WORKING PAPERS

RESEARCH DEPARTMENT

Satyajit Chatterjee
Federal Reserve Bank of Philadelphia

November 2003

FEDERAL RESERVE BANK OF PHILADELPHIA

Working Paper No. 03-24

On the Contribution of Agglomeration Economies to the Spatial Concentration of U.S. Employment

Satyajit Chatterjee[1]

Research Department

Federal Reserve Bank of Philadelphia

Philadelphia, PA 19106

November 10, 2003 (First Draft: April 7, 2003).

ABSTRACT

Why does the level of economic activity vary so much across space? One reason given is "agglomeration economies," meaning that a firm's or household's production costs (of market and home goods, respectively) are lower when production is carried out in close proximity to other firms and households. In this paper I explore, via a quantitative spatial macroeconomic model, the contribution of agglomeration economies to the observed spatial concentration of US employment. The approach is analogous to "business-cycle accounting" or "growth accounting." The results of the "spatial accounting" performed in this study depend on the details of the model used. The critical detail pertains to how the model rationalizes the stability of low-density localities. If it is rationalized via an appeal to restrictions on labor mobility, the accounting implies that the bulk of spatial concentration results from an unequal distribution of natural advantages. In contrast, if it is rationalized via an agglomeration threshold (an employment level below which local increasing returns do not operate), the accounting implies that the bulk of the spatial concentration results from increasing returns.

The bulk of the economic activity of an industrially developed country takes place in densely settled areas that make up a small portion of a country's overall territory. This striking spatial concentration is thought to result from two distinct sources. The first source is the need to extract and/or use a natural resource. The second source is the cost advantages conferred by spatial concentration itself, i.e., agglomeration economies or increasing returns.

If both fundamentals and increasing returns are relevant for the location of economic activity, how much of the observed spatial concentration of a country can be attributed to each of these two sources? In this paper, I answer this question using an approach that is novel in this context but common in quantitative macroeconomics. I start by developing a general equilibrium model with heterogeneous localities and local increasing returns. The model describes the determination of employment in each of many locations, locations that taken together exhaust the physical space available for economic activity. I restrict the model's parameters to match evidence on the likely magnitude of net agglomeration benefits (local increasing returns net of congestion costs) and the observed spatial distribution of U.S. employment for a recent year (1999). I then use the restricted model to determine how employment for that year would be distributed if agglomeration economies are counterfactually suppressed. A comparison of the counterfactual employment distribution with the actual one provides one indication of the influence agglomeration economies on the spatial distribution of employment.

The research reported in this paper complements the few existing studies that attempt to gauge the relative importance of natural advantages versus increasing returns. Kim (1995, 1999) uses a model of comparative advantage with immobile factors to calculate an upper bound on the contribution of agglomeration economies to regional variation in economic activity. Ellison and Glaeser (1997a) examine how their measures of industry concentration (introduced in Ellison and Glaeser (1999)) decline when account is taken of the availability of a limited set of natural resources. Rappaport and Sachs (2002) seek to identify the contribution of access to water (to a coast, navigable river or lake) for the spatial variation in employment. These studies conclude that natural advantages seem to play a major role in accounting for spatial variation in economic activity.[1]

---

[1] Relatedly, Hanson (2002) has looked for confirmation, in US data, of the "home market effect" predicted by Krugman's (1991) model of geographic concentration. Davis and Weinstein (2002a, b) have looked for evidence of

A common feature of these studies is that each attempts to determine the relative importance of natural advantages by conditioning spatial variation in the variable of interest (economic activity, employment, or measures of concentration) on proxies for specific natural resources. At least in the Kim and Ellison and Glaeser papers, the residual variation is viewed as an upper bound estimate of the impact of increasing returns. In contrast, the approach in this paper is to start with estimates of agglomeration economies (surveyed in Moomaw (1981) and more recently in Rosenthal and Strange (2003)) and determine, with the aid of a general equilibrium model, how much of the spatial variation in employment can plausibly be explained on the basis of these estimates. This approach has the advantage in that it's easier to establish a plausible range of variation for the degree of local increasing returns than it is to measure the (essentially limitless) variation in natural resources.

The investigation is disciplined by requiring that the model on which the counterfactual is performed account for the *actual* spatial distribution of employment. Both agglomeration economies and natural advantage are *centripetal* forces (to use terms popularized by Fujita, Krugman, and Venables (1999)) that work to concentrate economic activity into a relatively small number of locations. These centripetal forces are opposed by a set of *centrifugal* forces that work to disperse employment. The main centrifugal force in the model is the cost imposed by congestion. Congestion costs work to push employment away from dense areas toward less dense ones. For any given magnitudes of agglomeration economies and congestion costs, the requirement that the model account for the actual distribution of employment imputes a value to the strength of each locality's natural advantage. Thus, the empirical strategy is to treat natural advantage as a residual. When agglomeration economies are counterfactually suppressed, the imputed spatial pattern of natural advantage become the counterweight to the centrifugal force of congestion costs. In this paper, natural advantage should be interpreted to include the effects of local regulations on the production and consumption of goods.

The key challenge in implementing this approach stems from the fact that for plausible magnitudes of agglomeration economies and congestion costs, the model predicts that low-density localities are *unstable* and should either agglomerate up and become more dense or lose all their residents. Put differently, the model cannot account for the large numbers of low-density localities

---

multiple equilibria (predicted, again, by the Krugman model but implied more generally by agglomeration economies) in postwar Japan.

observed in reality as a stable equilibrium outcome. Instability of small localities is a well-known theme in urban economics, and this theme is present in this study as well.

I experiment with two ways of restoring the stability, or viability, of low-density localities. In the first alternative, a second centrifugal force is introduced in the guise of individuals who can only reside in particular places. This feature is reminiscent of the distribution of farmers in Krugman's (1991) model of economic geography, people who are modeled as being tied to a particular region and who do not make a location decision. It is also the feature invoked by Ciccone and Hall (1996) to make sense of their finding that the net benefits from agglomeration are positive. With positive net benefits to agglomeration, all economic activity should get concentrated in one giant location. Ciccone and Hall appeal to people's devotion to particular places as the force that prevents this extreme concentration from materializing.

The second alternative is to weaken the force of increasing returns for low-density localities. In this alternative, local employment is required to attain a certain minimum level for agglomeration economies to manifest themselves. Instability of low-density equilibrium results from the positive feedback induced by local increasing returns. By suppressing the positive feedback, an agglomeration threshold restores stability of low-density localities. Consequently, with an agglomeration threshold in place, less restriction on labor mobility is needed to account for low-density localities. Indeed, if the agglomeration threshold is chosen sufficiently high, no restriction on labor mobility is needed at all.

If the first alternative is chosen, the model reproduces the actual distribution of employment by imputing to each locality the strength of its natural advantage and whether existing residents of the locality are free to relocate. When agglomeration economies are counterfactually suppressed, these imputations of natural advantage and restrictions on labor mobility determine the new equilibrium distribution of employment. The new distribution of employment is unambiguously less concentrated than the actual distribution, a result that holds for every counterfactual performed in this study. However, in this case the quantitative magnitude of the deconcentration turns out to be somewhat modest. When agglomeration economies are suppressed, the model predicts that the spatial concentration of employment, as measured by the Gini concentration index, will be 24 percent less than the actual spatial concentration of employment. Furthermore, for plausible variation in the magnitude of agglomeration economies and congestion costs, the predicted decline in spatial concentration does not exceed 35 percent. Thus, a basic message from this set of counterfactuals

is that agglomeration economies are an important factor in spatial concentration but not the *most* important factor. The bulk of the spatial concentration appears to be due to spatial inequality in fundamentals.

If the second alternative is chosen, the model again reproduces the actual distribution of employment by imputing to each location the strength of its natural advantage and whether residents of the location are free to relocate. However, because the agglomeration threshold is an alternative way of restoring stability to low-density localities, the imputed restrictions on labor mobility are now generally much less severe. It turns out that this alternative implies a much larger contribution agglomeration economies. For a modest value of the threshold parameter, the baseline calibration predicts that if agglomeration economies are suppressed, spatial concentration will decline 63 percent. Furthermore, for plausible variation in the magnitude of agglomeration economies and congestion costs, the predicted decline in spatial concentration stays above 50 percent. For this set of experiments, then, agglomeration economies turn out to be a major factor in the spatial concentration of employment.

Since neither empirical nor theoretical literatures on agglomeration economies address the possibility of an agglomeration threshold, it's unclear which of the two sets of quantitative results is closer to the truth. However, the fact that the results depend on the presence or absence of an agglomeration threshold is in itself an important finding. Basically, it appears that to gauge the contribution of local increasing returns to spatial concentration, it's important to understand why low-density localities are a stable outcome of spatial equilibrium. If low-density localities survive because some people have a strong desire to live in these places, the contribution of agglomeration economies to spatial concentration may well be quite low. On the other hand, if low-density localities survive because the positive feedback from increasing returns is absent (or very weak) for these localities, the contribution of increasing returns to spatial concentration could well be much larger.

Since an agglomeration threshold *limits* the scope of increasing returns, it isn't obvious why it should imply a *larger* contribution of increasing returns. Briefly, the reason for this result is as follows. If there is no agglomeration threshold, currently large (and dense) localities benefit more from local increasing returns compared to currently small (and sparse) localities. Because the model is required to account for the current distribution of employment, it counterbalances the benefit of size by imputing relatively poor fundamentals (natural advantage) to currently dense localities. When agglomeration economies are suppressed, there is a tendency for employment to

reconcentrate around currently *sparse* localities because these localities have low congestion costs *and* relatively superior fundamentals. This tendency for concentration to re-emerge prevents the Gini index from declining steeply. In contrast, size confers less of an advantage when there is an agglomeration threshold and the model imputes relatively superior fundamentals to currently dense localities. In this case, suppression of agglomeration economies does not imply a tendency for employment to re-concentrate in currently sparse localities. Instead, absence of agglomeration economies leads to a more dispersed spatial distribution of employment and, consequently, a steeper decline in the Gini concentration index.

In accounting for the current spatial distribution of employment, the model also determines if the residents of each locality are free to relocate. These imputed restrictions on labor mobility also influence the results. When there is no agglomeration threshold, these restrictions become a barrier against the tendency of employment to *re*-concentrate; when there is an agglomeration threshold, the restrictions become a barrier against the tendency of employment to *de*-concentrate. In other words, the imputed restrictions on labor mobility prevent the spatial inequality in fundamentals to fully exert their influence. When the imputed spatial pattern of fundamentals is conducive to spatial concentration (as it is when there is no agglomeration threshold), the inability of people to move is a force *against* spatial concentration; when the imputed pattern is conducive to a dispersed distribution of employment (as it is when there is an agglomeration threshold), the inability of people to move is a force in *favor* of spatial concentration.

The paper is organized as follows. The next section describes the model. Section 3 has a detailed discussion of the nature of spatial equilibrium in the model. Section 4 describes how the model is mapped to U.S. data, and section 5 presents the results of all the counterfactuals. Section 6 summarizes the findings and suggests some questions for future research.

II.   THE MODEL ECONOMY

There are M distinct geographical areas indexed by $i = 1, 2, 3, ..., M$. These areas will be referred to as localities. The collection of localities is assumed to exhaust the physical space available for economic activity in the economy. Localities can differ with respect to area, with respect to the availability of natural resources, and with respect to laws and regulations that affect production and consumption possibilities. There are large number of individuals, $N > 1$, who live and work in

these localities.[2]

A. Firms

There is one costlessly transportable composite good. The firm-level production function for producing the transportable good in locality i is

$$y = \lambda \phi_i \beta(N_i) k^{1-\alpha} l^{\alpha}, \ 0 < \alpha < 1, \tag{1}$$

where k and l are the capital and labor used by a plant, λ is an economy-wide technology index, and ϕ_i is an index that captures the combined impact of locality-specific factors on production capabilities. For instance, the production advantages conferred by being on the coast or on a navigable river and the impact of local labor and environmental regulations will all be captured in ϕ_i. $\beta(N_i)$ is a function of total employment in locality i, denoted N_i, that takes into account the external economies in production resulting from the scale of the locality's labor pool. This is one way in which *agglomeration economies* enter the model. The function is taken to be

$$\beta(N_i) = \max\left\{\overline{N}^{\nu}, N_i^{\nu}\right\}, \ \overline{N} \geq 1, \nu > 0. \tag{2}$$

The specification assumes that there is a threshold level $\overline{N}$, potentially 1, above which agglomeration economies operate. In the range where agglomeration economies do operate the elasticity of agglomeration benefits with respect to change in local employment is a positive constant ν.[3]

Each locality can also produce a good that cannot be shipped to a different locality. The firm-level production function for the non-transportable good produced in locality i is

$$g = \xi_i \Gamma(N_i, A_i) y, \tag{3}$$

where y is the transportable good used as input by the firm, ξ_i is an index that captures the combined impact of locality-specific factors on production of the non-transportable good (analogous to ϕ_i) and $\Gamma(N_i, A_i)$ is a function that takes into account the *diseconomies* imposed by local congestion on the production of the non-transportable good. This function is taken to be

$$\Gamma(N_i, A_i) = e^{-\gamma(N_i/A_i)}, \ \gamma > 0, \tag{4}$$

---

[2]The model is adapted from Chatterjee and Carlino (2001).

[3]An alternative specification of $\beta(N_i)$ is $\max\{1, (N_i - \overline{N})^{\nu}\}$. In this case, the elasticity of agglomeration benefit with respect to local employment, for $N_i > \overline{N}$, is $\nu \cdot [(N_i - \overline{N})/N_i]$. Since estimates of ν do not take the possibility of a threshold effect into account, the specification in the text was chosen over this one.

where A_i is land area of locality i. Thus, according to (3) and (4), higher employment density in a given locality makes the production of the non-transportable good less efficient. An important property of Γ is that the absolute value of its elasticity with respect to employment density is increasing in employment density:

$$\frac{d(\ln \Gamma)}{d\ln(N_i/A_i)} = -\gamma \cdot (N_i/A_i). \tag{5}$$

All firms behave competitively. Producers of locality i's non-transportable good take the price of the transportable good and employment density in locality i as given. With the price of the transportable good as a numeraire, competitive production implies that price of the non-transportable good in location i, denoted p_i, cannot exceed its marginal cost:

$$p_i \leq \xi_i^{-1} e^{\gamma(N_i/A_i)}. \tag{6}$$

Producers of the transportable good in locality i take the level of local employment and the product wage in that locality, w_i, as given. They also takes r as given. Again, competitive production implies that the price of the transportable good cannot exceed its marginal cost of production in locality i:

$$1 \leq [\alpha^\alpha (1-\alpha)^{(1-\alpha)} \lambda \phi_i \beta(N_i)]^{-1} w_i^\alpha r^{(1-\alpha)}. \tag{7}$$

B. People

There are two types of individuals: mobile individuals who can move between localities and immobile individuals who cannot. I assume that there is a large number $\underline{N}_i > 1$ of immobile individuals in each locality i, although $\underline{N}_i/N_i$ need not be large. Both types have one unit of labor which they supply inelastically to firms producing the transportable good in their locality.

The utility of an individual living and working in locality i is given by:

$$U = \psi_i \pi(N_i) c^{1-\theta} g^\theta, \ 0 < \theta < 1, \tag{8}$$

where c and g are the individual's consumption of the transportable good and non-transportable good, respectively, ψ_i is an index that captures the combined impact of locality-specific amenity factors (such as climate) and laws and regulations on utility. The function $\pi(N_i)$ takes into account external benefits that workers derive from the scale of the local labor pool. This is the second way in which agglomeration economies enters the model. This function is taken to be

7

$$\pi(N_i) = \max\left\{\overline{N}^{\eta}, N_i^{\eta}\right\}, \ \overline{N} \geq 0, \eta > 0 \tag{9}$$

which parallels the specification of agglomeration benefits in the production of the transportable good and has similar properties.

For tractability, I assume that mobile individuals do not have any capital income. Conditional on the choice of locality, utility maximization implies that a mobile individual in locality i chooses $g = \theta(w_i/p_i)$ and $c = (1-\theta)w_i$. Thus, the indirect utility of a mobile individual residing in locality i is

$$V_i = [\psi_i \pi(N_i)(1-\theta)^{(1-\theta)}\theta^{\theta}]p_i^{-\theta}w_i. \tag{10}$$

Given costless mobility, a mobile individual will choose to live and work in location i only if

$$V_i = \max_j\{V_j\}. \tag{11}$$

The demand functions of immobile individuals who reside in locality i is similar to that of mobile individuals except that they derive income from capital as well. That is, an immobile individual with asset level x who resides in locality i will choose $g = \theta(w_i + rx)/p_i$ and $c = (1-\theta)(w_i + rx)$. For such an individual the indirect utility is

$$V_i(x) = [\psi_i \pi(N_i)(1-\theta)^{(1-\theta)}\theta^{\theta}]p_i^{-\theta}(w_i + rx).$$

Since they are immobile, their indirect utility is *not* required to satisfy a condition like (11).

C. Equilibrium

Since there is a large number of immobile individuals in every locality, there are individuals supplying labor inelastically to firms producing the transportable good in every locality. Therefore, in any equilibrium, there must be positive production of the transportable good in every locality. Then, it follows from (7) that

$$w_i = [\alpha^{\alpha}(1-\alpha)^{(1-\alpha)}\lambda\phi_i\beta(N_i)]^{\frac{1}{\alpha}}r^{\frac{-(1-\alpha)}{\alpha}}. \tag{12}$$

Additionally, since every locality has large number of immobile individuals with strictly positive income (note that $w_i > 0$), it follows that there will be a positive demand for the non-transportable

good in every locality for any $p_i \in [0, \infty)$. Therefore, in any equilibrium, there will be positive production of the non-transportable good in every locality. Then, it follows from (6) that

$$p_i = \xi_i^{-1} e^{\gamma(N_i/A_i)}. \tag{13}$$

Denote $\theta^\theta (1-\theta)^{(1-\theta)} \alpha^\alpha (1-\alpha)^{(1-\alpha)} \lambda^{1/\alpha} r^{\frac{-(1-\alpha)}{\alpha}}$ by $H(\alpha, \theta, \lambda, r)$, the product of locality-specific factors $\psi_i \cdot \xi_i^{-\theta} \cdot \phi_i^{1/\alpha}$ by S_i, ν/α by μ, and $\theta\gamma$ by δ. Then, substituting (12) and (13) into (10) and using (2) and (9), yields:

$$V_i = H(\alpha, \theta, \lambda, r) \cdot S_i \cdot \max\{\overline{N}^{\mu+\eta}, N_i^{\mu+\eta}\} \cdot e^{-\delta N_i/A_i}. \tag{14}$$

The r.h.s of equation (14) incorporates all the economic forces at work in this model. The first factor, $H(\alpha, \theta, \lambda, r)$ captures the effect on utility of economy-wide factors such as the level of technology λ and the level of interest rates r. The second factor, S_i, incorporates the effect of all local factors such as weather and amenities, availability of natural resources, and local regulations affecting the production of the transportable and non-transportable goods.[4] The third factor, $\max\{\overline{N}^{\mu+\eta}, N_i^{\mu+\eta}\}$, incorporates the positive effect of agglomeration economies. These economies arise both from the benefits that firms receive from producing in a locality with a bigger labor pool and the benefits that people enjoy from locating close to other people. The former channel is captured in the parameter μ (which depends, in turn, on α and ν) and the latter channel in the parameter η. When $\overline{N}$ is 1, these effects operate as long as there is more than 1 person employed in the location, but when $\overline{N}$ is greater than 1 it operates only when local employment exceeds the threshold $\overline{N}$. The final factor, $e^{-\delta N_i/A_i}$, incorporates the negative effect of congestion. As a locality gets more dense, the price of the locally produced non-transportable good rises. For a given rise in density, the reduction in utility is greater if the share of non-transportable good in the worker's budget, θ, is higher, if the adverse effect on the production of the non-transportable good, δ, is higher, and if initial density, N_i/A_i, is higher.

An equilibrium for this economy is a number V^* and a vector $(N_i^*)_{i=1}^M$ that satisfy the following conditions:

$$N_i^* \geq \underline{N}_i \ \forall \ i = 1, 2, 3 \ldots, M \tag{15}$$

---

[4]For instance, a mild climate or access to scenic spots would result in a higher S_i through a higher ψ_i, the availability of a valuable natural resource such as petroleum would result in a higher S_i through a higher ϕ_i, and tough labor and environmental regulations would result in a lower S_i through a lower ϕ_i and ξ_i but a possibly higher S_i through a higher ψ_i.

$$V^* = H(\alpha, \theta, \lambda, r) \cdot S_i \cdot \max\{\overline{N}^{\mu+\eta}, N_i^{*\mu+\eta}\} \cdot e^{-\delta N_i^*/A_i} \text{ if } N_i^* > \underline{N}_i \tag{16}$$

$$V^* \geq H(\alpha, \theta, \lambda, r) \cdot S_i \cdot \max\{\overline{N}^{\mu+\eta}, N_i^{*\mu+\eta}\} \cdot e^{-\delta N_i^*/A_i} \text{ if } N_i^* = \underline{N}_i \tag{17}$$

$$\sum_{i=1}^{M} N_i^* = N. \tag{18}$$

To see this, suppose that we have V^* and $(N_i^*)_{i=1}^{M}$ that satisfy (15) - (18). Choose $w_i^* = [\alpha^\alpha(1-\alpha)^{(1-\alpha)}\lambda\phi_i \max\{\overline{N}^\nu, N_i^{*\nu}\}]^{\frac{1}{\alpha}} r^{\frac{-(1-\alpha)}{\alpha}}$ and $p_i^* = \xi_i^{-1} e^{\gamma(N_i^*/A_i)}$. At these wages, production of the transportable good yields zero profits in every locality. So, production of the transportable good in any locality i can expand to the point where all N_i^* people are employed, i.e., at these wages the labor market in each locality can clear. Similarly, at the prices p_i^* the production of the non-transportable good yields zero profits in every locality. So, production of the non-transportable good in any locality i can expand to the point where total demand for the non-transportable good from mobile and immobile individuals is met, i.e., at these prices the non-transportable goods market in each locality can clear as well.[5] The only other markets in this model are those for the transportable good and capital and by assumption both are international markets with given prices. Finally, it is obvious that substituting w_i^* and p_i^* into (10) will yield $V_i = V^*$ for any i with $N_i^* > \underline{N}_i$ and $V_i \leq V^*$ for all other i. Therefore, mobile individuals do not have an incentive to move to a different location from their current one.

III. Agglomeration Economies, Instability, and Multiple Equilibria

While conditions (15) - (18) completely characterize an equilibrium, not all pairs $\left(V^*, (N_i^*)_{i=1}^{M}\right)$ that satisfy these conditions are economically meaningful. Because of increasing returns, an equilibrium may be unstable with respect to small perturbations. The aim of this section is to (i) explain how instability manifests itself in this model and how it can lead to multiple stable equilibria for a locality and (ii) to refine the definition of an equilibrium to exclude unstable equilibria.

It is convenient to work with logarithmic transforms. Let $\ln(N_i/A_i)$ be denoted by d_i, $\ln(\overline{N}/A_i)$ by $\overline{d}_i$, $\ln(H(\alpha, \theta, \lambda, r) \cdot S_i)$ by s_i, and $\ln(A_i)$ by a_i. Then, equation (14) can be written as

$$\ln(V_i) = s_i + (\mu + \eta)a_i + (\mu + \eta) \cdot \max\{d_i, \overline{d}_i\} - \delta \cdot e^{d_i}.$$

In what follows, I will treat N_i as a continuous variable. Then, the r.h.s. of the above equation can be viewed as a function of d, that is, define $v_i(d) : [\underline{d}_i, +\infty) \to R$, where $\underline{d}_i \equiv \ln(\underline{N}_i/A_i)$ is the

---

[5] The total demand for the non-traded good in location i is $\theta(w_i^*/p_i^*)(N_i^* - N_i^I) + \theta(w_i^* + r\overline{x}_i)/p_i^*)N_i^I$, where $\overline{x}_i$ is the average asset holdings of immobile individuals residing in locality i.

smallest employment density possible in locality i and

$$v_i(d) = s_i + (\mu + \eta)a_i + (\mu + \eta) \cdot \max\{d, \overline{d}_i\} - \delta \cdot e^d. \tag{19}$$

First consider the case where $\underline{d}_i < \overline{d}_i < \ln[(\mu + \eta)/\delta]$. Observe that $v_i(d)$ is continuous over its entire domain and differentiable everywhere except at $\overline{d}_i$ For $d \in [\underline{d}_i, \overline{d}_i]$, the implied employment level is too low for agglomeration economies to be present. Over this range, only the congestion effect operates so that the function is strictly decreasing and strictly concave:

$$\partial v_i/\partial d = \partial^2 v_i/\partial d^2 = -\delta \cdot e^d < 0. \tag{20}$$

For d greater than the $\overline{d}_i$ the implied employment level exceeds $\overline{N}$ and agglomeration economies are present. Over this range, the behavior of the function reflects the relative strengths of local increasing returns and congestion diseconomies. For $d \in (\overline{d}_i, +\infty)$, the first two derivatives with respect to d are

$$\partial v_i/\partial d = (\mu + \eta) - \delta \cdot e^d, \text{ and} \tag{21}$$

$$\partial^2 v_i/\partial d^2 = -\delta \cdot e^d < 0. \tag{22}$$

Thus, $v_i(d)$ continues to be strictly concave but the first derivative is positive between $(\overline{d}_i, \ln[(\mu + \eta)/\delta])$, zero at $\ln[(\mu + \eta)/\delta]$, and negative between $(\ln[(\mu + \eta)/\delta], \infty)$.

Figure 1 illustrates this case and shows one set of equilibrium possibilities for locality i. Let v^* be the equilibrium utility available to mobile individuals. There are three points where the $v_i(d)$ function intersects the horizontal "utility-at-other-locations" line, and so there are three interior equilibrium density levels.[6] However, the middle equilibrium, d^M, has the feature that *any* small exogenous increase or decrease in employment density around it raises the local utility level above or below what mobile individuals can get in other locations. If this property is coupled with the intuitive notion that a locality which offers a higher (lower) utility than v^* can expect to draw (lose) mobile individuals, then a transient perturbation in density around d_M will move the locality to either the d^L or d^H equilibrium. In this sense, the equilibrium d_M is *unstable*. In contrast, the d^L and d^H equilibrium have the feature that for small enough changes in density, local utility moves in a direction opposite to the change in density. Hence, these equilibria are stable. The instability of the (local) equilibrium results from the positive feedback introduced by agglomeration economies.

---

[6]This assumes that the particular equilibrium attained in location i has a vanishingly small effect on v^*.

Next, consider the case where $\underline{d}_i < \ln[(\mu + \eta)/\delta] \leq \overline{d}_i$. Now, (21) implies that the $v_i(d)$ function is strictly decreasing over its entire domain and diverges to $-\infty$ as d approaches $+\infty$. If $v_i(\underline{d}_i) > v^*$, there is a unique employment density strictly greater $\underline{d}_i$ for which $v_i(d)$ is exactly equal to v^*. If $v_i(\underline{d}_i) \leq v^*$, there can be no employment density strictly greater than $\underline{d}_i$ for which $v_i(d)$ will equal v^*. In this case, the only equilibrium is the corner equilibrium $\underline{d}_i$. Therefore when $\overline{d}_i \geq \ln[(\mu+\eta)/\delta]$, there is a unique local equilibrium for any value of v^* and $\underline{d}_i$ and this equilibrium is stable.

Next, consider the case where $\ln[(\mu + \eta)/\delta] \leq \underline{d}_i$. Then, regardless of the value of $\overline{d}_i$ it follows, from (21) again, that $v_i(d)$ function is strictly decreasing. Hence, as in the previous case, there will be a unique stable equilibrium.

Finally, consider the case where $\overline{d}_i \leq \underline{d}_i < \ln[(\mu + \eta)/\delta]$. This can happen if $\underline{d}_i$ is raised (i.e., $\underline{N}_i$ is raised) or $\overline{d}_i$ is lowered (i.e., $\overline{N}_i$ is is lowered). It follows from (21) that $v_i(d)$ is increasing right from the start, reaches a maximum at $\ln(\mu + \eta)/\delta$, and then declines. Figure 2 illustrates this case. A sufficient condition for there to be stable multiple equilibria is that $v_i(\underline{d}_i) < v^* < v_i(\ln[(\mu+\eta)/\delta])$. When multiple stable equilibria exist, the lower stable equilibria d^L is always the corner equilibrium $\underline{d}_i$. Even though $v_i(d)$ is increasing at $\underline{d}_i$, the lower equilibrium is stable because $v_i(\underline{d}_i) < v^*$; this gap implies that there is a small enough increase in d, say ε, for which $v_i(\underline{d}_i + \varepsilon)$ continues to remain below v^*.

In the rest of this paper, I will focus only on stable equilibria. To be clear, a stable equilibria is defined as follows.

Definition : The collection $\{d_i^*, v^*, s_i, \underline{d}_i, \overline{d}_i, N\}$ is a stable equilibrium if it satisfies the following conditions:

$$d_i^* \geq \underline{d}_i \tag{23}$$

$$d_i^* > \underline{d}_i \Rightarrow v^* = v_i(d_i^*) \tag{24}$$

$$d_i^* = \underline{d}_i \Rightarrow v^* \geq v_i(d_i^*) \tag{25}$$

$$\sum_{i=1}^{M} A_i \cdot e^{d_i^*} = N \tag{26}$$

If $\overline{d}_i < \ln[(\mu + \eta)/\delta]$, then

$$d_i^* > \underline{d}_i \Rightarrow d_i^* \notin [\overline{d}_i, \ln[(\mu + \eta)/\delta]] \text{ and} \tag{27}$$

$$d_i^* \in [\overline{d}_i, \ln[(\mu + \eta)/\delta]] \Rightarrow v^* > v_i(d_i^*). \tag{28}$$

Conditions (23)-(26) correspond to conditions (15)-(18). The stability requirements are incorporated in conditions (27) and (28). These conditions apply only when $\overline{d}_i < \ln[(\mu + \eta)/\delta]$, i.e., the

corresponding $v_i(d)$ function has a rising segment. If a location has mobile workers in equilibrium, (27) prohibits d_i^* from lying in the closed interval $[\bar{d}_i, \ln[(\mu + \eta)/\delta]$; as shown in Figure 1 and 2, this interval corresponds to the domain of d for which the $v_i(d)$ function is upward sloping. If a location's equilibrium density lies in $[\bar{d}_i, \ln[(\mu + \eta)/\delta]$, (28) requires that the utility to mobile individuals be strictly less than v^* (if a location has employment density is in $[\bar{d}_i, \ln[(\mu+\eta)/\delta]$ then by (27) it cannot have any mobile workers and by (25) it cannot deliver utility greater than v^*; the stability condition (28) rules out utility *equal* to v^*).

IV.   Mapping the Model Economy to U.S. Data

The goal of this section is to restrict the parameters of the model so that the behavior of the model economy matches the behavior of the US economy in as many dimensions as there are parameters. An important preliminary step in doing so is to decide what geographical areas in the U.S. correspond to localities in the model economy. There are two key assumptions about localities made in the model. First, the set of locations *exhaust* the physical space available for economic activity and, second, people live and work in the *same* locality. The second assumption suggests choosing the geographical areas so that there is little, or no, inter-area commuting. The first assumption suggests choosing a collection that is comprehensive enough to include most of U.S. territory. With these requirements in mind, the geographic areas were chosen to be the 17 consolidated metropolitan statistical areas (CMSAs), 258 metropolitan statistical areas (MSAs), and 2248 rural counties. Together, this gives a total of 2523 approximately self-contained labor market areas, covering all of the 48 contiguous states.[7]

Setting aside the agglomeration threshold $\overline{N}$, the numerical specification of the model described by (23) - (27) involves setting values for the five economy-wide parameters, namely, $\nu, \eta, \alpha, \theta$, and γ, and 3×2523 locality-specific parameters, namely, locality areas a_i, locality-specific factors s_i and locality-specific density of immobile individuals $\underline{d}_i$. Some of these parameters can be set by appealing to direct observations and to the findings of existing microstudies. These include the land area of each locality a_i, which is obtained by direct observation, and the parameters ν, α, and θ. Microeconomic studies (Sveikauskas (1976), Moomaw (1981), Rosenthal and Strange (2003)) indicate that ν varies between 0.02 and 0.09 for different industries, with $v = 0.03$ being the median value. The calibration will take 0.03 as the starting point but will examine sensitivity of

---

[7]All MSA and CMSA definitions pertain to 1990.

the results to higher and lower values. Expenditure shares of businesses and households suggest that $\alpha = 0.70$ (Gollin(2002)) and $\theta = 0.50$ (Jacobs and Shipp (1990)) are reasonable choices[8]. There are no estimates of η, so I assigned it a value of 0.01 on the ground that some positive effect of size on utility seems reasonable.

The congestion cost parameter γ controls the response of the price of the non-transportable good to variations in local employment density. The non-transportable good is a stand-in for goods and services that are locally produced and consumed, of which housing services is clearly the most important. Figure 3 plots the logarithm of 1990 median house values in metropolitan areas against metropolitan employment density. The plot shows a positive relationship and yields a regression coefficient of 9.3×10^{-4}. Roback (1982) estimated a relationship between the logarithm of the site price of residential land and population density, controlling for several city-specific factors. Her estimate implies a γ value of 3.6×10^{-4}.[9] The baseline calibration will take 10.0×10^{-4} as the starting point for the value of γ, but sensitivity of results to values of γ between 6.0×10^{-4} and 14.0×10^{-4} will be considered. These parameter choices and ranges are summarized in Table 1.

Table 1

Calibration Guidelines of Some Model Parameters

ν	$\{0.01, 0.02, \mathbf{0.03}, 0.04, 0.05\}$
η	0.01
α	0.70
θ	0.50
γ	$\{6, 8, \mathbf{10}, 12, 14\} \times 10^{-4}$
a_i	ln(1990 area of CMSA/MSA i)

I now describe how the remaining parameters, namely, s_i and $\underline{d}_i$ are restricted. Additional restrictions come from requiring that, for any given values of the parameters in Table 1 and for any given choice of $\overline{N}$, the model reproduce the actual 1999 employment densities for each locality as a *stable* equilibrium outcome.

Denote the observed employment density in each locality in 1999 by d_i^{obs} and suppose, without any loss of generality, that the utility received by mobile individuals, v^*, is ln(100). Observe that

---

[8] See Chatterjee and Carlino (2001) for a more extensive discussion of the evidence on ν and θ.

[9] The coefficient on the density variable in her regression is 2.0×10^{-4} (Table 3, p. 1272). Since the median employment to population ratio for metropolitan areas in my data set is 0.57, Roback's estimate of the density coefficient implies a γ value of 3.6×10^{-4}.

for any choice of values for parameters listed in Table 1, and for any choice of $\overline{N}$, the implied values of $(\mu + \eta)/\delta$ and $\bar{d}_i$ divide the set of localities into two mutually exclusive groups. In one group are localities for which $\bar{d}_i < \ln[(\mu + \eta)/\delta]$ *and* $d_i^{\text{obs}} \in [\bar{d}_i, \ln[(\mu + \eta)/\delta]]$. For localities in this group, it follows from (27) that

$$\underline{d}_i = d_i^{\text{obs}}. \tag{29}$$

That is, all localities in this group are in a corner equilibrium. Furthermore, (28) implies an upper bound on the strength of locality-specific factors, namely

$$s_i < \ln(100) - (\mu + \eta)a_i - (\mu + \eta) \cdot d_i^{\text{obs}} - \delta \cdot e^{d_i^{\text{obs}}}. \tag{30}$$

For the other group, (23) implies

$$\underline{d}_i \leq d_i^{\text{obs}}, \tag{31}$$

and (25), implies

$$s_i \leq \ln(100) - (\mu + \eta)a_i - (\mu + \eta) \cdot d_i^{\text{obs}} - \delta \cdot e^{d_i^{\text{obs}}}. \tag{32}$$

In addition, (24) implies that

$$\left(\underline{d}_i - d_i^{\text{obs}}\right) \cdot \left(s_i - \ln(100) + (\mu + \eta)a_i + (\mu + \eta) \cdot d_i^{\text{obs}} + \delta \cdot e^{d_i^{\text{obs}}}\right) = 0. \tag{33}$$

This last restriction asserts that if there is at least one mobile individual in locality i, s_i must attain its upper bound and, conversely, if s_i does not attain its upper bound, there must not be any mobile individuals in locality i. The conditions (29)-(33) exhaust the restrictions on $\{s_i, \underline{d}_i\}_{i=1}^{M}$ implied by the requirement that $\{d_i^{\text{obs}}\}_{i=1}^{M}$ be consistent with a stable equilibrium.

The next set of restrictions come from assuming that $\underline{d}_i$ in (31) is equal to $\ln(N_{\min} - 1) - a_i$, where $N_{\min}$ is the lowest level of local employment observed for 1999. This assumption ensures that regardless of which localities end up in the second group, each has at least one mobile individual and, hence, each location's s_i attains the upper bound in (32). The assumption, therefore, pins down both the $\underline{d}_i$ and the s_i for all localities in the second group. Since $N_{\min}$ is a small number (120 to be exact), the vast majority of workers in localities falling into the second group get classified as mobile. In the approach being followed in this paper, the hypothetical relocation of workers *is* the way in which the influence of agglomeration economies is revealed and hence it's reasonable to minimize restrictions on mobility to the extent possible.[10]

---

[10]One could, for instance, set $\overline{d}_i = d_i^{\text{obs}}$ for localities in the second group as well but then, regardless of the values of s_i, there will be no change in the distribution of employment when agglomeration economies are suppressed.

The remaining set of parameters to be pinned down are the s for the first group. As noted earlier, the model and the data do not restrict the level of s for these localities other than to say that each must be strictly less than the upper bounds implied by the stability of local equilibria. To make the s "observable", I assume that the distribution of s for the localities that are in a corner equilibrium (first group) is no different from the distribution of s for localities that are in interior equilibrium (second group). I then use the distribution of s for localities in the *second* group to assign to each i in the *first* group the average value of s conditional on s_i not exceeding its upper bound in (30).[11]

V. COMPUTATIONAL EXPERIMENTS

In order to assess the contribution of agglomeration economies to the spatial concentration of employment, I need a way to describe and summarize the degree of spatial concentration in employment. An attractive way to do this is by using Lorenz curves. In the present context, a Lorenz curve is constructed by first ordering the localities by their employment density, with the most dense locality being ranked first. Then, the cumulative percentage of land areas (running from 0 to 100) of localities so ordered is plotted against the cumulative percentage of employment. If employment were uniformly distributed over the U.S. continental landmass, this plot would coincide with the 45-degree line. But if employment is not uniformly distributed, the curve will be bowed above the 45-degree line. As Figure 4 shows, the curve is indeed heavily bowed. The top 1 percent of the densest U.S. territory accounts for about 15 percent of total employment, the top 2 percent accounts for about 25 percent, and the top 15 percent accounts for 50 percent. The Gini coefficient associated with this Lorenz curve is 0.78. This summary measure of spatial concentration is used in this study.[12]

A. Restoring the Stability of Low-Density Localities via Restrictions on Mobility

In the first set of experiments the agglomeration threshold $\overline{N} = 1$. For each set of experiments, the starting calibration sets values of the economy-wide parameters to those noted in Table 1 and, in the

---

[11]Instead of assigning the conditional expectation, one could a assign a random s_i conditional on s_i not exceeding the relevant upper bound. One could then do this random assignment for each i (in the first group) many times over and calculate an equilibrium for each set of assignment. For the baseline calibration, this procedure led, on average, to almost the same results as reported in the text.

[12]See Carlino and Chatterjee (2002) for a discussion of Lorenz curves based on employment density.

case of parameters for which a range is specified (ν and δ), to the values noted in boldface. Thus, the starting calibration implies $(\mu + \eta)/\delta$ is 106 workers per square miles (rounded). Since $\overline{N} = 1$, and all localities have positive employment, any location with density less than or equal to 106 falls in the first group of localities, i.e., the group of localities that are in a corner equilibrium with $\underline{d}_i = d_i^{\text{obs}}$. In this instance, this group contains $2,321$ localities and about 25 percent of employment. The remaining 202 localities fall into the second group and are (by the final assumption made in the previous section) at an interior equilibrium; the s_i and $\underline{d}_i$ for the second group are determined as described earlier. The distribution of the s's observed for this second group is used to pin down the s_i's for the first group in the manner described at the end of the previous section.

For the counterfactual, a new equilibrium is calculated with the agglomeration parameters ν and η set to zero and all other parameters unchanged. In the absence of agglomeration economies, the $v_i(d)$ function is downward sloping for every locality and, consequently, there is no possibility of multiple equilibria. In this new equilibrium, differences in employment density across localities stem entirely from differences in fundamentals, i.e., differences in the s_i, a_i and $\underline{d}_i$.

Figure 5 plots the actual Lorenz curve for 1999 along with the Lorenz curve corresponding to the new equilibrium. The new (counterfactual) Lorenz curve lies inside the actual one, indicating that in the model agglomeration economies contribute unambiguously to the concentration of employment. The Gini index for the new Lorenz curve is 0.59, about 24 percent less than the actual Gini index of 0.78. Although a decline in the spatial concentration of 24 percent is a sizeable effect, the more remarkable feature of Figure 5 is the extent of spatial concentration remaining in the model after agglomeration economies are suppressed. For this experiment at least, agglomeration economies fail to emerge as the major contributor to spatial concentration. The bulk of the spatial concentration appears to be due to the spatial inequality in the fundamentals.

The proximate cause of this result is apparent from Figure 6. The figure displays both the actual employment density across localities in 1999 (with locations ranked according to declining employment densities) and each locality's counterfactual employment density. The horizontal line drawn at $\ln(106)$ divides the set of localities into the two groups: localities with density less than or equal to 106 are in the first group, (i.e., the group that is initially at corner equilibrium) and localities with density greater than 106 are in the second group (i.e., the group that is initially at interior equilibrium). Among the second group, elimination of agglomeration economies generally leads to a drastic decline in size (and density); with a handful of exceptions, localities in this

group essentially disappear as they drop down to the lowest employment level permitted in the counterfactual (120 workers). Among the first group, elimination of agglomeration economies leads generally to a substantial increase in employment density, with low density locations tending to gain the most. Thus, the currently dense locations become very sparse and the currently sparse locations become very dense. It's this switch which prevents the Gini index from falling steeply. The reason the index falls at all is that the new dense locations are much smaller in land area than the currently dense locations; relative to the actual Lorenz curve, a bigger percentage of land area is needed to account for any given share of total employment.

This pattern of employment relocation can be understood by looking at the difference in (log) utility of a mobile individual in a currently high density location and a currently low-density location. From the $v_i(d)$ function, this difference is

$$v_l(d_l) - v_h(d_h) = (s_l - s_h) - (\mu + \eta) \cdot (d_h - d_l) - (\mu + \eta) \cdot (a_h - a_l) + \delta \cdot (e^{d_h} - e^{d_l}),$$

where h and l are any two currently high and low density locations, respectively. The calibration of the model assigns a non-positive value to this difference. But when agglomeration economies are eliminated, this differences reduces to

$$v_l(d_l) - v_h(d_h) = (s_l - s_h) + \delta \cdot (e^{d_h} - e^{d_l}).$$

Since $d_h > d_l$, the second term contributes positively to the difference, i.e., low density localities look more attractive because of low congestion costs. What is interesting is that the first term generally contributes positively as well. Figure 6 plots the logarithm of location specific factors imputed to each locality, with localities ranked according to declining employment density. There is a clear upward tilt to the plot, with currently low-density localities being assigned superior TFP and/or amenity terms.[13] It's this combination of low congestion costs and generally superior fundamentals that leads to a re-emergence of employment concentration in currently sparse localities.

The method of assigning s_i to localities in the first group will tend to assign higher values to lower density localities. To see this, observe that for low-density localities the $v_i(d)$ function is locally increasing. Therefore, a locality that has the same land area but lower density relative to another locality will have a higher upper bound for s. Consequently, the average s conditional on s_i not exceeding the upper bound will also be higher. That's one reason for the upward tilt

---

[13]Recall that $s_i = \log(H) + \alpha^{-1} \cdot \ln(\phi_i) + \theta \cdot \ln(\xi_i^{-1}) + \ln(\psi_i)$, where ϕ_i and ξ_i^{-1} are the locality-specific TFP terms in the production of the transportable and local good, respectively, and ψ_i is locality specific amenity term.

in Figure 6. But the relatively low values assigned to the s_i in localities in the second group (i.e., the currently dense locations) reflects the fact that these localities benefit a great deal from agglomeration economies. If these locations also had generally superior TFP and/or amenity factors (relative to low density localities), the general equilibrium of the model would imply a higher level of density for these localities than is actually observed. An implication of this is that the assignment of s_i for this group of localities depends importantly on the magnitude of the (net) benefits of agglomeration. This is the point to which I turn next.

Table 2
Impact of Net Agglomeration Benefits on Spatial Concentration, $\overline{N} = 1$

Parameters: $\nu, \gamma \times 10^4$	% Decline in Gini	Mean s of 20 Most Dense Localities $\div$ Mean s of 20 Least Dense Localities	% Workers who are Immobile	# of Corner Localities
0.01, 14	35	1.04	12	1,920
0.02, 12	31	1.01	18	2,184
0.03, 10	**24**	**0.97**	**25**	2,321
0.04, 8	17	0.93	35	2,426
0.05, 6	10	0.88	55	2,478

For any given density level, the (net) benefit from agglomeration increases with $\mu + \eta$ and decreases with δ. Therefore, for the parameter ranges noted in Table 1, the net benefit from agglomeration is lowest when $\nu = 0.01$ and $\gamma = 14 \times 10^{-4}$ and highest when $\nu = 0.05$ and $\gamma = 6 \times 10^{-4}$. Table 2 shows that as the net benefit from agglomeration rises, the contribution of agglomeration economies to spatial concentration (as measured by the decline in the Gini index) *falls* from 32 percent to 9 percent. The third column, which measures the ratio of the average s_i assigned to the 20 most dense localities to the average s_i assigned to the 20 least dense localities, indicates the main reason why this happens. When the net benefits from agglomeration is low, the model replicates the observed concentration of employment by making the currently dense locations more attractive in term of fundamentals than the currently sparse localities. Conversely, when the net benefit of agglomeration is high, currently dense locations are assigned less favorable fundamentals than currently sparse localities. Consequently, eliminating agglomeration economies results in the re-emergence of concentration in the currently sparse localities when agglomeration benefits are high.

The final two columns in Table 2 report the fraction of workers who are assigned to be immobile and the number of localities in a corner equilibrium. The total number of workers classified as immobile is the sum of workers in all localities in a corner equilibrium plus $N_{\min}$ times the number of localities in an interior equilibrium. The percentage of immobile workers is simply this sum as a percentage of total employment. When net benefit from agglomeration is low, this fraction is low because the number of localities in a corner equilibrium is low. As net benefit from agglomeration rises, the number of localities in corner equilibrium rises and percentage of immobile workers rises. For the last row, over 50 percent of the workers are immobile. Because there is a tendency for employment to reconcentrate in currently low-density localities, these restrictions contribute to the decline in the Gini index. If these restrictions were to be removed, the Gini index would decline less (or might not not decline at all).

B. Restoring Stability of Low-Density Localities via an Agglomeration Threshold and Limited Restriction on Mobility

As noted earlier, empirical specifications of agglomeration economies do not consider threshold effects explicitly. However, it's worth exploring the implications of an agglomeration threshold for two reasons. First, empirical studies on agglomeration effects examine urban areas only and there is, therefore, an implicit assumption that agglomeration economies are not relevant for the large number of rural areas present in this study. By choosing $\overline{N}$ appropriately, agglomeration economies can be eliminated for rural areas. Second, an agglomeration threshold means that many low-density localities can be in a stable *interior* equilibrium because the $v_i(d)$ function is now locally decreasing (recall Figure 2). For these low-density localities, the calibration pins down the location-specific parameters and, hence, reduces the uncertainty about fundamentals.

Several values of the agglomeration threshold are examined in this section. The value of $\overline{N}$ I focus on is $\overline{N} = 35,000$. This value is only slightly smaller than the smallest employment level among metropolitan localities. Thus, with this value, rural counties do not experience any benefits of agglomeration while all urban localities do. Figure 8 displays the actual and counterfactual Lorenz curves with all all other parameters held at their starting values (i.e., at the values for which Figure 5 was drawn). The counterfactual Lorenz curve is now noticeably closer to the 45 degree line. The Gini index declines from 0.78 to 0.29, a decline of about 63 percent. Figure 9 shows the actual and counterfactual employment densities. Once again, there are some currently

high density localities that become very sparse but now there are no currently sparse localities that become very dense. Figure 10 plots the imputed s. There is clear endency for the imputed s to decline with density, and the mean s of the 20 most dense localities is now somewhat higher than the mean s of the 20 least dense localities.

The contrast between Figures 5 and 8 shows that the presence of a threshold for agglomeration economies can significantly raise the contribution of local increasing returns to spatial concentration. Since a threshold limits the scope of agglomeration economies, this result seems counter-intuitive. The key is to understand how an agglomeration threshold changes the imputation of s for low-density localities. A low-density locality is always attractive because it has low congestion costs. For a low-density locality to remain low density in a stable equilibrium, this attraction must be balanced by some disadvantage. In the absence of an agglomeration threshold, a low-density locality can be at a disadvantage simply because it's small and has low agglomeration benefits; in particular, it could have good fundamentals and still be low density. When there is an agglomeration threshold, the scope of agglomeration economies is constrained and a low employment level implies less of a productivity disadvantage for low-density localities. Consequently, to account for a low-density locality as a stable equilibrium, the disadvantage of small size now has to be *supplemented* with disadvantage of poor fundamentals. This is the reason why Figures 7 and 10 look so different. This difference implies that when agglomeration economies are suppressed, employment rises substantially in most low-density localities. The failure of concentration to re-emerge in a few currently sparse locations is the reason the Gini index declines more steeply.

There is another effect that also works toward generating a larger contribution of local increasing returns to spatial concentration. With an agglomeration threshold in place, the number of localities that are in a corner equilibrium drops from around $2,321$ to only 230. Consequently, the percentage of individuals who are immobile drops from 25 to 12 percent. Since more individuals are free to relocate, the pattern of s_i's has a larger influence on the distribution of individuals in the counterfactual and, as already discussed, this pattern is conducive to a dispersed distribution of employment rather than a concentrated one.

Table 3

Impact of Net Agglomeration Benefits on Spatial Concentration, $\overline{N} = 35,000$

Parameters: $\nu, \gamma \times 10^4$	% Decline in Gini	Mean s of 20 Most Dense Localities÷ Mean s of 20 Least Dense Localities	% Worker Who Are Immobile	# of Corner Localities
0.01, 14	35	1.07	2	47
0.02, 12	54	1.05	6	124
0.03, 10	**63**	**1.02**	**12**	230
0.04, 8	64	0.99	22	329
0.05, 6	53	0.96	42	380

Table 3 shows how increases in the net benefit of agglomeration alter the contribution of agglomeration economies. It's helpful to compare Table 3 and Table 2, row by row. For each row, the percentage decline in the Gini index is higher in Table 3 than in Table 2, the ratio of the mean s of the 20 most dense localities to the mean s of 20 least dense localities is higher in Table 3 than in Table 2, and the percentage of workers who are immobile is lower in Table 3 than in Table 2. These results are in line with the fact that an agglomeration threshold implies a more adverse imputation of s for currently sparse localities and less restrictions on mobility.

Reading down the columns of Table 3, one finds that the contribution of local increasing returns to spatial concentration is generally increasing in the net benefits of agglomeration. When the net benefits are low, the model imputes relatively generous s to the currently dense locations (otherwise, these localities would not be as dense as they are). Consequently, eliminating agglomeration economies does not lead to a big decline in the Gini index because the currently dense locations retain a fairly large fraction of workers. As the net benefit rises, the model imputes less favorable s to the currently dense localities and, consequently, leads to a more dispersed employment distribution in the counterfactual. This pattern is borne out until we reach the last row, at which point there is a fall in the contribution of agglomeration economies. This fall is due to the fact that for this parametrization more than 40 percent of workers are immobile.

Table 4

Impact of Threshold Size on Spatial Concentration, $\nu = 0.03, \gamma = 10 \times 10^{-4}$

Parameter $\overline{N}$	% Decline in Gini	Mean s of 20 Most Dense Localities÷ Mean s of 20 Least Dense Localities	% Workers Who Are Immobile	# of Corner Localities
1	**24**	**0.96**	**25**	**2321**
20,000	63	1.01	17	513
35,000	**63**	**1.02**	**12**	**230**
50,000	62	1.03	10	143
100,000	56	1.03	6	47
200,000	44	1.04	3	15
300,000	36	1.05	2	6

Table 4 displays how the contribution of agglomeration economies changes with the agglomeration threshold holding ν and γ fixed at 0.03 and 10×10^{-4}, respectively. Clearly there is nothing too special about the $\overline{N} = 35,000$; somewhat larger or somewhat smaller values also lead to a similar contribution of agglomeration economies. However, substantially higher values of $\overline{N}$ lead to smaller contributions. For $\overline{N} = 100,000, 200,000,$ and $300,000$, the Gini index declines 56, 44, and 36 percent, respectively. The reason the contribution of agglomeration economies fall off with substantially higher values of the agglomeration threshold is because the scope of agglomeration economies gets increasingly limited. There is simply less increasing returns with higher threshold values and therefore less contribution of increasing returns to spatial concentration.

As noted earlier, an agglomeration threshold can substitute for mobility restrictions in accounting for the stability of low-density localities. It's worth finding out what the contribution of increasing returns would be if the threshold is so high that there are no localities in a corner equilibrium. For each row in Table 5, the agglomeration threshold is set to the smallest value (to the nearest 100) consistent with no locality being in a corner equilibrium. The contribution of agglomeration economies is fairly modest in each case, and rises for the first three rows. Over this range, the agglomeration threshold does not change but the net benefit from agglomeration goes up leading to a rise in the contribution of increasing returns to spatial concentration. For the fourth and fifth row, there are substantial jumps in the agglomeration threshold. The agglomeration threshold jumps up because with relatively high net benefits from agglomeration, there are several large cities with density less than $(\mu+\eta)/\delta$. Since the agglomeration threshold needs to be set above

the maximum employment of localities with density less than $(\mu + \eta)/\delta$, it becomes larger as we go down the rows. Because of these jumps, which limit the scope of agglomeration economies, there is now a decline in the contribution of agglomeration economies to spatial concentration. There is essentially no contribution of increasing returns in the final row because very few localities actually benefit from agglomeration economies.

Table 5

Impact of Net Agglomeration Benefits with No Restrictions on Mobility

Parameters: $\nu, \gamma \times 10^4$	% Decline in Gini	Mean s of 20 Most Dense Localities÷ Mean s of 20 Least Dense Localities	$\overline{N}$	# of Corner Localities
$0.01,\ 14$	6	1.09	$880,500$	0
$0.02,\ 12$	12	1.08	$880,500$	0
$\mathbf{0.03, 10}$	**19**	**1.06**	$880,500$	0
$0.04,\ 8$	13	1.06	$1,879,700$	0
$0.05,\ 6$	0	1.05	$8,868,700$	0

It is of some interest to compare the third row of Table 5 with the third row of Table 2. For this row, ν and γ are at their baseline values. When there is no agglomeration threshold, the contribution of increasing returns to spatial concentration is 24 percent; if there is an agglomeration threshold that's high enough to eliminate all restrictions on labor mobility, the contribution of increasing returns is 19 percent. Thus, in these two polar cases, the contribution of increasing returns is quite similar. This comparison is of interest because the one study that did estimate something close to an agglomeration threshold found it to be in the vicinity of 1.1 million workers.[14]

---

[14]Segal (1976) estimated the production function (1) directly using labor and capital stock data on 58 SMSAs for 1967. He used SMSA population rather than employment as the size variable and distinguished between large and small MSAs. He found measurable productivity differences for SMSAs with populations above and below 2 million and estimated that difference to be about 8 percent in favor of large SMSAs. Because Segal used broad population categories to measure SMSA size, his estimate of an 8 percent productivity differential in favor of large SMSAs is not an estimate of ν. On the other hand, his finding that productivity differences are discernible between the group of SMSAs with populations greater than 2 million may be interpreted as evidence that $\overline{N}$ is around 1.1 million workers (in my dataset, the ratio of employment to population is around 0.56),

VI. Summary and Conclusions

This paper presented a novel approach to answering a scientifically deserving question in economic geography: What is the relative contributions of natural advantage and increasing returns in the spatial concentration of employment? The question is scientifically deserving not only because it's obviously central to a proper understanding of economic geography but also because it is not easy to answer. If one observes densely populated localities, does one conclude that the density results primarily from natural advantages or does one conclude that it results primarily from local increasing returns?

In the past, researchers answered this question by measuring variation in natural resources and then determining how much of the observed variation in economic activity (or concentration of economic activity) can be explained on the basis of the observed variation in natural resources. The residual variation is taken as an upper bound on the contribution of increasing returns. As noted in Ellison and Glaeser (1999), this approach requires taking a stand on the set of natural resources that matter for economic geography. Otherwise, one could eliminate *any* role for increasing returns by including as many dimensions of variation in natural resources as there are observations! However, the more one restricts the set of natural resources, the looser is the upper bound on the contribution of increasing returns. It's fair to say that this approach is well-suited to measuring the impact of spatial distribution of specific natural resources but less well-suited for measuring the contribution of increasing returns to spatial concentration.
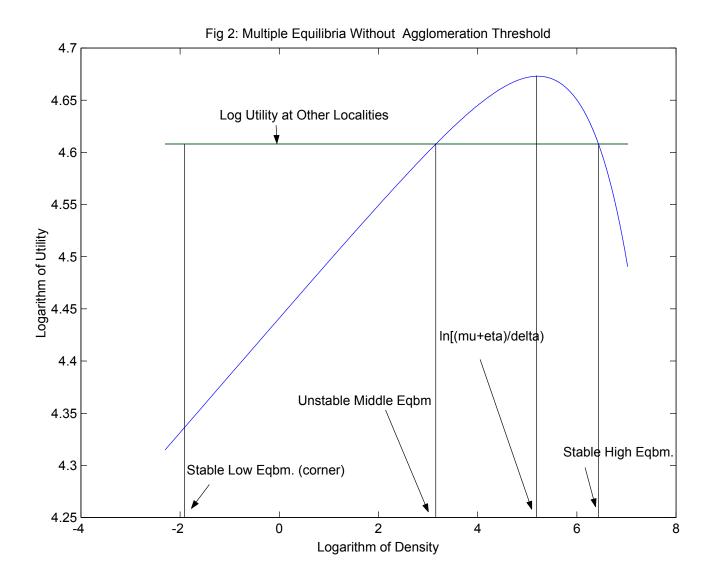
In contrast, the approach taken in this paper is to start with estimates of the magnitude of agglomeration economies and use a general equilibrium model to determine how much of the spatial variation in economic activity can be plausibly explained on the basis of these estimates. The approach is analogous to "business-cycle accounting" or "growth accounting," familiar to us from the quantitative macroeconomics literature. As is the case for these other accounting exercises, the results of the "spatial accounting" performed in this study depends on the details of the model used. In this instance, the critical detail pertains to how the model rationalizes the stability of low-density localities. If it is rationalized via an appeal to restrictions on labor mobility, the accounting implies that the bulk of spatial concentration results from an unequal distribution of natural advantages. In contrast, if it is rationalized via an agglomeration threshold (an employment level below which local increasing returns do not operate), the accounting implies that the bulk of the spatial concentration results from increasing returns.

In terms of research priorities, a better understanding of the reasons for the persistence of low-density localities is warranted. In particular, locating evidence, or lack thereof, of an agglomeration threshold would be a worthwhile research effort. Although the literature on agglomeration economies hasn't stressed the possibility of an agglomeration threshold, threshold effects have received attention in the wider literature on economic geography. For instance, geographers who study the spatial diffusion of innovations using epidemiological models have observed that an innovation might fail to diffuse if the population density is below some critical threshold.
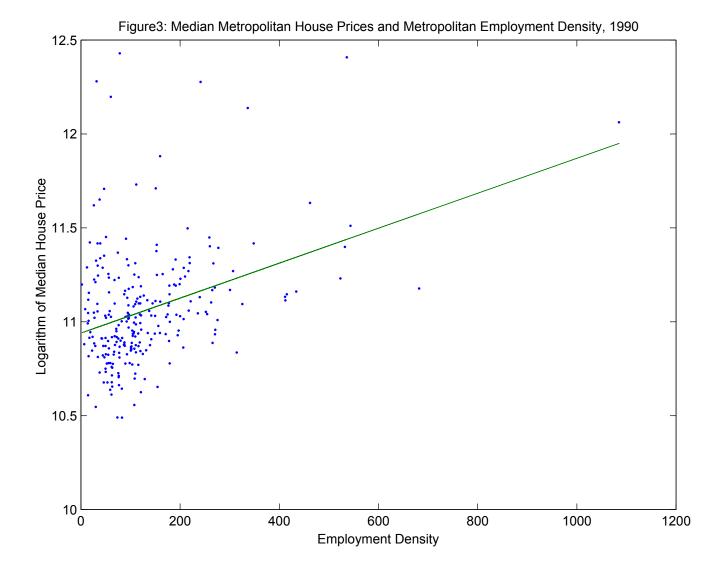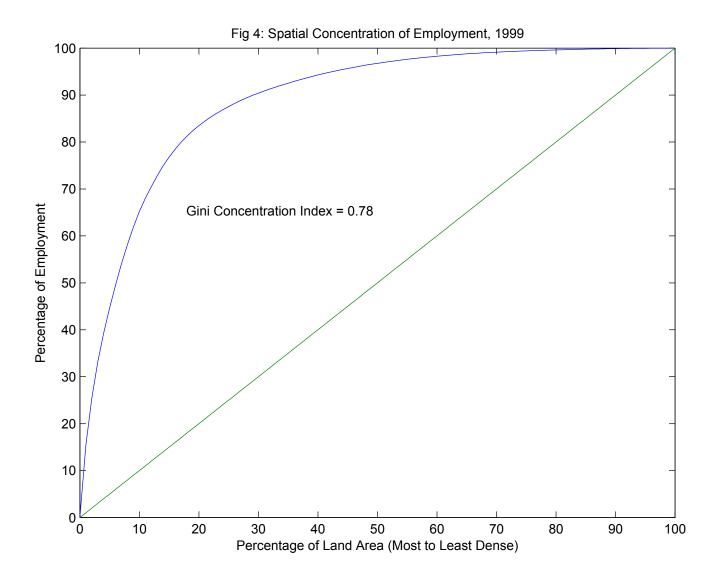
References

[1] Carlino, Gerald and Satyajit Chatterjee. "Employment Deconcentration: A New Perspective on America's Postwar Urban Evolution," *Journal of Regional Science*, 42, pp. 455-476, 2002.

[2] Chatterjee, Satyajit and Gerald Carlino. "Aggregate Metropolitan Employment Growth and the Deconcentration of Metropolitan Employment," *Journal of Monetary Economics*, 48, pp. 549-583, 2001.

[3] Ciccone, A. and Hall, R. E., 1996, Productivity and the Density of Economic Activity, *American Economic Review* 86, 54-70.

[4] Davis, Donald R., and David E. Weinstein. "Bombs, Bones, and Breakpoints: The Geography of Economic Activity," *American Economic Review*, 2002a.

[5] ——. "A Search for Multiple Equilibria in Urban Industrial Structure," Columbia University, 2002b.

[6] Ellison, Glenn and Edward L. Glaeser, "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?" *American Economic Review*, May, 1999, 311-316.

[7] Fujita, Masahisa, Paul R. Krugman, and Anthony J. Venables. *The Spatial Economy: Cities, Regions, and International Trade.* Cambridge: The MIT Press, 1999.

[8] ——, and Jean-Francoise Thisse. *Economics of Agglomeration.* Cambridge: Cambridge University Press, 2002.

[9] Gollin, Douglas. "Getting Income Shares Right," *Journal of Political Economy*, 110, 2002, pp. 458-474.

[10] Hanson, Gordon, H. "Market Potential, Increasing Returns, and Geographic Concentration," University of California, San Diego, December 2001.

[11] Jacobs, Eve E., and Shipp, Stephanie, 1990, "How Family Spending Has Changed in the US," *Monthly Labor Review*, March, 20-27.
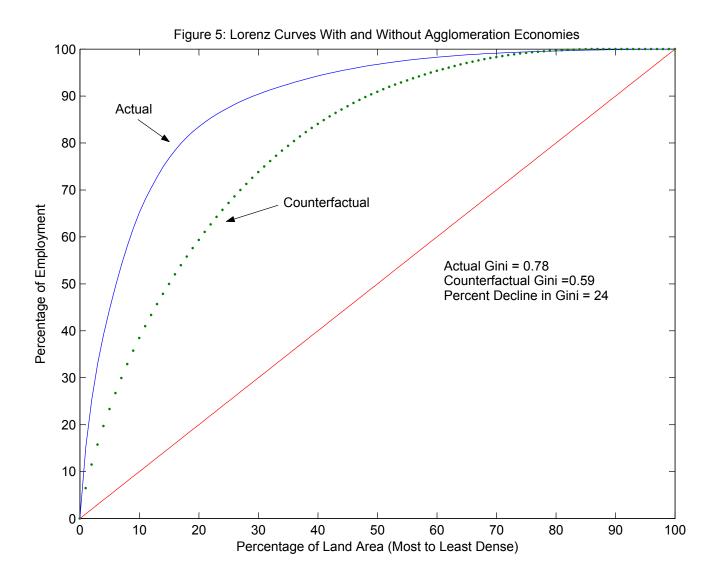
[12] Kim, Sukkoo. "Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in U.S. Regional Manufacturing Structure, 1860-1987," *Quarterly Journal of Economics*, 110, 1995, 881-908.

[13] _____. "Regions, Resources, and Economic Geography: Sources of U.S. Regional Comparative Advantage, 1880-1987," *Regional Science and Urban Economics*, 29, 1999, 1-32.

[14] Krugman, Paul R. "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99, pp. 483-499, 1991.

[15] Moomaw, Ronald L., 1981, Productivity and City Size: A Critique of the Evidence, T*he Quarterly Journal of Economics*, 116, 675-89.

[16] Rappaport, Jordan and Jeffrey D. Sachs. "US as a Coastal Nation," Federal Reserve Bank of Kansas City, RWP 01-11, October 2002.

[17] Roback, Jennifer, 1982, Wages, Rents and the Quality of Life, *Journal of Political Economy*, 90, 1257-78.

[18] Rosenthal, Stuart S. and William C. Strange. "Evidence on the Nature of and Sources of Agglomeration Economies," mimeo, April 21, 2003. Forthcoming in the *Handbook of Urban and Regional Economics*, Volume IV.

[19] Segal, David, 1976, Are There Returns to Scale in City Size? *Review of Economics and Statistics*, 58, 339-50.

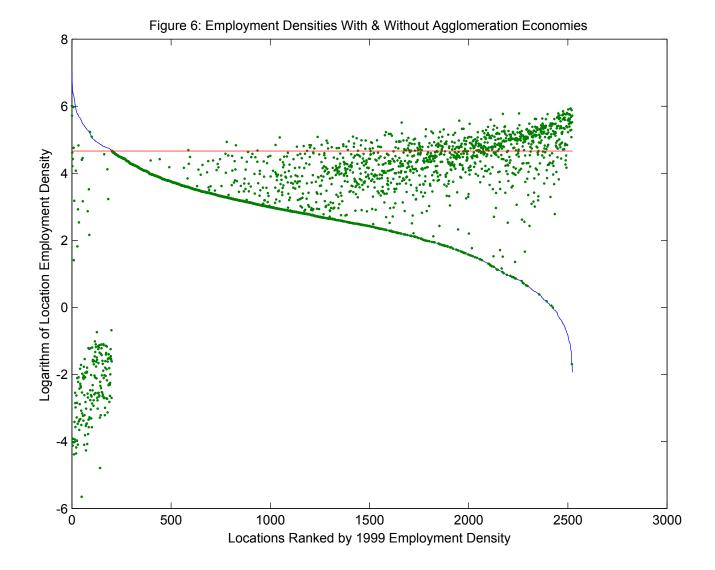[20] Sveikauskas, Leo, 1975, The Productivity of Cities, *The Quarterly Journal of Economics*, 89, pp. 393-413.

Figure 1: Multiple Equilibria With a Positive Agglomeration Threshold

Fig 2: Multiple Equilibria Without Agglomeration Threshold

Figure3: Median Metropolitan House Prices and Metropolitan Employment Density, 1990

Fig 4: Spatial Concentration of Employment, 1999

Gini Concentration Index = 0.78

Percentage of Employment

Percentage of Land Area (Most to Least Dense)

Figure 5: Lorenz Curves With and Without Agglomeration Economies

Actual

Counterfactual

Actual Gini = 0.78
Counterfactual Gini =0.59
Percent Decline in Gini = 24

Percentage of Employment

Percentage of Land Area (Most to Least Dense)

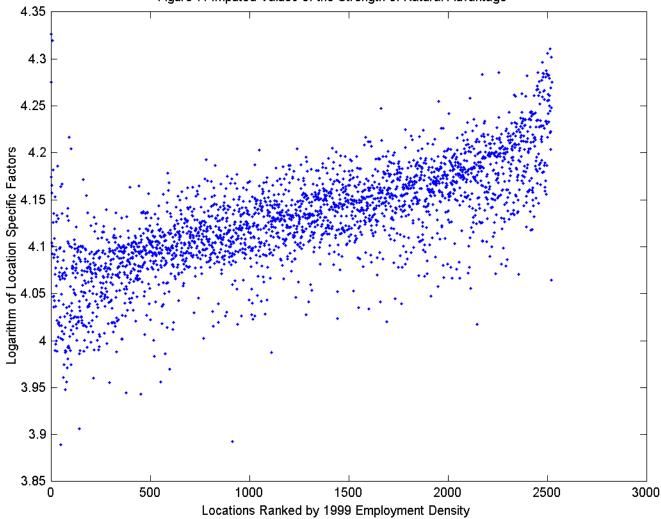Figure 6: Employment Densities With & Without Agglomeration Economies

Figure 7: Imputed Values of the Strength of Natural Advantage

Figure 8: Lorenz Curves With and Without Agglomeration Economies, Agglomeration Threshold = 35,000

Actual

Counterfactual

Actual Gini = 0.78
Counterfactual Gini = 0.29
Percent Decline in Gini = 63

Percentage of Employment

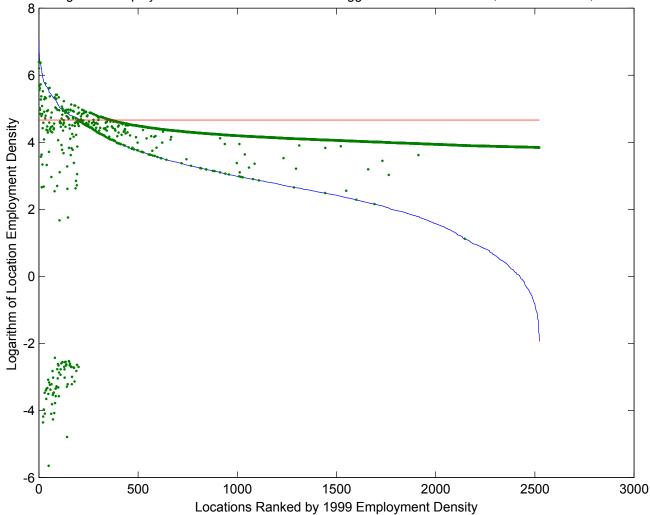Percentage of Land Area (Most to Least Dense)

Figure 9: Employment Densities With & Without Agglomeration Economies, Threshold = 35,000
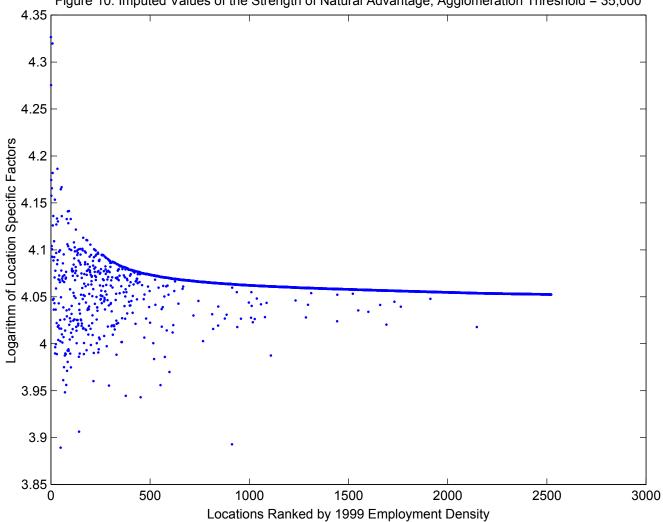
Figure 10: Imputed Values of the Strength of Natural Advantage, Agglomeration Threshold = 35,000