A Dynamic Firm Location Choice Model with Agglomeration Externalities

April 21, 2008

Jeff Brinkman
Tepper School of Business
H. John Heinz III School of Public Policy
Carnegie Mellon University

Committee Members:
Daniele Coen-Pirani, Tepper School of Business
Holger Sieg, Tepper School of Business
Martin Gaynor, H. John Heinz III School of Public Policy

1. Introduction

This paper outlines a dynamic model of firm location choice, in an environment which includes both agglomeration effects and stochastic productivity shocks. The model follows closely the methods of Hopenhayn (1992) and his treatment of firm level entry and exit decisions. However, this model adds the complexity of a second location, as well as the existence of spillover effects, or agglomeration externalities. Agglomeration externalities occur when firms get benefits from being located close to other firms. These externalities, in theory, are due to knowledge spillovers, labor pool sharing, or proximity of complementary activities.

This paper seeks to explore the movement of firms in an economy with more than one location. The hypothesis is that firms change over time in terms of their need for proximity and the benefits it brings such as labor pool sharing, knowledge spillovers, or services. This would suggest that firms’ relative valuation of proximity verses low rent changes over time. This paper describes and analyses a model which uses agglomeration externalities to explain the movement of firms between two locations in an environment with random productivity shocks.

The first section of this paper gives an overview of some of the existing literature on agglomeration theory and dynamic models. This is followed by a description of the firm location model. Next, an algorithm is presented which can be used to solve for equilibrium in an economy with two locations, random productivity shocks, and agglomeration effects. Here special attention is paid to the form of the externality. It will be shown that the externality must change over the life of the firm in order to display

---

movement between locations. Finally, data is introduced to calibrate the model and analyze the level of agglomeration as well as some of the characteristics of the dynamics of firm location in the face of agglomeration externalities.

2. Background and Existing Literature

Agglomeration theory attempts to explain why economic activity tends to cluster close together geographically. It seems fairly obvious that some places are inherently better for certain types of economic activity because of innate geographic advantages, e.g. shipbuilding in Norfolk or energy companies in Texas. However, natural resources do not appear to completely explain the clustering of economic activity. One theory which might explain this clustering is that certain types of activity produce externalities which benefit firms located nearby. In other words, spatial proximity to other firms increases production. The literature on agglomeration economies is well established and it has application in urban, regional, national, and global economic development.

The effect of agglomeration externalities might be separated into two popular theories. The first theory was articulated by Marshall (1920), and suggests that agglomeration effects are due to knowledge spillovers from like industries. The second, attributed to Jacobs (1969), points to the diversity of production in cities as the driving factor in firm concentration.

Much of the previous work on agglomeration externalities has focused on the nature and extent of agglomeration. For example, Henderson (1994) analyzes whether these externalities are due to specialization or diversity, akin to the Marshall vs. Jacobs question. Rosenthal and Strange have looked at the determinants of agglomeration (2001). In terms of the dynamics of agglomeration, Puga and Duranton (2001) present a model which discusses the dynamics of growth and innovation in terms of spillover effects, using diversity as a driver of innovation.

This paper attempts to apply a general equilibrium framework to agglomeration theory using a simple two location model. The model does not try to explain the determinants of agglomeration, but instead explores the valuation and level of agglomeration, how this changes over the life of a firm, and how the change drives location choices of firms. The framework is an extension of the model presented in Hopenhyan (1992) and the solution methods are an extension of the algorithm presented in Hopenhyan and Rogerson.

---

These papers study equilibrium conditions of a dynamic model of firm entry and exit with random productivity shocks, and they use cost of entry conditions to find a stationary equilibrium with entry and exit. This paper uses the same entry conditions, but adds the complexity of a second location, and introduces agglomeration externalities to explain movement between locations.

3. Description of the Model

In the economic model, there are two locations, denoted $j = 1, 2$. In each period a firm chooses to stay where it is, relocate to the other location, or shutdown. In each time period, a firm is subject to a productivity shock, $\varphi$, which is specific to each firm and changes over time according to a Markov process with a conditional distribution $F(\varphi' | \varphi)$. Assuming that the firms are identical except for this productivity shock, the production function of a firm in location $j$ can then be written as:

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

where $q$ is the output, $n$ is labor, $l$ is land, and $e_j$ is the agglomeration externality in location $j$. The form of this externality will be discussed in further detail later, but it should depend on some aggregate measure of land and labor use in location $j$, and/or some moment of the distribution of $\varphi$ values in location $j$. Generally, we could write this externality as some exogenous function:

$$ e_j = \Theta(\mu_j, L_j, N_j) $$

where $\mu_j(\varphi)$ is the distribution of firms with productivity shock $\varphi$ in location $j$, and $L_j$ and $N_j$ are the aggregate levels of land and labor used, respectively, in location $j$.

For simplification, we will assume that output prices, $p$, and wages, $w$, are constant and determined exogenously, and therefore can be treated as parameters. The rent prices, however, are determined by a function of the amount of aggregate land used in each location, and denoted:

$$ r_1(L_1), r_2(L_2) $$

The inverse supply function could be different for each location, but should be increasing in $L$ so that equilibrium is possible. This specification of inverse land supply functions assumes that the amount of land is elastic, or that land can be used for other uses such as residential. Firms will choose to locate in close proximity to one another in order to reap

---

the benefits of the agglomeration externality, but as they do, the rent increases, offsetting the benefit.

Assuming that the firms make decisions on land and labor usage after they have already observed their productivity shock, $\varphi$, for that period, the static first order conditions for profit maximization can be used to obtain the demand for inputs as a function of $\varphi$:

\[
n(\varphi; e_j, r_j)
\]

\[
l(\varphi; e_j, r_j)
\]

and profit function:

\[
\pi(\varphi; e_j, r_j)
\]

Given the static choices for land and labor use for each firm, we can calculate the aggregate level of land and labor by the following formulas respectively.

\[
L_j = S_j \int l(\varphi; e_j, r_j) \mu_j(d\varphi)
\]

(1)

\[
N_j = S_j \int n(\varphi; e_j, r_j) \mu_j(d\varphi)
\]

(2)

Where $S_j$ is the number of firms located in location $j$. After choosing labor and land inputs, each firm faces the decision of whether to stay in its current location, move to the other location, or shut down. This decision implies the following Bellman equations.

\[
V_1(\varphi) = \pi(\varphi; e_1, r_1) + \beta \max \left\{ 0, \int V_1(\varphi') F(d\varphi' | \varphi), \int V_2(\varphi') F(d\varphi' | \varphi) - c_r \right\}
\]

(3)

and,

\[
V_2(\varphi) = \pi(\varphi; e_2, r_2) + \beta \max \left\{ 0, \int V_2(\varphi') F(d\varphi' | \varphi), \int V_1(\varphi') F(d\varphi' | \varphi) - c_r \right\}
\]

(4)

where $V_j(\varphi)$ is the value function in location $j$, $\beta$ is the discount factor, and $c_r$ is the cost of relocating from one location to another. In each period, firms with make the location choice which has the highest expected value, which implies decision rules of the following form for firms currently in location $j$: 


If the value functions are known for each location, then we can also write down conditions for entry costs. The model assumes that there is a pool of firms waiting to enter the economy that receive an initial productivity shock $\phi$, drawn from a distribution denoted by $\nu$. The expected value for firms entering the economy must be less than or equal to the entry costs. Otherwise, the mass of firms entering will be infinite. This condition can be written as:

$$c_e \geq \int V(\phi)\nu(d\phi), \ j = 1,2.$$  
(5)

If entry exists, then this condition holds with equality.

### 3.1 Definition of Equilibrium

Given this notation, we can now specify equilibrium for this economy. A stationary equilibrium for this economy consists of the following:

1. Rents in each location, $r_j^*(j = 1,2)$
2. Masses of entrants in each locations, $M_j^*(j = 1,2)$
3. Distributions of firms in the two locations, $\mu_j^*(\phi)$
4. Externalities in each location, $e_j^*$
5. Optimal decision rules for land, $l_j^*(\phi; r_j, e_j)$
6. Optimal decision rules for labor, $n_j^*(\phi; r_j, e_j)$
7. Value Functions for each location, $V_j^*(\phi)$
8. Optimal decision rules for location in the next period, $x_j^*(\phi)$

Such that:

1. The dynamic programming problem is satisfied:

$$V_1(\phi) = \pi(\phi; e_1, r_1) + \beta \max \left\{0, \int V_1(\phi')F(d\phi'|\phi), \int V_2(\phi')F(d\phi'|\phi) - c_r \right\}$$

$$V_2(\phi) = \pi(\phi; e_2, r_2) + \beta \max \left\{0, \int V_2(\phi')F(d\phi'|\phi), \int V_1(\phi')F(d\phi'|\phi) - c_r \right\}$$

2. The value functions satisfy,
\[ c_e \leq \int V_j(\varphi)\nu(d\varphi), j = 1,2 \]

3. Rents satisfy,
\[ r_j^* = r(L_j^*), j = 1,2 \]
where,
\[ L_j^* = S_j^* \int l^*(\varphi; e_j, r_j)\mu_j^*(d\varphi) \]

4. The externalities satisfy,
\[ e_j^* = \Theta(\mu_j^*, L_j^*, N_j^*), j = 1,2 \]

5. The distributions of firms \( \mu_j^*, j=1,2 \) are consistent over time and consistent with firms’ decision rules.

4. Solution Algorithm and Functional Forms

4.1 Overview of Solution Algorithm

The general method for finding equilibrium is fairly straight forward. The first step is to guess a scalar vector of equilibrium values for rents, mass of entrants, and externalities in each location \((r_1, r_2, M_1, M_2, e_1, e_2)\). Given the rents and externalities, we can first solve for the firms static choices for land and labor and the indirect profit functions for each location. We can then solve the dynamic programming problem by choosing an initial value function and iterating until the value function satisfies the dynamic programming problem from equations (1) and (2). The value functions also imply the location decision rules.

Once we have firm’s decision rules for land, labor, and location, we guess an initial mass of entrants in each location and simulate the economy until the distribution of firms, \( \mu_j \), does not change over time. Once \( \mu_j \) is found we can then calculate the aggregates in the economy and check to make sure that conditions 2, 3, and 4 in the definition of equilibrium above are satisfied. If they are not, we update the vector of scalars and repeat the process until all of the conditions for equilibrium are satisfied.

While the above description gives a general method for solution of the equilibrium of the economy, it turns out that careful choices for the functional forms allow for more efficient implementation of the solution algorithm. More detail on the solution algorithm is given in section 5.
4.2 Functional Forms

While the solution algorithm above is relatively straight forward, it is computationally cumbersome, even for this simplified model. The above algorithm requires that we simulate a large number of firms and search over a six dimensional parameter space. The solution algorithm can be streamlined through careful choice of functional forms. The description of the algorithm assumes that we make choices for the following functional forms.

1. The production function,
   \[ q = f(\varphi, n, l, e_j) \]

2. The externalities
   \[ e_j = \Theta(\mu_j, L_j, N_j) \]

3. The land supply functions,
   \[ r_j(L_j) \]

Rosenthal and Strange (2003)\(^9\) provide some thoughts on the nature of the externality. They suggest that the externality acts as a multiplier on the production function, and we maintain this assumption. This assumption implies the following functional form for the profit function, using a standard Cobb-Douglas production function, with parameters \(\alpha\) and \(\gamma\), and fixed cost \(c_f\).

\[
\pi_j = \varphi e_j n^\alpha l^\gamma - wn - r_j l - c_f
\]

By solving the static profit maximization problem for a firm, it can be shown that the ratio of labor and land is given by:

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

This is convenient because the ratio is not dependent on the shock \(\varphi\) or the externality \(e_j\). Appendix 1 shows the calculation of these first order conditions.

Keeping this in mind, let us turn to the question of the functional form of the externality. As mentioned earlier, we expect that the functional form of the externality should be dependent on aggregate levels of land and labor, and/or moments of the distribution of

---

shocks $\varphi$. Since agglomeration theory suggests that increased concentration of productivity in a location leads to external spillover effects, it is appealing to think that the externality depends on the density of employment given by the ratio of the aggregate levels of employment and land use, $N/L$.

If this is true, the externalities, $e_j$, can be solved directly from the rents, $r_j$. This has two distinct advantages in the solution algorithm. The first advantage is that the rents can now be solved for without simulating the economy to find the aggregates. Note that the value function iteration implied in equations (3) and (4) only requires calculation of the profit function, which, given the above specification for the externality, is not dependent on the aggregates. Second, this specification effectively removes two variables from the solution algorithm, because the externalities are found directly from the rents, which significantly decreases computational burden. If the externality and production functions can be written in this form, the solution algorithm is effectively reduced from searching for one 6-dimensional vector to searching for two 2-dimensional vectors separately. This is a significant advantage. Writing the externality as a function of rents, $r_j$, we get:

$$e_j = \Theta(\mu_j, L_j, N_j) = \Theta(\frac{N_j}{L_j}) = \Theta(\frac{\alpha r_j}{\gamma w})$$

or,

$$e_j = e(r_j)$$

Where $e(r_j)$ is some function of the ratio of land to labor, which is dependent only on the rent and other constant parameters. Furthermore, it can be shown that if the externality is multiplicative, solving the first order conditions of the profit maximization leads to the following indirect profit function:

$$\pi_j = A(\varphi e(r_j))^{\frac{1}{1-\alpha-\gamma}} r_j^{-\gamma} (1-\alpha-\gamma) - c_f$$  \hspace{1cm} (8)$$

where $A$ is a constant that depends only on parameters $\gamma$ and $\alpha$, and wage $w$. The derivation of this indirect profit function is shown in Appendix 2.

There are several other requirements for the externality that might not be as obvious. We would expect that the economy has certain characteristics. First, we expect that some firms enter into both locations. Second, we expect that at least some firms will move from one location to the other. In order for this to be true we need the functional form externality to have certain characteristics.
Consider the following functional form of the externality, which is consistent with Lucas and Rossi-Hansberg’s (2002)\textsuperscript{10} treatment of agglomeration externalities and maintains the structure of the Cobb-Douglass production function, is:

\[
e(r_j) = \left( \frac{N_j}{L_j} \right)^\theta = \left( \frac{\alpha r_j}{\gamma w} \right)^\theta.
\]

This leads to the flowing indirect profit function.

\[
\pi_j = A \left( \varphi \left( \frac{\alpha r_j}{\gamma w} \right)^\theta \right) \frac{1}{1-\alpha-\gamma} r_{j}^{-\gamma} - c_f
\]

At first glance, this form of the externality seems reasonable. As the employment density increases, the externality increases. However, note that unless \(\theta\) is equal to \(\gamma\), firms always prefer one location to the other. Furthermore, this specification does not allow for firms to move between locations. In order for this model to allow for movement between locations, it must be that the importance of the externality to the firm is dependent in the productivity shocks. That is to say, that \(\theta\) is dependent on \(\varphi\). This leads to following specification:

\[
e(r_j) = \left( \frac{N_j}{L_j} \right)^{\theta_1-\varphi} = \left( \frac{\alpha r_j}{\gamma w} \right)^{\theta_1-\varphi}
\]

and the following indirect profit function,

\[
\pi_j = A \left( \varphi \left( \frac{\alpha r_j}{\gamma w} \right)^{\theta_1-\varphi} \right) \frac{1}{1-\alpha-\gamma} r_{j}^{-\gamma} - c_f
\]

This specification is appealing in that it is consistent with the initial intuition that as firms grow (\(\varphi\) increases) the relative valuation of the externality and rents changes. It will be shown that this specification allows for equilibrium where there is entry, exit, and movement between locations.

The final functional form choice that needs to be made is the choice of the inverse land supply function. A reasonable assumption is that as the amount of land consumed for commercial use increases, the rents increase. This suggests the following functional form.

\[ r_j = A_j L_j^\delta, \quad j = 1, 2 \]  

(11)

where \( L_j \) is the aggregate land use in each location and \( A_j \) and \( \delta \) are parameters. In this form, notice that the parameter, \( \delta \), is the same in both locations. This assumption simplifies the solution of the stationary distribution \( \mu \) in each location. First note that the stationary distribution satisfies,

\[
S_1 \mu_1 (\varphi') = S_1 \int F(\varphi'| \varphi) 1 \{ x_1 (\varphi) = 1 \} \mu_1 (d\varphi) + S_2 \int F(\varphi'| \varphi) 1 \{ x_2 (\varphi) = 1 \} \mu_2 (d\varphi) + M_1 \nu (\varphi') \\
S_2 \mu_2 (\varphi') = S_1 \int F(\varphi'| \varphi) 1 \{ x_1 (\varphi) = 2 \} \mu_1 (d\varphi) + S_2 \int F(\varphi'| \varphi) 1 \{ x_2 (\varphi) = 2 \} \mu_2 (d\varphi) + M_2 \nu (\varphi')
\]  

(12)

Recall that the aggregate land use can be calculated with the following:

\[ L_j = S_j \int l(\varphi; e_j, r_j) \mu_j (d\varphi) = M_2 \int l(\varphi; e_j, r_j) \hat{\mu}_j (d\varphi) \]

Define,

\[ s_1 = \frac{S_1}{M_2}, s_2 = \frac{S_2}{M_2}, m = \frac{M_1}{M_2} \]

and define the “adjusted distribution” as,

\[ \hat{\mu}_j = s_j \mu_j. \]

Given these definitions we are now able to solve for the aggregates of the economy by first solving for \( m \), and then solving for \( M_2 \). Using equation (12) above, along with the conditions,

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

and,
\[ M_2 = \left( r_1 \frac{A_1}{A_2} \right)^{\beta} \int l(\varphi; r_2) \hat{\mu}_2(d\varphi) \] 

(13)

By using this functional form for the inverse land supply, we are effectively able to reduce the problem to solving for one variable, \( m \), numerically, while solving the other, \( M_2 \), analytically. This is a great reduction in computational burden. The precise methods used to solve for equilibrium are outlined in the next section.

5. Algorithm and Results

In this section a detailed description of the algorithm and simulation is provided in order to clarify and examine the model. The particular parameter values used for the plots were chosen so as to demonstrate the solution of equilibrium with entry, exit, and movement between locations. In fact, it is not always possible to find equilibrium for arbitrary parameter values.

The solution method consists of three separate parts. The first part simply simulates the AR(1) process in order to obtain a transition probability matrix for discretized \( \varphi \) values on a grid. The second step finds the decision rules of the firms and the rents in each location. This is done by solving for the value functions iteratively, and then checking the entry conditions from equation (5). The final step is to find the aggregates of the economy including, flows into, out of, and between the two locations, and the stationary distribution of firms in each location. MatLab was used to perform the simulations and computations, and the code is included in the appendix.

5.1 Finding the Transition Matrix

The first step is to simulate the AR(1) process. This is done by choosing a grid of \( \varphi \) values which covers the bulk of the distribution of the AR(1) process. Then simulate a large number of \( \varphi' \) values for each point on the grid using:

\[ \log \varphi' = \rho \log \varphi + \varepsilon' \]

where \( \rho \) is the correlation between consecutive shocks, and \( \varepsilon \) is a normally distributed random draw with parameters \( \mu_\varepsilon \) and \( \sigma_\varepsilon \). The \( \varphi' \) values for each \( \varphi \) are then put into bins centered around the points on the grid. This gives a discretized \( n \) by \( n \) transition matrix,
\[ F(d\varphi|\varphi) \]

where \( n \) is the number of grid points. The following plot shows a typical transition matrix. Each curve represents a conditional distribution of shocks in the next period for a given current shock.

![Transition Distributions](image)

Figure 1

### 5.2 Finding the Rents

The next step is to use the following discretized dynamic programming problem, which is analogous to the continuous problem presented in equations (3) and (4),

\[
V_1(\varphi) = \pi(\varphi; r_1) + \beta \max \left\{ 0, \sum_n V_1(\varphi') F(d\varphi'|\varphi), \sum_n V_2(\varphi') F(d\varphi'|\varphi) - c_r \right\}
\]

and,

\[
V_2(\varphi) = \pi(\varphi; r_2) + \beta \max \left\{ 0, \sum_n V_1(\varphi') F(d\varphi'|\varphi) - c_r, \sum_n V_2(\varphi') F(d\varphi'|\varphi) \right\}
\]

(14)

(15)
as well as the cost of entry condition,

\[ c_e \geq \int V_j(\varphi)\nu(d\varphi), \quad j = 1,2 \]

to find the rents for each location and the decision rules of the firms.

The process is to first guess rents in each location and then iterate recursively until the two functions converge to stable values. Then the expected value for entrants is calculated in each location. These expected values are compared to the entry costs. The rents are then updated repeatedly until the entry condition holds.

Figure 2 – Equilibrium value functions

Figure 3 shows a set of value functions in equilibrium along with the entrant distribution. Note that in equilibrium, the expected value of entrants must be equal in each location. Therefore the integral of the value functions multiplied by the entrant distribution must be equal.
5.3 Finding Stationary Distributions and Flows

The next step is to guess an initial ratio of entrants, $m$, and use the decision rules and transition matrix to solve the following equations,

\[
\hat{\mu}_1(\varphi') = \sum_n F(d\varphi'|\varphi)1\{x_1(\varphi) = 1\} \hat{\mu}_1(\varphi') + \sum_n F(d\varphi'|\varphi)1\{x_2(\varphi) = 1\} \hat{\mu}_2(\varphi') + m\nu(\varphi)
\]

\[
\hat{\mu}_2(\varphi') = \sum_n F(d\varphi'|\varphi)1\{x_1(\varphi) = 2\} \hat{\mu}_1(\varphi') + \sum_n F(d\varphi'|\varphi)1\{x_2(\varphi) = 2\} \hat{\mu}_2(\varphi') + \nu(\varphi)
\]

(16)

where $x_j(\varphi)$ are decision rules for locations $j = 1, 2$, and $\hat{\mu}_j(\varphi)$ is an “adjusted distribution” of $\varphi$ in each location which is defined as the distribution $\mu(\varphi)$ multiplied by a scale factor $s_j$. These equations are the discretized and normalized (divided by $M_2$) versions of
equation (12). Once this adjusted distribution is found, then it needs to be checked against the following condition,

\[
\frac{r_1}{r_2} = \frac{A_1}{A_2} \left[ \int l(\varphi; r_1) \mu_1(d\varphi) \right]^{\delta} \left[ \int l(\varphi; r_2) \mu_2(d\varphi) \right]
\]

Where \( l(\varphi; r_j) \) is the static land use decision for firms in each location. If the condition does not hold, then \( m \) is updated until equilibrium is found. Keep in mind that these integrals are done numerically in practice. Once the ratio, \( m \), is found, then we can solve for all of the flows and distributions in the model. The following equations are used for those calculations.

The mass of entrants can be found by,

\[
M_2 = \left( \frac{r_1}{A_1} \right)^{\delta} \left[ \int \frac{1}{l(\varphi; r_2) \mu_2(d\varphi)} \right]
\]

\[M_1 = mM_2\]

\( S_j \), the total number of firms in each location, can be found by integrating over the adjusted distribution and multiplying by \( M_2 \),

\[S_j = M_2 \int \mu_j(d\varphi)\]  

(17)

The total land use in each location can be found by,

\[L_j = M_2 \int l(\varphi; r_j) \mu_j(d\varphi)\]  

(18)

The total employment in each location is given by,

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

(19)

The average employment in each location is given by,

\[Mean \ Employment = \frac{N_j}{S_j}, j = 1,2\]  

(20)
The flow from location 1 to location 2 is given by,

\[ \text{flow}_{12} = M_2 \int 1\{x_1(\varphi) = 2\} \hat{\mu}_1(\varphi) \, d\varphi \]  
(21)

The flow from location 2 to location 1 is given by,

\[ \text{flow}_{21} = M_2 \int 1\{x_2(\varphi) = 1\} \hat{\mu}_2(\varphi) \, d\varphi \]  
(22)

The mass of shutdowns from location 1 is given by,

\[ \text{flow}_{10} = M_2 \int 1\{x_1(\varphi) = 0\} \hat{\mu}_1(\varphi) \, d\varphi \]  
(23)

The mass of shutdowns from location 2 is given by,

\[ \text{flow}_{20} = M_2 \int 1\{x_2(\varphi) = 0\} \hat{\mu}_2(\varphi) \, d\varphi \]  
(24)

Finally, the average land use per firm in each location is given by,

\[ \text{Mean}(l_j) = \frac{L_j}{S_j} \]  
(25)

The plot below demonstrates how firms might be distributed in an economy under the assumptions of the model. Note that shock distributions do not translate directly into size distributions, because the firms employment choices will be different in the two locations given the difference in rents. In other words, in the lower rent area, firms could have higher shock values, but because of the cheaper land, they will use more land relative to workers.
6. Data and Calibration

This section seeks to analyze the observed data in terms of the above model. First, the data shows how much economic activity is concentrated spatially. Second and perhaps more interestingly, the data can tell us some things about the scale and importance of the agglomeration externality, as well as how this externality changes as a firm grows.

Before beginning this task, I would like to submit one caveat. The strongest assumption of this model is that firms are identical. However, firms clearly have different needs for land and labor, and also presumably for the externality we are discussing. Given this heterogeneity, there are no doubt compositional effects in the aggregation of data, particularly given the fact that we are concerned about spatial distribution. The ideal data set could overcome this problem, but given the abstract nature of this model, along with the complexity of firm behavior over time and space, it is more fruitful to develop a general understanding of the distribution and dynamics of economic activity spatially in terms of land use, labor, and prices. Furthermore, some intuition about compositional effects can be obtained by looking at data from different sectors. Despite these difficulties, the data and calibration exercise described below, gives clear insight into the level, relative scale, and dynamics of agglomeration externalities.
6.1 The Data

For this analysis, I focus on Allegheny County, Pennsylvania in 2004. In order to analyze the data in terms of the model, I defined the boundaries for two separate locations. Location 1 is the City of Pittsburgh, and location 2 is the rest of Allegheny County. Because of the need to use zip code level data, and the fact that zip codes do not align directly with city boundaries some judgments were made. The city boundary is used for several reasons. First, we want the two locations to differ significantly in both density and price. Second, using the city boundary provides a little more assurance that zoning laws, development incentives, tax rates, and other government activity are at least consistent within the locations. Finally, the model assumes that the two locations are infinitely far apart and infinitely compact. In realistic terms, we want high transportation and transaction costs between the locations and low transportation and transaction costs within locations. These two locations satisfy these requirements relatively well.

The data is drawn from two main sources. First, I use commercial property value data from the Allegheny County Assessors Office.11 This data provides the total amount of commercial property in each location, and also gives an idea of the relative value of land in each location. The rest of the data comes from Census Bureau data sets. The Zip Code Business Patterns data provides the number of establishments, total payroll, and employment size categories by sector and by zip code.12 The number and distribution of entrants and exits are taken from the Business Dynamics data.13

6.2 Characteristics of the Economy

In order to gain intuition about the effect of agglomeration externalities, it is important to understand the level and scale of concentration of economic activity geographically. Furthermore, it is necessary to understand how this concentration affects prices. While the fact that there is higher density and higher prices in the city is well known, the scale of this concentration and the relative value of rents might be surprising.

Figure 6 shows the employment by zip code in Allegheny County. This map clearly shows that employment is heavily concentrated in the center of the city. In fact, three zip codes near the center of the city with relatively small land areas account for 160,520 employees, or approximately 23% of the total employment.

---

11 Allegheny County Real Estate Website, http://www2.county.allegheny.pa.us/realestate/, 2004
12 Zip Code Business Patterns, US Census Bureau Website http://www.census.gov/econ/census02/guide/g02zip.htm 2004
Figure 5 - Employment in Allegheny County by zip code 2004

Figure 7 shows the assessed value per square ft. of commercial land by zip code. Note that these assessed values were not used in the analysis because of the difficulty of relating assessed values to real rents, but the relative values should be similar. Again, the center city neighborhoods have land values many times higher than in suburban areas. In
fact, the values range from $22 per sq. ft. in the central business district, to $1 per sq. ft. or less in many of the outlying areas. Table 1 summarizes the characteristics of the two locations.

Figure 6 - Assessed land value in Allegheny County
Table 1 - Aggregate characteristics of locations

<table>
<thead>
<tr>
<th></th>
<th>Pittsburgh City</th>
<th>Allegheny County - Not Pittsburgh</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Employment</strong></td>
<td>247,107</td>
<td>449,949</td>
</tr>
<tr>
<td><strong>Number of Firms</strong></td>
<td>9,400</td>
<td>28,801</td>
</tr>
<tr>
<td><strong>Average Firm Size</strong></td>
<td>26.29</td>
<td>15.62</td>
</tr>
<tr>
<td><strong>Total Commercial Land (sq. ft.)</strong></td>
<td>142,372,010</td>
<td>1,748,214,558</td>
</tr>
<tr>
<td><strong>Average Price (per sq. ft.)</strong></td>
<td>$5.93</td>
<td>$1.32</td>
</tr>
<tr>
<td><strong>Employment Density (employment per sq mile commercial land)</strong></td>
<td>48,387</td>
<td>7,175</td>
</tr>
</tbody>
</table>

In order to examine the model we also need some measure of the distribution of firms in the economy. Specifically, we need to know the distribution of productivity shocks in the economy, the distribution of entrants, and the AR(1) shock process. Clearly, these cannot be observed directly, but we can observe firm size distributions which can be related to the shock values for a given parameter set. Figure 7 shows the distributions of establishment sizes in each location. The graphs are normalized to show the percentage of firms in each category so that the two locations can be compared.

From this graph it can be seen that in the city, the firms are larger than in the county, although the distributions are very similar. This would suggest that the need for the externality increases as firms grow. This is not entirely surprising, as certain types of businesses would presumably need to be close to other firms. For example, large financial firms, law firms, or corporate headquarters would need access to one another. However, I would caution that the results above are probably more a compositional
effect. For evidence of this consider the manufacturing and financial sectors, shown in figures 8 and 9 respectively. In manufacturing, smaller establishments are much more concentrated in the city, and all four establishments with 1000 plus employees are located outside of the city. Conversely, larger firms in the finance and insurance industry tend to prefer the city with five of six establishments with over 1000 employees located in the city.

![Figure 8](image)

**Figure 8 – Percentage of manufacturing establishments in each location in each employment size category**

![Figure 9](image)

**Figure 9 - Percentage of finance and insurance establishments in each location and each size category**

The above graphs show that there are clearly differences between types of firms in terms of how they value proximity relative to rent. However, the above figures are also encouraging, in that they show a clear trend, at least within sectors, toward a change in
valuation of agglomeration externalities over time. Due to lack of sufficient data for each sector, the rest of the data analysis and calibration presented here is carried out for the economy as a whole, but the compositional effects should be kept in mind.

One more useful data set available through the Census Bureau is firm entry and exit dynamics. The entry and exit of firms in Allegheny County is shown in Table 2 for the years available.

<table>
<thead>
<tr>
<th>Year</th>
<th>Births</th>
<th>Deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>2,946</td>
<td>2,958</td>
</tr>
<tr>
<td>2000</td>
<td>2,821</td>
<td>2,857</td>
</tr>
<tr>
<td>2001</td>
<td>2,920</td>
<td>2,989</td>
</tr>
<tr>
<td>2002</td>
<td>3,327</td>
<td>3,416</td>
</tr>
<tr>
<td>2003</td>
<td>2,871</td>
<td>2,989</td>
</tr>
</tbody>
</table>

The number of births closely matches the number of deaths, and the data is consistent over time with an average of 3041.8. This result is encouraging with respect to the model. Unfortunately, data on the size distribution is not available at the county level, however, we do know that the average entrant in the MSA has 10.65 employees and the average exit has approximately 11.36 employees in 2004. Both of these numbers are considerably lower than the average employment size in the county which is approximately 19.8. Furthermore, data is available for the distribution of entrants on the national level. Figure 10 shows the distribution of entrants relative to the distribution of firms in the county.

The sizes of entering and exiting firms are significantly smaller than firms in general, while the exiting firms are slightly larger than the entering firms on average. This
difference between births and deaths turns out to be important, because it allows for identification of the shock process, which will be discussed further in the next section.

These are the overall characteristics of the economy. Missing from this data set is the movement of firms between locations. Nonetheless, these basic statistics alone provide substantial evidence that agglomeration externalities exists, and modest evidence that these externalities depend in some way on characteristics of firms that change over time.

6.3 Calibration and the Agglomeration externality

This section seeks to describe the dynamics of the economy by performing a rough calibration of the parameters in the model. While the lack of ideal data creates difficulties, calibration can tell us something about the importance of the agglomeration externality and also describes methods which could be used in further study given ideal data, or to analyze the differences between sectors in terms of the importance of agglomeration externalities.

The parameter set for this model is large and includes all of the moments of the distributions in each location and for entrants and exits. Also, it includes the parameters in the AR(1) process, three unknown cost parameters, production function parameters, and the discount factor. Finally, in addition to the data described above we need to have information on wages and rents.

In order to make this task a little more tractable, I start by using some parameter values from the literature which are rather uncontroversial and not directly the focus of this paper. First, the production function parameters for land and labor ($\gamma$ and $\alpha$) are set to be .06 and .65 respectively, based on values used in previous literature. I use .95 for the discount factor, and .95 for the correlation parameter in the shock process, given a 1-year period. The wage is the average yearly wage in the county, and is taken directly from the census data.

All of the other parameters are found by matching the model and algorithm to observed data points in the economy. The rents are set ahead of time using the first order conditions which imply that,

$$r = \frac{L \gamma}{w N \alpha},$$

where $L$ and $N$ are the observed total land and labor use in a given location.

The rest of the parameters are found within a search algorithm in order to insure that the parameters provide the best fit to the data and to the conditions for equilibrium within the model. The basic calibration algorithm is as follows.

---

14 The labor share used regularly in the literature and generally ranges from .6 to .8.
1. **Guess the externality parameters, $\theta_1$ and $\theta_2$**

In this step it is important to note that we already know something about the approximate level of $\theta_1$ and $\theta_2$. Recall from the indirect profit function (equation 10) that the profit function will cross when $\theta_1 + \theta_2 \varphi = \gamma$. We need the profit functions to cross somewhere in the middle of the distribution or else firms will always prefer one location to the other. Furthermore the entry conditions require that the value functions (which are closely related to profits) cross somewhere in the heart of the shock distribution. This makes initial choices for these parameters a little easier.

2. **Use the first order conditions to map labor choices to the shock values**

The mapping from shock values to labor is given by:

$$
I = \left[ \left( \frac{\alpha \gamma}{\gamma w} \right)^{1-\alpha} \left( \frac{w}{\varphi \alpha} \right) \right]^{\frac{1}{\alpha-1+\gamma}}
$$

and

$$
N = l \frac{\alpha r_j}{\gamma w}
$$

Note that the function does not have a global inverse, so it is important to understand the region of $\varphi$ values in which the algorithm is working.

3. **Find distributions of shock values which best match the observed distributions in the economy.**

The distributions are matched by assuming that the shocks are distributed log normal and finding the parameters of the distribution which best match the observed distributions for entrants, exit, and each location.

4. **Use the distributions to determine the shock process**

Identification of the shock process is dependent on setting the correlation coefficient ahead of time. With this, we can use stationarity of the distributions to get the following relationships.

$$
\hat{\mu}(\phi') = \hat{\mu}(\phi) \sum_n F(d \phi' | \phi) - \mathbf{1} \{x_1(\phi) = 0\} \hat{\mu}(\phi) \sum_n F(d \phi' | \phi) + M \nu(\phi')
$$

$$
\hat{\mu}(\phi') = \hat{\mu}(\phi)
$$
In the above equations the total distribution, distribution of entrants, and distribution of exits are known. This leaves only the transition matrix which is known analytically given two unknown shock parameters. We can then search for these unknown parameters.

5. **Update entry costs, relocation costs, and fixed costs which best fit the entry conditions and mass of entrants given calculated value functions and decision rules.**

The costs parameters are found using the original algorithm which finds the value function and decision rules. Using the value functions and the entry distribution we can check that the entry conditions hold in both locations, given by equation 5. This stage of the algorithm also checks to make sure that the decision rules are consistent with the total number of firms exiting the economy given the stationary distributions.

6. **Repeat until the parameters converge.**

### 6.4 Calibration Results

Table 3 displays calibrated parameters of interest in addition to calculated rents. There are several results of note in this table. First, although it is hard to develop an intuition about the absolute values of rent, it is encouraging that the ratios of the calculated rents are very similar to the assessed land values, despite the fact that they were determined separately. The two externality parameters are appealing in the sense that they are on the same scale as the land share parameter. Furthermore, as suspected from the distributions $\theta_2$ is positive, meaning that on average larger firms desire places with more employment density.

<table>
<thead>
<tr>
<th>Table 3 - Calibrated Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter</strong></td>
</tr>
<tr>
<td>Rent 1 (city)</td>
</tr>
<tr>
<td>Rent 2</td>
</tr>
<tr>
<td>$\theta_1$</td>
</tr>
<tr>
<td>$\theta_2$</td>
</tr>
<tr>
<td>$\alpha$</td>
</tr>
<tr>
<td>$\gamma$</td>
</tr>
<tr>
<td>Entry Cost</td>
</tr>
<tr>
<td>Fixed Cost</td>
</tr>
<tr>
<td>Relocation Cost</td>
</tr>
<tr>
<td>$\mu_{\epsilon}$</td>
</tr>
<tr>
<td>$\sigma_{\epsilon}$</td>
</tr>
</tbody>
</table>
The costs parameters are more difficult to interpret. It would be expected that the entry costs and fixed costs would be on a much larger scale than relocation costs. However, I suspect that the relocation cost parameter estimate is low, probably due to imperfect data. Finally, the AR(1) process parameters are displayed. The mean parameter is certainly consistent with the average firm size in the economy. However, the variance parameter is probably much higher than suspected. This again is likely due to the assumption of identical firms. It should be noted that the variance of the shock process and the relocation costs may not be identified without more data on movement between locations. These two parameters simultaneously determine the overlap of the distributions in the two locations. This also may be an explanation for the unexpected estimates of these two parameters.

The distributions and dynamics of the economy can be better understood graphically. Figure 11 shows the distribution of shocks for entrants, exits, and for each location. In addition, the decision rules are overlaid on the distributions. The decision rules are the same for each location given the small relocation costs. In other words, firms will always decide to move if it will increase their expected values.

![Figure 11 - Distributions of shocks and decision rules](image)

In order to interpret this graph, first consider the different distributions. Note that the entrants and exits are distributed over considerably smaller shock values than the total
distributions in each location. Also note as the shocks get bigger, more firms are located in the city. The decision rules show that shock values at which firms switch from one location to another. They will exit at very low shock values, and will switch between locations at a value of approximately $1.34 \times 10^5$. This is appealing in the sense that the distributions are cut in a way that matches the total number of firms in each location. However, the decision rules coupled with low relocation costs and high variance in the shock process would lead to very high movement between locations, which is unlikely in the real economy. The reasons for this anomaly were described earlier and are most likely a combination of simultaneity, non-identical firms, and lack of data on relocation. Overall, however, these results are encouraging in the sense that they show that the model can explain movement, the level of the externality, as well as capture the total mass of firms that wish to locate in each location, enter, and exit the economy.

7. Conclusions and Further Study

The above results demonstrate the computation of equilibrium for an economy with two locations, stochastic productivity shocks, and agglomeration externalities. The results also show that agglomeration effects might partially explain the movement of firms between locations. While this model is relatively simple, it points toward new ways of studying firm movement as well as opening up new methods for studying agglomeration effects.

This model supports the underlying theory that as firms grow the relative valuations of low rents and high agglomeration externalities change. An analysis of the data points to the fact that there is heavy concentration of economic activity geographically, and that this concentration presents itself in the land rents for different locations. Furthermore, the calibration exercise demonstrates the level of agglomeration externalities as well as how firms might value this externality as their productivity changes. Overall, there is evidence that this model could be used to further study differences in sectors in terms of need for spillover effects as well as the composition of firms within a metropolitan area.

There are several paths which might be taken to extend this research. The first is to add complexity to the model. This could include adding heterogeneity to firms, focusing on individual sectors, making labor markets endogenous, or adding to the number of jurisdictions in order to better capture the spatial distribution of economic activity. Furthermore, longitudinal firm level data could help to better understand the actual decisions of firms in terms of land, labor, and location, and how these decisions relate to agglomeration externalities.
Appendices

A.1 Derivation of static land and labor decisions in location j

Profit function:
\[ \pi_j = \varphi \alpha n^{-1} l^{\gamma} - wn - r_j l - c_f \]

First Order Conditions with respect to:
\[ n : \quad \varphi \alpha n^{-1} l^{\gamma} = w \]
\[ l : \quad \varphi \gamma n^{\alpha-1} l^{\gamma} = r_j \]

Solve:
\[ \frac{\alpha}{\gamma} = \frac{wn^{\alpha-1} l^{\gamma}}{r_j n^{\alpha-1} l^{\gamma}} \Rightarrow \frac{n = \alpha r_j}{l = \gamma w} \]
\[ n = l \frac{\alpha r_j}{\gamma w} \]

substitute above;
\[ \left( \frac{\alpha r_j}{\gamma w} \right)^{\alpha-1} l^{\gamma} = \frac{w}{\varphi \alpha} \]
\[ l^{\alpha-1+\gamma} = \left( \frac{\alpha r_j}{\gamma w} \right)^{\alpha-1} \left( \frac{w}{\varphi \alpha} \right) \]
\[ l = \left( \frac{\alpha r_j}{\gamma w} \right)^{1-\alpha} \left( \frac{w}{\varphi \alpha} \right) \]
and
\[ n = l \frac{\alpha r_j}{\gamma w} \]
A.2 Derivation of Indirect Profit Function

Using the static choices for land and labor above, we can derive the indirect profit function.

\[
l = \left[ \left( \frac{\alpha r_j}{\gamma w} \right)^{1-\alpha} \left( \frac{w}{\phi e_j \alpha} \right)^{1} \right]^{\frac{1}{\alpha-1+\gamma}}
\]

\[
l = \left( \frac{\alpha r_j}{\gamma w} \right)^{\frac{1-\alpha}{\alpha-1+\gamma}} \left( \frac{w}{\phi e_j \alpha} \right)^{\frac{1}{\alpha-1+\gamma}}
\]

\[
l = \left( \frac{\alpha}{\gamma w} \right)^{\frac{1-\alpha}{\alpha-1+\gamma}} r_j^{\frac{1-\alpha}{\alpha-1+\gamma}} \left( \frac{w}{\alpha} \right)^{\frac{1}{\alpha-1+\gamma}} \left( \phi e_j \right)^{\frac{1}{1-\alpha-\gamma}}
\]

\[
l = A_i r_j^{\frac{1-\alpha}{\alpha-1+\gamma}} \left( \phi e_j \right)^{\frac{1}{1-\alpha-\gamma}}
\]

where:

\[A_i = \left( \frac{\alpha}{\gamma w} \right)^{\frac{1-\alpha}{\alpha-1+\gamma}} \left( \frac{w}{\alpha} \right)^{\frac{1}{\alpha-1+\gamma}}\]

and

\[
n = \left( \frac{\alpha r_j}{\gamma w} \right) l
\]

\[
n = A_i \left( \frac{\alpha r_j}{\gamma w} \right)^{\frac{1-\alpha}{\alpha-1+\gamma}} \left( \phi e_j \right)^{\frac{1}{1-\alpha-\gamma}}
\]

\[
n = A_i \left( \frac{\alpha}{\gamma w} \right)^{\frac{1-\alpha}{\alpha-1+\gamma}} r_j^{\frac{1-\alpha}{\alpha-1+\gamma}} \left( \phi e_j \right)^{\frac{1}{1-\alpha-\gamma}}
\]
The profit function is given by:

\[ \pi_j = \varphi \epsilon_j \gamma n^a l^\gamma - w_n - r_j l - c_f \]

substitute:

\[
\pi_j = \left[ A_l \left( \frac{\alpha}{\gamma} \right) r_j^{\alpha - \frac{a}{1+a+\gamma}} (\varphi \epsilon_j)^{\frac{1}{1-a-\gamma}} \right]^a \left[ A_l \frac{\alpha^{1-a+\gamma}}{\gamma^{1-a+\gamma}} (\varphi \epsilon_j)^{\frac{1}{a+\gamma}} \right]^\gamma \varphi \epsilon_j - w_n - r_j l - c_f
\]

\[
\pi_j = \left[ A_l \left( \frac{\alpha}{\gamma} \right) \right]^a A_l^\gamma r_j^{\frac{a-\alpha}{a+\gamma} \frac{\gamma - \alpha}{\alpha + \gamma} (\varphi \epsilon_j)^{\frac{1}{1-a-\gamma}} \varphi \epsilon_j - w_n - r_j l - c_f
\]

let \( A_x = \left[ A_l \left( \frac{\alpha}{\gamma} \right) \right]^a \), then,

\[
\pi_j = A_x r_j^{\frac{\gamma}{a+\gamma} (\varphi \epsilon_j)^{\frac{1}{1-a-\gamma}}} - w \left( A_l \left( \frac{\alpha}{\gamma} \right) r_j^{\frac{1}{a+\gamma} (\varphi \epsilon_j)^{\frac{1}{1-a-\gamma}}} - r_j \left( A_l \frac{\alpha}{a+\gamma} (\varphi \epsilon_j)^{\frac{1}{1-a-\gamma}} \right) - c_f \right)
\]

\[
\pi_j = (\varphi \epsilon_j)^{\frac{1}{1-a-\gamma}} \left( A_x r_j^{\frac{\gamma}{a+\gamma}} - w A_l \left( \frac{\alpha}{\gamma} \right) r_j^{\frac{\gamma}{a+\gamma} - A_l r_j^{\frac{\gamma}{a+\gamma}}} \right) - c_f
\]

\[
\pi_j = (\varphi \epsilon_j)^{\frac{1}{1-a-\gamma}} \left( A_x r_j^{\frac{\gamma}{a+\gamma}} - w A_l \left( \frac{\alpha}{\gamma} \right) r_j^{\frac{\gamma}{a+\gamma} - A_l r_j^{\frac{\gamma}{a+\gamma}}} \right) - c_f
\]

\[
\pi_j = A_l \left( \frac{\alpha}{\gamma} \right) r_j^{\frac{\gamma}{a+\gamma} (\varphi \epsilon_j)^{\frac{1}{1-a-\gamma}}} - w A_l \left( \frac{\alpha}{\gamma} \right) r_j^{\frac{\gamma}{a+\gamma} - A_l r_j^{\frac{\gamma}{a+\gamma}}} - c_f
\]

\[
\pi_j = \left( A_x - w A_l \left( \frac{\alpha}{\gamma} \right) - A_l \right) (\varphi \epsilon_j)^{\frac{1}{1-a-\gamma}} r_j^{\frac{\gamma}{a+\gamma}} - c_f
\]

\[
\pi_j = A (\varphi \epsilon_j)^{\frac{1}{1-a-\gamma}} r_j^{\frac{\gamma}{a+\gamma}} - c_f
\]

where \( A = \left( A_x - w A_l \left( \frac{\alpha}{\gamma} \right) - A_l \right) \)
A.3 Computer Code

The following programs compute all of the variables, functions, and distributions of interest for the model. The programs should be run in order. Run ar1matrix.m, rentfinder.m, and ratiofinder.m, in that order. The parameters are established at the beginning of each program. There are two other functions, fastfirmchoice2.m and stationaryeq2.m, which are called inside the main program.

ar1matrix.m

% this simulates an ar1 process to produce a transition matrix;

clear
ro = .9;
sig = .2;
mu = 0;
n = 2000000;
phihigh = 6;
philow = .1;
nphi=201;

%phi1 = ones(1,n);
inter = (phihigh-philow)./(nphi-1);

phi1 = philow:inter:phihigh;        %grid of phi values
s = length(phi1);

for kk = 1:nphi;
clear phi2 trans1
eps1 = sig.*randn(1,n);         %calculate vector of random epsilons
phi2 = phi1(kk).^ro.*exp(eps1); %calculate next phi values
trans1 = hist(phi2,phi1);       %place in bins
trans2(kk,:) = trans1;
end
trans = trans2./n;                  %normalize, sum = 1
transparam = [sig,ro];
save phi1.txt phi1 -ASCII
save trans.txt trans -ASCII
save transparam.txt transparam -ASCII

figure(1)
plot(phi1,trans)
title('Conditional Distributions for Each Phi Value')
xlabel('Phi')
ylabel('Conditional Distributions')

rentfinder.m

%This code simulates a stochastic dynamic location choice model where there are two
%jurisdictions. There is a distribution of firms who choose at each time
%period to stay put, move, or shut down.

%This program plots the value functions and decision rules for a firm
%located in location 1 and location 2. It also outputs the entry costs for
%each location, which is equal to the expected value for a firm given entry
distributions.

clear
global alp gm thta1 thta2 ro sigeps intsig intmu w p
global beta cr ce cf tol n eps phi inter intdist
global v1 v2 gl g2 l1 l2 n1 n2 exval1 exval2 trans

load phi1.txt phi1 -ASCII %phi values loaded from simulation
load trans.txt trans -ASCII %probability transition matrix from simulation
load transparam.txt transparam -ASCII %ar1 process parameters from simulation

%PARAMETERS
r1 =2.5; %initial guess for rent in location 1
r2 = 1; %initial guess for rent in location 2
alp = .3; %exponent on labor (alpha)
gm = .3; %exponent on land (gamma)
thta1 = .66; %exponent on externality ej; for now, this is the same for all firms.
thta2 = .4; %determines phi's influence on externality
sig = transparam(1); %st. deviation of epsilon
ro = transparam(2); %correlation between consecutive shocks
intsig = .1; %initial st. dev. of log of phi of entrants
intmu = -.2; %initial mean of log of phi of entrants
w = .5; %wage rate, could be determined endogenously
p = 1; %price of output, could be determined endogenously (but probably not)
beta = .9; %time discount

%SETUP
phi = phi1; %grid of phi values; (Phi is the random shock in each period)
inter = (phi(5)-phi(4));
intdist = 1./sqrt(2.*pi).*(1./(phi.*intsig)).*exp(-1./2.*((log(phi)-intmu)./intsig).^2); %initial dist. of phi of entrants

%FIND RENTS
f = @fastfirmchoice2;

r0 = [r1, r2];

r = fminsearch(f,r0) %call function fastfirmchoice2 to calculate value function and return entry costs

%calculate decision rules --> shut down = 0, location1 = 1, location 2 = 2

for k = 1:n
    [val,g11] = max([0,exval1(k),exval2(k)-cr]);
g1(k) = g11-1;
[val,g22] = max([0,exval1(k)-cr,exval2(k)]);
g2(k) = g22-1;
end

dat = [phi',g1',g2',l1',l2'];
par = [ro,sig,intsig,intmu,gm,alp,w,thta1,thta2];

%Output important data to text files
save g1l1.txt dat -ASCII
save rents.txt r -ASCII
save parameters.txt par -ASCII

figure(1)
plot(phi,v1,phi,v2,phi,intdist);
%title('Value Functions')
xlabel('Phi')
ylabel('Value Functions')
figure(2)
plot(phi,g1,phi,g2)
legend('Decision Rule 1','Decision Rule 2');
%title('Value Functions')
xlabel('Phi')
ylabel('Desision Rules')
axis([0 2.5 -1 3])

fastfirmchoice2.m

function equalentry = fastfirmchoice2(r)

%This function takes in rent values for each location, calculates the
%expected values for entrants in each location and outputs the difference
%between those values some predetermined entry cost.

global alp gm thta1 thta2 ro sigeps intsig intmu w p
global beta cr ce cf tol n eps phi inter intdist
global v1 v2 g1 g2 l1 l2 n1 n2 exval1 exval2 trans

r1 = r(1);
r2 = r(2);

%Specifications for the externality and stochastic shock:
e1 = phi.*(alp./gm.*r1./w).^((thta1-thta2.*phi); %externality in location 1 from FOC's,
1+N/L, phi is in exponent,
e2 = phi.*(alp./gm.*r2./w).^((thta1-thta2.*phi); %externality in location 2 from FOC's,
1+N/L, phi is in exponent,
l1 = (((alp.*r1)./(gm.*w)).^(1-alp).*(w./(e1.*alp))).^(1./(alp-1+gm)); %static choice for
land
l2 = (((alp.*r2)./(gm.*w)).^(1-alp).*(w./(e2.*alp))).^(1./(alp-1+gm)); %static choice for
land
n1 = l1.*alp.*r1./w./gm;    %static choice for labor
n2 = l2.*alp.*r2./w./gm;    %static choice for labor
prod1 = n1.^alp.*l1.^gm.*e1;   %production function; location 1
prod2 = n2.*alp.*l2.^gm.*e2;   %production function; location 2
prof1 = prod1 - w.*n1 - r1.*l1 - cf;     %profit function; location 1
prof2 = prod2 - w.*n2 - r2.*l2 - cf;      %profit function; location 2
v1 = ones(1,n);
v2 = ones(1,n);
dif = tol + 1;
count = 0;

%VALUE FUNCTION ITERATION

while dif > tol
    vold1 = v1;
vold2 = v2;
    v11 = repmat(v1,n,1);              %replicate to multiply with transition matrix
    v22 = repmat(v2,n,1);
    intfunc1 = v11 .* trans;          %multiply value function with transition matrix
    intfunc2 = v22 .* trans;
    exval1 = sum(intfunc1');          %calculate expected value of location 1, for given phi
    exval2 = sum(intfunc2');          %calculate expected value of location 1, for given phi
    %calculate new value function
    for k = 1:n
        v1(k) = prof1(k) + beta .* max([0,exval1(k),exval2(k)-cr]);
        v2(k) = prof2(k) + beta .* max([0,exval2(k),exval1(k)-cr]);
    end
    dif1 = max(abs(v1-vold1));
    dif2 = max(abs(v2-vold2));
    dif = abs(max(dif1,dif2));
    count = count + 1;
end             %End of value function iteration loop

%ENTRY COSTS

%Integrate over initial distribution of entrants to entry costs
%multiply the value functions by the distribution function
intfunc11 = v1 .* intdist;
intfunc22 = v2 .* intdist;
%Calculate Expected Value of entrants
exval11 = sum(intfunc11) .* inter;
exval22 = sum(intfunc22) .* inter;
equalentry = abs(ce-exval11) + abs(exval22-ce);

ratiofinder.m

%This program takes in decision rules and rents and calculates masses of
%entrants.
clear

global A1 A2 del intdist xx11 xx12 xx21 xx22 n tol landuse m
global trans mu1 mu2 phi l1 l2 r1 r2 land1 land2 rentratio inter
A1 = .17;
A2 = .01;
del = .5;
mlow = 2;          %initial high and low guesses for m = M1/M2
mhigh = 5;

load trans.txt trans -ASCII %probability transition matrix from simulation
load -ASCII g1l1.txt
phi = g1l1(:,1); %shock values
g1 = g1l1(:,2); %decision rules location 1
g2 = g1l1(:,3); %decision rules location 2
l1 = g1l1(:,4); %land use choices location 1
l2 = g1l1(:,5); %land use choices location 2
n1 = g1l1(:,6); %employment choices location 1
n2 = g1l1(:,7); %employment choices location 2
load -ASCII rents.txt
r1 = rents(1); %rent in location 1
r2 = rents(2); %rent in location 2
load -ASCII parameters.txt
ro = parameters(1);
sigeps = parameters(2);
intsig = parameters(3);
intmu = parameters(4);
gm = parameters(5);
alp = parameters(6);
w = parameters(7);
thta1 = parameters(8);
thta2 = parameters(9);

n = length(g1);
tol = 10^-4;

%modified decision rules
x11 = zeros(1,n); %Firms in 1 who stay in 1
x21 = zeros(1,n); %Firms in 2 who go to 1
x12 = zeros(1,n); %Firms in 1 who go to 2
x22 = zeros(1,n); %Firms in 2 who stay in 2
x10 = zeros(1,n); %Firms in 1 who shutdown
x20 = zeros(1,n); %Firms in 2 who shutdown

%Calculate modified decision rules
for k = 1:n;
    if g1(k) == 1
        x11(k) = 1;
    end
    if g1(k) == 2
        x12(k) = 1;
    end
    if g2(k) == 1
        x21(k) = 1;
    end
    if g2(k) == 2
        x22(k) = 1;
    end
    if g1(k) == 0
        x10(k) = 1;
    end
    if g2(k) == 0
        x20(k) = 1;
    end
end

xx11 = repmat(x11,n,1); %replicate decision rules to allow for the use of matrix operations
xx12 = repmat(x12,n,1);
xx21 = repmat(x21,n,1);
xx22 = repmat(x22,n,1);
\[
\text{intdist} = 1./\sqrt{2.\pi} \cdot (1./\phi \cdot \text{intsig}) \cdot \exp(-1./2. \cdot ((\log(\phi) - \text{intmu})/\text{intsig})^2);
\]
% initial dist. of phi of entrants

meq = stationaryeq2(mhigh,mlow) % call function stationaryeq2 calculate distributions
and return ratio of entrants

figure(1)
plot(phi,mu1,phi,mu2)
legend('Stationary Dist. - Loc 1','Stationary Dist - Loc 2');
title('Adjusted Stationary Distributions')
xlabel('\Phi')
ylabel('s*\mu')

% CALCULATE ALL OF THE AGGREGATES

% entrants in location 2
M2 = (1./land1).*(r1./A1).^(1./del)
% entrants in location 1
M1 = meq.*M2
% "scale factors"
s1 = sum(mu1).*inter
s2 = sum(mu2).*inter
% total land use in each location
L1 = s1.*M2.*land1
L2 = s2.*M2.*land2
% total employment
N1 = L1.*((alp.*r1)./(gm.*w))
N2 = L2.*((alp.*r2)./(gm.*w))
% flows
flow12 = M2.*sum(x12.*mu1).*inter
flow21 = M2.*sum(x21.*mu2).*inter
flow10 = M2.*sum(x10.*mu1).*inter
flow20 = M2.*sum(x20.*mu2).*inter
% total number of firms
S1 = s1.*M2;
S2 = s2.*M2;
% average employment per firm in each location
navg1 = N1./S1;
navg2 = N2./S2;
% all aggregates as a fraction of S1 + S2
rM1 = M1./(S2+S1);
rM2 = M2./(S2+S1);
rl1 = L1./(S2+S1);
rl2 = L2./(S2+S1);nN1 = N1./(S2+S1);
rN2 = N2./(S2+S1);
rf12 = flow12./(S2+S1);
rf21 = flow21./(S2+S1);
rf10 = flow10./(S2+S1);
rf20 = flow20./(S2+S1);

\textbf{stationaryeq.m}

function meq = stationaryeq2(mhigh,mlow)

% This program takes in ratio of entrants, decision rules and rents and calculates
stationary distributions in each location

global A1 A2 del intdist xx11 xx12 xx21 xx22 n tol rentratio m
global trans mu1 mu2 phi l1 l2 r1 r2 land1 land2 landuse inter
gs = (3-sqrt(5))/2; % golden ratio
momid = gs*(mhigh-mlow)+mlow;
m = [mlow,mhigh,mmid];
check1 = ones(1,3);
check = 2;

mu1 = zeros(1,n);  %initial guesses for "adjusted distributions"
mu2 = zeros(1,n);

while check > tol  %Main loop uses golden section method to find ratio m
    m = [mlow,mhigh,mmid];
    for k = 1:3
        dif = 1;
count = 0;
        %Main loop finds the ratio that satisfies the land supply functions
        while dif > tol
            muold1 = mu1;
muold2 = mu2;
mu11 = repmat(mu1,n,1);
mu22 = repmat(mu2,n,1);

            %Recursive Equations
            mu1 = sum((trans'.*xx11.*mu11)')+sum((trans'.*xx21.*mu22)')+m(k).*intdist';
mu2 = sum((trans'.*xx12.*mu11)')+sum((trans'.*xx22.*mu22)')+intdist';

dif1 = max(abs(mu1-muold1));
dif2 = max(abs(mu2-muold2));
dif = abs(max(dif1,dif2));
count = count+1;
        end
        inter = (phi(5)-phi(4));

        land1 = sum(l1'.*mu1).*inter;  %integrate to find "scaled" land use in each location
        land2 = sum(l2'.*mu2).*inter;

        landuse = A1./A2.*(land1./land2).^del;
        rentratio = r1./r2;

        check1(k) = (landuse - rentratio);
    end

    if check1(1).*check1(3) <= 0  %These 'if' statements update the golden section points
        check = abs(check1(3))
mhigh = mmid;
        mmid = gs*(mmid-mlow)+mlow;
    elseif check1(3).*check1(2) <= 0
        check = abs(check1(3))
        mlow = mmid;
        mmid = gs*(mhigh-mmid)+mmid;
    else
        error = 1111
        check = 0;
    end
end

meq = m(3);