ to the stationary distribution of firm level productivities. In particular, there exist stationary densities of firm productivities conditional on location, age, and the endogenous entry, exit, and relocation decisions of firms. We can, therefore, construct a likelihood function based on these conditional densities.\footnote{Note that one would need to account for measurement error in output to obtain a well-behaved likelihood function.} This likelihood estimator is difficult to implement in practice since the stationary distribution of firm productivities does not have an analytical solution. The corresponding density is hard to compute with high accuracy. Moreover, the revenue estimates provided by Dun and Bradstreet may be inaccurate, in particular for small firms.

We, therefore, adopt a simulated method of moments approach to estimate the parameters of our model that treats output as a latent variable.\footnote{Alternatively, we could estimate an auxiliary model using semi-nonparametric estimation as developed by Gallant and Nychka (1987) and then match the scoring functions of the auxiliary and the structural model (Gallant and Tauchen, 1989).} The estimation strategy relies on the idea that the structural model should replicate the observed joint empirical distribution function of age, facility size, and employment conditional on location choice.

The observed joint empirical distribution function of age, facility size, and employment conditional on location choice can be captured by moments that are based on histograms. In practice, these moments are constructed by placing establishments into categories, such as firms with 5 to 8 employees, 11 to 20 years old, located in the city. These type of moments then are calculated as the percentage of firms in
a given category relative to the number of establishments in the entire metropolitan area. In addition, we use the total percentage of firms in the CBD relative to the full metropolitan area.

Combine all moments used in the estimation procedure into one vector $\boldsymbol{m}_N$ and denote with $\boldsymbol{m}_S(\theta)$ their simulate counterparts where $S$ denotes the number of simulations. The orthogonality conditions are then given by

$$g_{N,S}(\theta) = \boldsymbol{m}_N - \boldsymbol{m}_S(\theta)$$

(20)

Following Hansen (1982), $\theta$ can be estimated using the following moments estimator:

$$\theta_{S,N} = \arg\min_{\theta \in \Theta} g_{S,N}(\theta)' A_N g_{S,N}(\theta)$$

(21)

for some positive semi-definite matrix $A_N$ which converges in probability to $A_0$. Since we can make the simulation error arbitrarily small, we suppress the dependence of our estimator on $S$. The estimator $\theta_N$ is a consistent estimator of $\theta_0$ and:

$$N^{1/2} (\theta_N - \theta_0) \overset{d}{\to} N(0, (\tilde{A}_0 D_0)^{-1} \tilde{A}_0 V \tilde{A}_0' (\tilde{A}_0 D_0)^{-1'})$$

(22)

where $\tilde{A}_0 = D_0' A_0$, $D_0 = E [\partial \boldsymbol{m}(\theta) / \partial \theta_0]$ and $V$ is asymptotic covariance matrix of the vector of sample moments, and $\theta_0$ denotes the parameter of the data generating process. The most efficient estimator is obtained by setting $A_N = V_N^{-1}$. In this case:

$$N^{1/2} (\theta_N - \theta_0) \overset{d}{\to} N(0, (D_0' V^{-1} D_0)^{-1})$$

(23)

Furthermore, standard J-statistics can be used for hypothesis and specification tests.\footnote{Strictly speaking, one would need to correct for the sampling error induced into the estimation procedure by the simulations. However, if the number of simulations is large, these errors will be negligible. For a review see Gourieroux and Monfort (1993).}
5 Some Preliminary Estimation Results

Table 3 reports the parameter estimates and the estimated standard errors for a variety of model specifications. Wages in our model are equal to $48,661 which corresponds the average yearly income in the financial/service sectors, based on census business patterns data. We also set $\alpha$, the labor share of to be equal to 0.65.$^{23}$ The facility supply elasticity, denoted by $\delta$, is set equal to 0.2.$^{24}$ Last, we set the exogenous exit probability to 0 (or $\xi = 1$) and the discount factor equal to 0.95, taking a year as the relevant unit of time.

We consider three different model specifications. The first model reported in column I of Table 3 assumes that relocation costs are constant and equal $200,000. The second specification reported in column II assumes that relocation costs are proportional to output. This specification accounts for the fact that large firms face higher relocation costs than small firms. The third specification uses a concave relocation cost function.

First, consider the version of the model with relocation costs equal to $200,000. We find that the estimate of the land share parameter is 0.0884. The parameter estimate of the agglomeration externality is 0.102. We have seen before that the restriction that $\theta > \gamma$ is necessary to get an equilibrium sorting pattern in which high productivity firms prefer locations with high agglomeration externalities. These parameters are estimated with high precision.

The fixed costs of operation are $52,000, or a quarter of the costs of relocating to a different community. Entry costs differ by location and are estimated to be $764,000 and $786,000, respectively. The minimum land requirement is approximately 1210

\[^{23}\text{Estimates about the land share are reported in Deckle and Eaton (1993), Adsera (2000), and Caselli and Coleman (2001).}\]

\[^{24}\text{Estimates vary for this rent elasticity of supply for office space, but are generally accepted to be significantly greater than unity. See Wheaton (1999) and Henderschott et al (1999) for estimates.}\]
Table 3: Parameter Estimates

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>(c_r)</td>
<td>200,000</td>
<td>(0.01 \times q)</td>
<td>(q^{13})</td>
</tr>
<tr>
<td>(\rho)</td>
<td>0.9773 (0.001)</td>
<td>0.9772</td>
<td>0.9770</td>
</tr>
<tr>
<td>(\sigma_{\epsilon})</td>
<td>0.137 (0.001)</td>
<td>0.132</td>
<td>0.137</td>
</tr>
<tr>
<td>(\mu_{\epsilon})</td>
<td>0.239 (0.001)</td>
<td>0.240</td>
<td>0.243</td>
</tr>
<tr>
<td>(\mu_{\text{ent}})</td>
<td>11.67 (0.001)</td>
<td>11.63</td>
<td>11.62</td>
</tr>
<tr>
<td>(\sigma_{\text{ent}})</td>
<td>0.230 (0.001)</td>
<td>0.306</td>
<td>0.343</td>
</tr>
<tr>
<td>(l_{\text{min}})</td>
<td>1210 (344)</td>
<td>1058</td>
<td>1120</td>
</tr>
<tr>
<td>(\theta)</td>
<td>0.102 (0.001)</td>
<td>0.104</td>
<td>0.110</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>0.088 (0.001)</td>
<td>0.090</td>
<td>0.093</td>
</tr>
<tr>
<td>(c_f)</td>
<td>51900 (2880)</td>
<td>52590</td>
<td>59831</td>
</tr>
<tr>
<td>(ce_1)</td>
<td>764000 (12900)</td>
<td>780000</td>
<td>844000</td>
</tr>
<tr>
<td>(ce_2)</td>
<td>786000 (13000)</td>
<td>785000</td>
<td>849000</td>
</tr>
<tr>
<td>(A_1/A_2)</td>
<td>1.6327 (0.189)</td>
<td>1.8411</td>
<td>1.7830</td>
</tr>
</tbody>
</table>
square foot with an estimate standard error of 344. The productivity shocks are highly correlated across time. The point estimate of 0.977 is consistent with previous estimates in the literature (Hopenhayn and Rogerson, 1993).

The parameter estimates are similar for the other two specifications of the model reported in columns II and III. Fixed costs and relocation costs are higher in these specifications while minimum land requirement and thus the location specific fixed costs are lower. However, the three specifications imply equilibria that differ in some qualitative and quantitative features. Table 4 reports rental price ratios and relocation patterns implied by the three model specifications.

Table 4: Rental Price Ratios and Relocation

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_r$</td>
<td>200,000</td>
<td>0.01 * q</td>
<td>$q^{24}$</td>
</tr>
<tr>
<td>ratio of rents</td>
<td>1.216</td>
<td>1.228</td>
<td>1.114</td>
</tr>
<tr>
<td>Reloc. cost per emp. ($)</td>
<td>732.18</td>
<td>813.67</td>
<td>1214.12</td>
</tr>
<tr>
<td>% est. Move (2 to 1)</td>
<td>0.03 %</td>
<td>0.09 %</td>
<td>0.07 %</td>
</tr>
<tr>
<td>% emp Move (2 to 1)</td>
<td>1.02 %</td>
<td>1.21 %</td>
<td>1.36 %</td>
</tr>
<tr>
<td>%t est. Move (1 to 2)</td>
<td>0.00 %</td>
<td>1.68 %</td>
<td>1.11 %</td>
</tr>
<tr>
<td>% emp. Move (1 to 2)</td>
<td>0.00 %</td>
<td>0.53 %</td>
<td>0.21 %</td>
</tr>
</tbody>
</table>

The equilibrium associated with the first model specification implies that the rental rate for office in the CBD is approximately 21 percent higher than the rate outside the CBD. This estimated price ratio is similar to the one reported by the Building Owners and Managers Association. The price ratio along with the estimate of the externality parameter, $\theta$, implies that firms located in the CBD receive a 2.02 percent productivity gain over firms located elsewhere due to the local agglomeration externality. For the third model specification the price ratio is smaller and we obtain a 1.2 percent productivity gain. Overall, the magnitude of these productivity gains

\textsuperscript{25}The Building Owners and Managers Association collects information on expenses and income for office space throughout North America. They report a rent ratio of 1.22 for suburban and CBD office space for the United States. See www.boma.org for more information.
are small, but not irrelevant.

**Figure 3: Stationary Distributions and Decision Rules**

The upper panel of Figure 3 plots the stationary distribution of firms in both locations as well as the distribution of entrants for specification III. Note that neither of these distributions must integrate to one since the mass of firms and entrants are equilibrium outcomes. The lower panel of Figure 3 plots the optimal decision rules. The equilibrium implies that firms with high productivity shocks relocate to the CBD while low productivity firms leave the CBD to operate in cheaper locations.
Table 4 quantifies the impact of relocation. Specification 1 of our model implies that there is only relocation of large firms from outside the CBD into the CBD. While only a small fraction of firms move to the CBD, they account for approximately one percent of employment in the metro area. In specifications II and III, the equilibrium implies that small firms will relocate from the CBD to locations outside the CBD. Approximately one percent of all small firms leave the CBD per year.

Table 5: Age-Employment distributions of establishments by location, as a percentage of total establishments in the entire county (computational moments in parenthesis)

<table>
<thead>
<tr>
<th>Inside CBD</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 to 10</td>
</tr>
<tr>
<td>1 to 3</td>
<td>1.48</td>
</tr>
<tr>
<td>Employment</td>
<td>4 to 8</td>
</tr>
<tr>
<td></td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
</tr>
<tr>
<td>9 to 16</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
</tr>
<tr>
<td>&gt; 16</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outside CBD</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 to 10</td>
</tr>
<tr>
<td>1 to 3</td>
<td>16.25</td>
</tr>
<tr>
<td>Employment</td>
<td>4 to 8</td>
</tr>
<tr>
<td></td>
<td>3.59</td>
</tr>
<tr>
<td></td>
<td>(7.86)</td>
</tr>
<tr>
<td>9 to 16</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>(2.25)</td>
</tr>
<tr>
<td>&gt; 16</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>(0.95)</td>
</tr>
</tbody>
</table>

Finally, we evaluate the within sample fit of our model. Table 5 reports the distribution of firms by age and employment size for our sample and the one predicted by our model for specification I of the model. We have seen before that the data
suggest that firms located in the city are older and larger than firms located in the rest of the metro area. Moreover, the firms in the city have more employees holding age constant. Overall, our model fits this feature of the data reasonably well.

Table 6: Age-Facility Size establishment distributions by location as percentage of total establishments in the county, (computational moments in parenthesis)

<table>
<thead>
<tr>
<th>Inside CBD</th>
<th>Age</th>
<th>1 to 10</th>
<th>11 to 20</th>
<th>21 to 30</th>
<th>&gt; 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facility (sq ft)</td>
<td>1 to 1250</td>
<td>0.82</td>
<td>0.73</td>
<td>0.44</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>1251 to 2150</td>
<td>0.64</td>
<td>1.18</td>
<td>0.59</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>2151 to 3850</td>
<td>0.75</td>
<td>1.13</td>
<td>0.70</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>&gt; 3850</td>
<td>0.60</td>
<td>0.98</td>
<td>0.91</td>
<td>1.63</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outside CBD</th>
<th>Age</th>
<th>1 to 10</th>
<th>11 to 20</th>
<th>21 to 30</th>
<th>&gt; 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facility (sq ft)</td>
<td>1 to 1250</td>
<td>9.57</td>
<td>9.76</td>
<td>5.51</td>
<td>4.20</td>
</tr>
<tr>
<td></td>
<td>1251 to 2150</td>
<td>7.08</td>
<td>7.56</td>
<td>4.86</td>
<td>4.74</td>
</tr>
<tr>
<td></td>
<td>2151 to 3850</td>
<td>3.81</td>
<td>6.01</td>
<td>4.14</td>
<td>4.18</td>
</tr>
<tr>
<td></td>
<td>&gt; 3850</td>
<td>2.30</td>
<td>4.09</td>
<td>3.45</td>
<td>5.13</td>
</tr>
</tbody>
</table>

Table 6 focuses on the age-facility size establishment distributions. Employment and facility size are highly correlated in the data as predicted by our model. As a consequence, Table 6 reinforces our previous findings about the model fit. Firms in the city face higher rental rates for land and office space. As a consequence they operate with a higher employee per land ratios. Our model explains these features of the data well.
Our analysis has some important policy implications. Relocation costs prevent establishments from moving because the gains for the individual firm are not worth the moving cost. However, this decision may not be efficient since firms ignore the external benefits of density and agglomeration to other firms when making locational decisions. A relocation subsidy may help firms to sort more efficiently in the urban economy. However, it is also possible that relocation policies are not desirable. In large metropolitan areas, there are often many independent communities that compete among each other to attract business using targeted subsidies. It is not obvious that this type of tax and subsidy competition among communities increases economic welfare.

To evaluate these types of policies, we need to measure economic welfare. In our model, establishments have zero expected profits. Hence, the most useful measure of welfare in this economy is surplus that arises to land owners in equilibrium. This surplus can be measured as the area between the rent and the land supply curve:

\[
Surplus = \int_0^{L_1^*} (r_1^* - A_1 L^\delta) \, dL + \int_0^{L_2^*} (r_2^* - A_2 L^\delta) \, dL.
\]  

(24)

– to be continued –

6 Conclusions

We have developed a new dynamic general equilibrium model to explain firm entry, exit, and relocation decisions in an urban economy with multiple locations. We have characterized the stationary distribution of firms that arises in equilibrium. The parameters of the model can be estimated using a nested fixed point algorithm by matching the observed distribution of firms by location and the one implied by our model. We have implemented the estimator using unique data collected by Dun
and Bradstreet for the Pittsburgh metropolitan area. Firms located in the central business district are older and larger than firms located outside the urban core. They use more land and labor in the production process. However, they face higher rental rates for office space which implies that they operate with a higher employee per land ratio. Our estimates imply that agglomeration externalities increase the productivity of firms by one to two percent. Economic policies that subsidize firm relocations can potentially large effects on economic growth and firm concentration in central business districts. We view the findings of this paper as promising for future research. Our model can also be used to study relocations of firms across metropolitan area. Moreover, we extend the modeling framework in a number of useful directions to analyze investment and innovation decisions.
References


A Firm Location Choices in a Sample of Large U.S. Cities

To get some additional quantitative insights into firms sorting behavior, we collected Census data for a number of metro areas. We define a business district within the metropolitan area as those zip codes within a city that have a high density of firms signifying local agglomeration. To make this concept operational, we use an employment density of at least 10,000 employees per square mile. These locations need not be contiguous, as some metropolitan areas exhibit multiple dense business districts.

Table 7: Concentration of employment in dense business districts

<table>
<thead>
<tr>
<th>MSA</th>
<th>Total Emp. Outside CBD</th>
<th>Total Emp. in CBD</th>
<th>Avg. Emp. outside CBD</th>
<th>Avg. Emp. in CBD</th>
<th>% Services outside CBD*</th>
<th>% Services in CBD**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlanta</td>
<td>1,115,398</td>
<td>229,002</td>
<td>15.79</td>
<td>29.25</td>
<td>45.24%</td>
<td>63.31%</td>
</tr>
<tr>
<td>Boston</td>
<td>1,728,075</td>
<td>531,349</td>
<td>15.66</td>
<td>39.01</td>
<td>41.99%</td>
<td>59.90%</td>
</tr>
<tr>
<td>Chicago</td>
<td>3,070,387</td>
<td>528,529</td>
<td>15.86</td>
<td>24.47</td>
<td>41.85%</td>
<td>66.50%</td>
</tr>
<tr>
<td>Columbus</td>
<td>705,534</td>
<td>63,278</td>
<td>18.69</td>
<td>23.73</td>
<td>42.88%</td>
<td>58.64%</td>
</tr>
<tr>
<td>Hartford</td>
<td>499,718</td>
<td>18,783</td>
<td>17.26</td>
<td>26.95</td>
<td>40.31%</td>
<td>61.41%</td>
</tr>
<tr>
<td>Houston</td>
<td>1,720,625</td>
<td>286,574</td>
<td>16.38</td>
<td>28.47</td>
<td>42.86%</td>
<td>65.51%</td>
</tr>
<tr>
<td>Jacksonville</td>
<td>491,959</td>
<td>24,315</td>
<td>15.24</td>
<td>25.38</td>
<td>43.09%</td>
<td>66.28%</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>4,257,269</td>
<td>974,693</td>
<td>15.02</td>
<td>19.39</td>
<td>44.16%</td>
<td>52.39%</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>1,921,626</td>
<td>196,428</td>
<td>15.91</td>
<td>27.66</td>
<td>43.99%</td>
<td>55.74%</td>
</tr>
<tr>
<td>Phoenix</td>
<td>1,551,921</td>
<td>64,793</td>
<td>18.31</td>
<td>27.78</td>
<td>47.79%</td>
<td>71.01%</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>822,013</td>
<td>157,009</td>
<td>14.58</td>
<td>40.04</td>
<td>39.16%</td>
<td>60.90%</td>
</tr>
<tr>
<td>Salt Lake</td>
<td>440,239</td>
<td>53,086</td>
<td>15.22</td>
<td>21.08</td>
<td>45.64%</td>
<td>58.90%</td>
</tr>
<tr>
<td>San Antonio</td>
<td>655,740</td>
<td>26,572</td>
<td>17.21</td>
<td>20.49</td>
<td>43.22%</td>
<td>56.59%</td>
</tr>
<tr>
<td>Seattle</td>
<td>1,260,335</td>
<td>179,230</td>
<td>14.55</td>
<td>20.33</td>
<td>42.07%</td>
<td>58.97%</td>
</tr>
<tr>
<td>St Louis</td>
<td>1,253,959</td>
<td>84,034</td>
<td>16.38</td>
<td>42.57</td>
<td>41.41%</td>
<td>52.43%</td>
</tr>
<tr>
<td>Wash. DC</td>
<td>1,930,848</td>
<td>303,770</td>
<td>15.42</td>
<td>21.68</td>
<td>49.96%</td>
<td>60.05%</td>
</tr>
</tbody>
</table>

Source: 2006 Zip Code Business Patterns, U.S. Census

*Percentage of establishments outside the CBD that are in the service industries.

**Percentage of establishments in the CBD that are in the service industries.

Table 7 shows the concentration of employment in dense business districts for a
sample of U.S. cities. First, we report statistics using all firms that are located in the metro area. We find that there is a significant amount of heterogeneity among the cities in our sample. There are some cities such as Phoenix and Hartford where employment is not concentrated in dense business districts. Most larger cities in the U.S. such as Los Angeles, Chicago, Boston, Washington, Philadelphia, and Houston have a significant fraction of firms located in high density central business districts. This finding is also true for a variety of mid-sized cities such as Pittsburgh and Seattle. Focusing on the differences between firms located in and out of the CBD, we find firms in the CBD are larger than the MSA average. This indicates that they have higher levels of productivity. This finding is common among all cities in our sample. In addition, firms in the service sector are more concentrated in the CBD compared to firms in general, suggesting that service oriented firms benefit more from local agglomeration than other sectors.

B Analytical Properties of Equilibrium

To get some additional insights into the properties of our model it is useful to simplify the structure of the model and shut down the future productivity shocks. We can then characterize the equilibrium of the model almost in closed-form.\textsuperscript{26} Let us impose the following additional assumptions.

Assumption 8

1. The shock is drawn upon entry once and for all from a uniform distribution in $[0, 1]$: $\nu(\varphi) = 1$ for $\varphi \in [0, 1]$.  

\textsuperscript{26}The model cannot be entirely solved in closed form because the equilibrium $r_2$ has to satisfy a highly non-linear equation. Sufficient conditions on the model’s parameters for $r_2$ to exist and be unique are imposed instead. Conditional on $r_2$, everything else can be solved for analytically.
2. There are no fixed cost of operation: \( c_f = 0 \).

3. Importance of externality: \( \theta = 1 - \alpha > \gamma \)

Let 1 denote the high rent location and 2 the low rent one (1=city, 2=suburb). We show how to construct a unique equilibrium in which \( r_1 > r_2 \) and firms move from location 2 to location 1, but not vice versa. Firms who enter in location 1 stay there all the time or exit.

First note that under assumptions 2 and 3 above the indirect profit functions can be written as:

\[
\pi_j(\varphi) = r_j(\Delta \varphi^n - \bar{l}), \quad j = 1, 2,
\]

where \( \Delta > 0 \) and \( \eta > 1 \) are known functions of the parameters of the model. Consider location in the city. We have the following result.

**Proposition 1** If \( r_1 > r_2 \),

a) then firms in location 1 follow a simple cut-off rule. Firms below a threshold \( \varphi_l \) exit while firms above the threshold stay in location 1 forever. The cut-off is defined as:

\[
\varphi_l = \left( \frac{\bar{l}}{\Delta} \right)^{\frac{1}{\eta}}.
\]

b) then firms in location 2 follow a simple cut-off rule. Firms below the threshold \( \varphi_l \) exit, firms with shocks between \( \varphi_l \) and \( \varphi_h \) stay in location 2, and firms with shocks larger than \( \varphi_h \) move to location 1. The cut-off \( \varphi_h \) is defined as:

\[
\varphi_h = \left( \frac{\bar{l}}{\Delta} + \frac{c_r (1 - \beta \xi)}{\Delta (r_1 - r_2)} \right)^{\frac{1}{\eta}}.
\]

Proof:

a) Note that static firm profits are monotonically increasing in \( \varphi \). Define \( \varphi_l \) such that
\( \pi_1(\varphi_l) = 0 \). Then firms with \( \varphi < \varphi_l \) exit immediately. It is straightforward to show that

\[
V_1(\varphi_l) = \pi_1(\varphi_l) + \beta \xi \max\{0, V_1(\varphi_l), V_2(\varphi_l) - c_r\} = \pi_1(\varphi_l) = 0
\]  
(29)

where the second equality follows from the fact that the productivity cut-off for switching to location 2 is less than cut-off for exit if \( r_1 > r_2 \) as assumed. Firms with \( \varphi > \varphi_l \) stay in location 1 as long as they survive the exogenous destruction shock \( \xi \). Their payoffs are:

\[
V_1(\varphi) = \frac{\pi_1(\varphi)}{1 - \beta \xi} > 0
\]  
(30)

b) Next consider the decision rule of firms located in the suburb. Firms with \( \varphi < \varphi_l \) exit immediately:

\[
V_2(\varphi_l) = 0
\]  
(31)

Firms with shocks in \((\varphi_l, \varphi_h)\) stay in 2 forever (as long as they survive the exogenous destruction shock). Firms with high shock move to 1. The indifference condition for staying vs moving is:

\[
\frac{\pi_2(\varphi_h)}{1 - \beta \xi} = \pi_2(\varphi_h) + \beta \xi (V_1(\varphi_h) - c_r).
\]  
(32)

This equation defines the cut-off value \( \varphi_h \). The lemma then follows from the result that benefits of switching to location 1 monotonically increase with \( \varphi \). Q.E.D.

Next we consider the free entry conditions and show that these conditions determine the rents in both locations. We have the following result:

**Proposition 2** There is at most one set of rental rates \((r_1, r_2)\) that are consistent with the entry in both locations. Conditions on the parameter values guarantee existence of \((r_1, r_2)\).

---

\(^{27}\)Note that equation (26) implies that \( z_l \) does not depend on the location.
Proof (of uniqueness):
First consider the free entry condition in location 1 which is given by

\[ \int V_1(\varphi) \nu(\varphi) \, d\varphi = c_e. \] (33)

Substituting in our optimal decision rule and simplifying we obtain the equilibrium rent in location 1:

\[ r_1 = \frac{c_e (1 - \beta \xi) (\eta + 1)}{\Delta (1 - \beta \xi \varphi_h^{\eta + 1}) - I (1 - \beta \xi \varphi_1) (\eta + 1)}. \] (34)

Free entry in location 2 requires:

\[ \int V_2(\varphi) \nu(\varphi) \, d\varphi = c_e. \] (35)

Replacing the value function in location 2 and taking into account the definition of \( \varphi_h \) in (28) this equation simplifies to:

\[ \eta \varphi_h^{1+\eta} - (1 + \eta) \varphi_h^{\eta} - K = 0, \] (36)

where \( K \) represents a non-positive combination of the parameters and is defined in the appendix. The left hand side of this equation is positive when \( K = 0 \) and has a negative first derivative. Thus, if a solution for \( \varphi_h \) exists it must be unique. In turn, \( \varphi_h \) is monotonically related to \( r_2 \) by equation (28):

\[ r_2 = r_1 - \frac{c_r (1 - \beta \xi)}{\Delta \varphi_h^{\eta} - I}. \] (37)

Thus, if the solution \( \varphi_h \) to equation (36) is unique, the equilibrium value of \( r_2 \) is also unique. The appendix provides sufficient conditions on the parameters for this solution to exist. Q.E.D.
Next we characterize the equilibrium distribution of firms in each location.

**Proposition 3** For each value of $M_2$, there exists a unique stationary equilibrium distributions of firms in each location.

Proof:
Without loss of generality, let us normalize the model so that entry in location 2 is always equal to $M_2 = 1$. This implies a specific choice of $A_2$. Given this the mass of firms in location 2 is $\hat{\mu}_2(\varphi)$:

$$\hat{\mu}_2(\varphi) = \begin{cases} 
  z_l & \text{if } \varphi < z_l \\
  \frac{1}{1-\xi} (z_h - z_l) & \text{if } \varphi \in [z_l, z_h] \\
  1 - z_h & \text{if } \varphi > z_h
\end{cases}$$

(38)

Note that firms in location 2 with $\varphi < z_l$ exit and there is a measure $z_l$ of them. Firms with $\varphi > z_h$ move to 1, and there is a measure $1 - z_h$ of them. Firms in the middle group $\varphi \in [z_l, z_h]$ remain in 2 forever subject to surviving the death shock $\xi$.

Let $m$ denote entry in location 1. The mass of firms in location 1 is:

$$\hat{\mu}_1(\varphi) = \begin{cases} 
  m z_l & \text{if } \varphi < z_l \\
  \frac{m}{1-\xi} (z_h - z_l) & \text{if } \varphi \in [z_l, z_h] \\
  \frac{(\xi + m)}{1-\xi} (1 - z_h) & \text{if } \varphi > z_h
\end{cases}$$

(39)

Firms in the first group exit immediately. Firms in the middle group stay in 1 forever. Firms with $\varphi > z_h$ come from 2 sources: 1. firms who entered in 1 and stayed there forever subject to death shock $m (1 - z_h) / (1 - \xi)$ plus firms who entered in location 2 last period, survived the shock and moved to 1 where they remain forever: $\xi (1 - z_h) / (1 - \xi)$. Q.E.D.

Finally, we have the following result:
Proposition 4 There is at most one value of $m$ such that the relative demand for land equals the relative supply of land. Under conditions on the parameters, $m$ is shown to exist.

Proof:

Given the equilibrium distributions, we can solve for equilibrium value for entry, denoted by $m$. Note that given the assumptions the demand for labor is:

$$l_j (\varphi) = \bar{l} + \left( \frac{\alpha}{w} \right) ^{\eta} \varphi^{\frac{1}{1-\alpha}} r_j^{\frac{1-\alpha}{\gamma}}.$$ (40)

The equilibrium value of $m$ is such that it solves the relative land equilibrium condition which can be written as

$$\int l_1 (\varphi) \hat{\mu}_1 (d\varphi) = \frac{A_2 r_1}{A_1 r_2} \int l_2 (\varphi) \hat{\mu}_2 (d\varphi)$$ (41)

where the right hand side does not depend on $m$. The left-hand side depends linearly in $m$ through the mass $\hat{\mu}_1 (\varphi)$ in an increasing way. This means that if $m$ exists it is unique.

For $m \to \infty$ the left hand side of (41) goes to infinity. For $m \to 0$ the left hand side is strictly positive. To show that it is less than the right hand side $A_1$ must be sufficiently small. Since the rest of the equilibrium is independent of $A_1$ one can always choose $A_1$ small enough in order to guarantee existence. Thus, there exists a unique value of $m$. Q.E.D.

In what follows we present the equilibrium of the model in a numerical example.

Result 1 Consider the following parameter values: $\beta = 0.5$, $\alpha = 0.65$, $\theta = 0.35$, $\xi = 0.9$, $\bar{l} = 0.01$, $\gamma = 0.01$, $\eta = 2.94$, $w = 1$, $\Delta = 0.0367$, $A_1 = 0.5$, $A_2 = 1.0$, $c_e = 0.1$, $c_r = 0.01$, $\delta = 1$. Then, the unique equilibrium of the model is characterized by the following: $\varphi_l = 0.64$, $\varphi_h = 0.69$, $r_1 = 37.18$, $r_2 = 34.98$, $m = 0.21$. 
The analysis of this section shows that there exists a unique (up to scale) equilibrium with entry in both locations. Our analysis in the previous section reinforces the notion that equilibria with entry in both locations are often locally unique.