Agglomeration Economies: The Spark That Ignites a City?

BY SATYAJIT CHATTERJEE

In industrially developed countries, employment is heavily concentrated in cities. A concentration of workers and businesses in one location — what economists call agglomeration economies — lowers production costs. In fact, most economists believe that in the absence of agglomeration economies, the spatial distribution of employment would be much more even. In this article, Satyajit Chatterjee discusses his research, which questions this belief. He finds that while agglomeration economies are an important factor, they're *not* the most important one. The combined effects of factors unrelated to agglomeration economies, such as the availability of natural resources and local economic policies, appear to account for the bulk of the spatial concentration of U.S. employment.

The bulk of an industrially developed country's economic activity takes place in cities. Typically, these cities make up a relatively small portion of the country's overall territory. For instance, 83 percent of total employment in the U.S. is located in metropolitan areas, and these areas account for 24 percent of the total land area of the country.

Satyajit Chatterjee is a senior economic advisor and economist in the Philadelphia Fed's Research Department.

Why is employment so heavily concentrated in selected areas of the country? Economists think that spatial concentration of employment (or, more generally, economic activity) develops for two very different reasons. The first reason — and one that comes most readily to mind — is that a location attracts people and businesses because of the presence of some valuable natural resource. Petroleum, coal, lumber, minerals, and proximity to a navigable river or to the coast are all examples of valuable natural resources. Because such resources are not available everywhere, people and businesses end up flocking to resource-rich areas.

However, the natural resource reason does not explain the full extent of the remarkable spatial concentration we see in reality. For instance, access to a deep harbor was no doubt important for the emergence of Philadelphia as a colonial city, but can it be the main reason for Philadelphia's subsequent evolution into one of America's pre-eminent metropolitan areas? Studies of urban evolution suggest a second reason for spatial concentration: A concentration of workers and businesses in one location lowers production costs because proximity permits workers and businesses to save on the costs of transporting goods and people. Economists refer to this cost advantage as economies of spatial concentration, or *agglomeration* economies, for short.

Agglomeration economies can be a powerful force for attracting large numbers of people to a given location. They can cause a location with some small advantage in terms of natural resources to become a place with a large concentration of diverse businesses and households. While the natural resource initially attracts businesses and households to the location, this original group then becomes the factor that attracts other businesses and households to that location. As the location grows in size, business costs fall and the location's attractiveness as a potential spot for other businesses and households rises, and more people and businesses move in.

Although rising congestion eventually chokes off the inflow of people, agglomeration economies can be the spark that ignites the development of a city.

Economists generally believe that agglomeration economies are the primary factor that leads to the large clusters of people and jobs we see in the real world. In other words, most economists believe that in the absence of agglomeration economies, the spatial distribution of employment would be much more even.

In this article I discuss my research, which tried to determine if this belief is, in fact, accurate. My research indicates that while agglomeration economies are an important contributor to the spatial concentration of employment, they're *not* the most important factor. Contrary to expectations, factors other than agglomeration economies appear to account for the bulk of spatial concentration. It's not clear exactly what these other factors are, but they could be differences in the availability of natural resources across metropolitan areas, differences in economic policies across cities and states, or some other advantage of spatial concentration distinct from agglomeration economies. Whatever the case, my research suggests that agglomeration economies are probably just one of several important factors affecting spatial concentration of employment.

THE FACT OF SPATIAL CONCENTRATION

To determine the contribution of agglomeration economies to spatial concentration, we need a measure of the extent of spatial concentration in U.S. employment. An effective way to do this is by using a Lorenz curve, a graphical tool originally developed to show the extent to which income is unevenly distributed across people.[1] But Lorenz curves can also be used to show how unevenly employment is distributed across space.

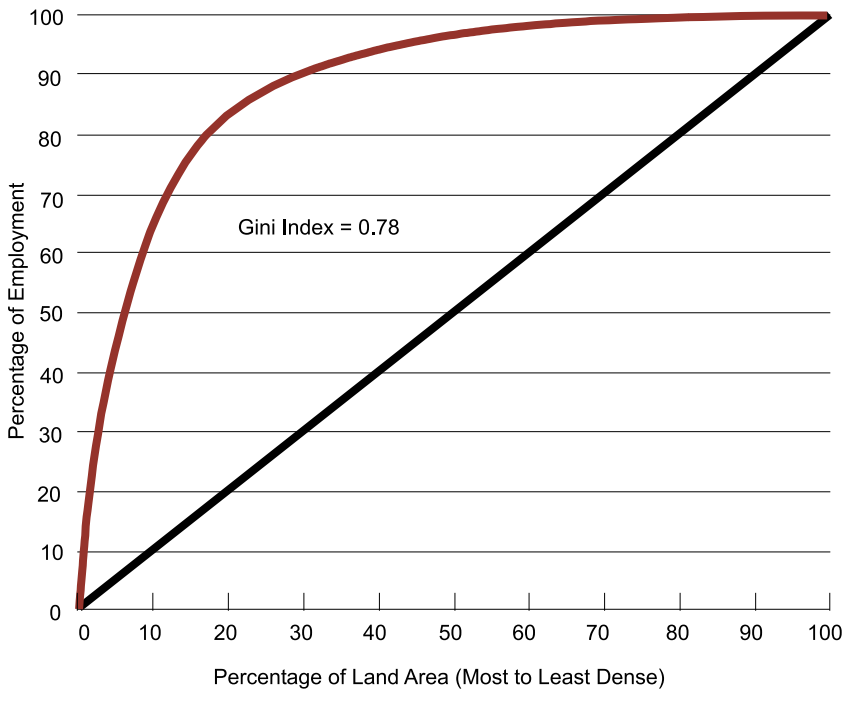
To construct a Lorenz curve of spatial concentration, I first ranked metropolitan areas and rural counties in the continental United States by their employment density, the densest areas being ranked first. Using this ranking, I then calculated the percentage of employment accounted for by the first, or top, 1 percent of the total continental land area, then the top 2 percent, and so on. The Lorenz curve is simply a graph that plots these calculations (Figure 1). If employment were uniformly distributed over the continental landmass, this graph would coincide with the 45-degree line shown in the figure. That is, the top 1 percent of the continental land area would account for 1 percent of employment, the top 2 percent of the area would account for 2 percent of employment, and so on. But if employment is not uniformly distributed, the graph will be bowed above the 45-degree line — as, in fact, it is.

As Figure 1 indicates, the top 1 percent of total continental land area accounts for about 15 percent of employment, the top 2 percent accounts for about 25 percent, and so on. Indeed, by the time we include the top 20 percent of the continental land area, we can account for more than 80 percent of total employment! Clearly,

[1] The statistician Max O. Lorenz (1880-1962) developed the Lorenz curve. The curve is probably the tool most used to analyze income and other distributions. Remarkably, Lorenz came up with the idea of the curve in his undergraduate thesis at the University of Iowa, circa 1894, at the age of 14! He went on to have a distinguished career, becoming the chief statistician of the Interstate Commerce Commission in Washington, D.C.

FIGURE 1

Spatial Concentration of U.S. Employment, 1999



U.S. employment is very unevenly distributed over space.

The Lorenz curve is an effective visual representation of the degree of spatial concentration of employment. It also provides the basis for the Gini index, a well-known index of concentration. The Gini index is a number between zero and one, and it is a measure of the difference between the Lorenz curve and the 45-degree line. It is computed by dividing the area between the Lorenz curve and the 45-degree line by the total triangular area above the 45-degree line. When employment is uniformly distributed, the Lorenz curve coincides with the 45-degree line, and the Gini index is zero. The more unevenly employment is distributed, the more bowed the Lorenz curve and the larger the area between the curve and the 45-degree line. Thus, the Gini index is higher for a more uneven distribution of employment and lower for a more even one. In Figure 1, the value of the Gini index is 0.78, which means the area between the 45-degree line and the bowed line represents close to 80 percent of the total area above the 45-degree line. This is the measure of spatial concentration I used in my research.

NATURE AND MAGNITUDE OF AGGLOMERATION ECONOMIES

As mentioned earlier, agglomeration economies arise because proximity permits workers and businesses to save on the costs of transporting goods and people. In this section I'll highlight one way in which this happens, then discuss what economists know about the magnitude of agglomeration economies in the U.S.

One reason agglomeration economies arise is that a large concentration of workers allows a business to deal more effectively with fluctuations in the volume of sales. Consider a business whose future demand can be either high or low, with equal probability. When demand is high, the business needs four workers; when demand is low, it needs only two. The business has to hire workers *before* it knows how large demand will be. Suppose the business chooses to hire three workers. If demand turns out to be low, workers work at two-thirds capacity, and all demand is met. If demand turns out to be high, all workers work at full capacity, but one-quarter of demand is not met. So there is a 50 percent chance that every worker works at less than full capacity.

Agglomeration economies arise because proximity permits workers and businesses to save on the costs of transporting goods and people.

Now imagine that another enterprise in the same line of business moves into the area and this enterprise faces a similar uncertainty with respect to demand. However — and this is the key assumption — the level of the new firm's demand is independent of the level of the first firm's demand. This may happen if the firms have different sets of customers and serve different markets. This means that the combinations of demand across the two firms can take one of four possibilities, all with equal probability: (high, high), (high, low), (low, high), and (low, low). Now, when the two businesses have different levels of demand (which happens with probability one-half), the firm with low demand has an incentive to rent out its one excess worker to the firm with high demand. This is feasible because both firms are in the same location and the cost of moving workers between firms is presumably low. If the two firms shifted workers between them in this way, the only time any worker would work at less than full capacity is when demand at *both* firms is low, which happens with probability one-quarter.

The movement of workers between businesses in the same location does happen in reality, although it takes the guise of contract workers selling their services to businesses on a temporary basis. For instance, we might have a situation where both businesses hire two permanent employees, and each business has the option to hire additional contract employees in the event the level of demand is high. In this arrangement, there are four permanent workers and two contract workers. The permanent workers always work at full capacity while contract workers have a 75 percent chance of working at full capacity or a 25 percent chance they won't work at all. Contract workers take on the risk of unemployment, but if the two firms use some of their cost savings to pay contract workers more than full-time employees, contract workers might feel compensated for the risk.

To summarize, physical proximity makes it possible for firms to share workers and so allows businesses to take advantage of the fact that the combined demand of several firms is more stable than the demand of a single firm. This stability permits a group of businesses to better utilize workers than a single business. The improved utilization of workers lowers business costs and provides a reason for firms and workers to cluster together.

Let's turn now to a description of the strategies economists have used to estimate the magnitude of ag-

glomeration economies that stem from better utilization of workers. The most direct way to do this is to measure changes in the utilization of workers due to spatial concentration. However, because it's not easy to directly measure how hard employees work, economists have used more indirect methods. Let's look at two of these methods along with the estimates of agglomeration economies obtained using each one.

The first method uses information on labor hours and equipment purchased (also called capital) and goods (output) sold by different industries in different metropolitan areas. For any given industry, labor and capital purchased will have a higher utilization rate in metro areas with a large concentration of workers and firms. Thus, for any given industry and for any given amounts of labor and capital, more output will be produced in a large metro area than in a small one. The estimate we get from this method suggests that agglomeration economies make businesses in metro areas with more than 2 million people 8 percent more productive than businesses in metro areas with less than 2 million people.[2]

The second method uses information on hourly wages businesses pay to workers. Businesses that use workers more effectively face lower costs and so make higher profits. Given that, a business would be motivated to locate in a large metro area rather than a small one. But when businesses do so, they compete with one another and end up paying more for each worker they hire. In other words, in a competitive environment, higher worker productivity will result in higher wages being paid to workers in large metro areas. By measuring the wages paid to similarly skilled workers in metro areas of varying sizes, we can estimate how much more productive workers are due to agglomeration effects. Studies that follow this approach have found that as a metro area doubles in size, the productivity of its workers rises 3 percent.[3]

AGGLOMERATION ECONOMIES' CONTRIBUTION TO SPATIAL CONCENTRATION

Given these estimates of the magnitude of agglomeration economies, the question is: How important are these agglomeration effects for the spatial concentration of employment? Answering this question involved two steps.

First, I constructed an economic model of local employment that can exactly reproduce the Lorenz curve in Figure 1, which gives the distribution of workers across metropolitan areas and rural counties in 1999. Second, I constructed a new Lorenz curve for a model economy that's identical to the one in the first step except that in this model, there are no agglomeration economies. If the Lorenz curve for this new model economy turns out to be close to the 45-degree line, I can reasonably conclude that agglomeration effects account for the bowed shape of the Lorenz curve in Figure 1. More generally, any difference between the Lorenz curve in Figure 1 and the Lorenz curve predicted by the model with no agglomeration effects can be attributed to the effects of agglomeration economies. In particular, the difference between the Gini indexes for the two Lorenz curves is a measure of the contribution of agglomeration effects to the spatial concentration of U.S. employment.

Description of the Model Economy. Briefly, the macroeconomic model in the first step has the following features.[4] There is a given set of locations, corresponding to the 275 metropolitan areas and 2,248 rural counties in the continental U.S.[5] Each location can produce two types of goods. One type, which I call *traded goods*, can be shipped without cost to other locations; the second type, which I call *local goods*, cannot be shipped at all. A household living in a given location derives benefit (or what economists call utility) from the consumption of the traded good and from consumption of the local good produced in that location. (The household cannot consume the local good of other locations because local goods cannot be shipped.)

Locations differ in terms of natural resources. In my model, the natural resources available to a location affect the productivity of labor and capital employed in the production of the traded good in that area. It may also affect how much enjoyment a household gets from living there. A location that has high productivity due to the presence of some natural

[2] Reported in David Segal's article.

[3] This estimate is the median value of agglomeration economies across manufacturing industries reported in Leo Sveikauskas's article.

[4] With some modifications, this is the same model I have used in previous research. The details of the model are in my article with Gerald Carlino.

[5] The 275 metropolitan areas consist of 258 primary metropolitan areas and 17 consolidated metro areas. A consolidated metropolitan area is a group of neighboring primary metro areas between which there is a significant amount of commuting.

resources will attract firms making the traded good; an area that's pleasant to live in because it has some other natural amenity will attract households.

As a location with some natural advantage attracts businesses and households, it gains employment. The rise in employment generates agglomeration economies and lowers business costs. This serves to make the location more attractive to businesses, and more businesses move in and create jobs. However, the people who move in to take these jobs make the location increasingly congested, and this congestion causes the price of the local good to rise. The rising price of the local good reduces the purchasing power of the wages workers receive in that location and limits the inflow of workers. The migration of workers between locations will make the wage (adjusted for amenities) equal across all metro areas, and every person seeking work will be employed in some location.

In this model, the distribution of employment across locations reflects the availability of natural resources in each area, the magnitude of agglomeration economies, and the magnitude of congestion costs. The magnitude of the agglomeration effects in the model is consistent with the evidence on agglomeration effects noted in the previous section. Also, the magnitude of congestion costs is consistent with the evidence on congestion costs that researchers have found for U.S. metro areas.

Finally, the model's parameters use values that determine the effects of natural resources on employment, so that the employment density in each metro area and rural county in the model exactly matches the employment density of that metro area or rural county in reality. This final step makes it possible for the model to exactly reproduce the Lorenz curve shown in Figure 1.

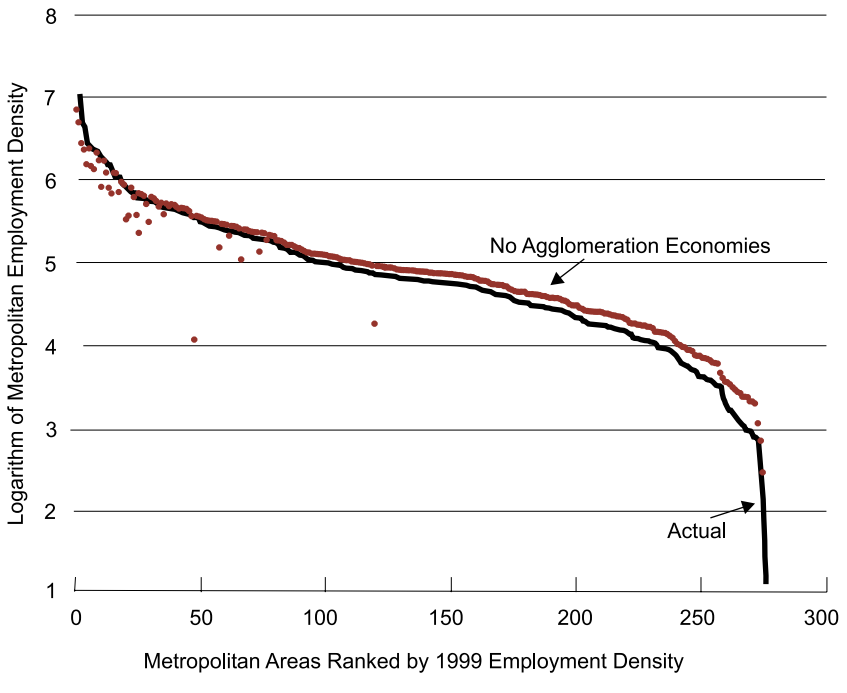
What Does the Model Say About the Role of Agglomeration Economies in Spatial Concentration?

Using this model I can investigate the role of agglomeration economies in the spatial concentration of U.S. employment. As noted earlier, my strategy for doing this is to examine what happens to the spatial distribution of employment in my model when I eliminate the reduction in production costs due to agglomeration economies while keeping all other aspects of the model unchanged. The solid black line in Figure 2 plots actual employment densities for metro areas in 1999; the dotted line plots what happens to employment densities in these metro areas when agglomeration effects are removed. As the figure shows, a relatively small set of high-density locations become less dense and a large set of relatively low-density locations become denser. The first set includes large metro areas, which benefit the most from agglomeration economies. These metro areas shed employment because they can no longer productively employ as many workers. Workers from these metro areas end up moving to smaller metro areas (and also to rural counties not shown in the figure), and consequently, these areas become denser.

The table lists the top 20 metro areas for which agglomeration economies seem most important. As one would expect, big cities like New York, Los Angeles, Chicago, and Atlanta are on the list. Los Angeles appears to be the city that benefits most from agglomeration economies in that almost 80 percent of its jobs would disappear if agglomeration economies were absent; Phoenix-Mesa is another area that appears to owe a lot of its employment to agglomeration econo-

FIGURE 2

Metropolitan Employment Densities With And Without Agglomeration Economies



mies. Philadelphia also makes the list and appears to owe 20 percent of its jobs to agglomeration economies.

Clearly, agglomeration economies appear to be very important for the development of specific cities, especially Los Angeles and Phoenix-Mesa. But how important is it generally? Figure 3 helps to answer this question. It compares the Lorenz curve when agglomeration effects are removed from the model constructed in step 1 with the Lorenz curve from Figure 1. The new Lorenz curve is less bowed, indicating that in the absence of agglomeration economies, employment is more evenly distributed. The Gini index declines about 16.5 percent, from 0.78 to 0.65.

The most striking feature of the new Lorenz curve is that it's still pretty far from the 45-degree line. Even in this world without agglomeration economies (but which is otherwise similar to the U.S. in important respects), there is considerable spatial concentration of employment. In other words, although the contribution of agglomeration economies is substantial, it's not as large as we might have expected. Recall that most economists consider agglomeration economies the most important reason for spatial concentration. But my model predicts that the U.S. would continue to be spatially concentrated, that is, have very dense areas, even if agglomeration economies were completely absent. Apparently, agglomeration economies are generally not needed to spark the development of cities! [6]

What, Then, Are the Other Determinants of Spatial Concentration? If agglomeration economies are not the key contributor to spatial concentration, what is? Taken at face value, my model suggests that it's the uneven distribution of natural resources that accounts for the bulk of spatial concentration. Indeed, some researchers have suggested that access to a navigable river or coast is, in fact, a key determinant of spatial concentration in the U.S.[7] Nevertheless, it's not accurate to say that any concentration left unexplained by agglomeration economies must result from the effects of natural resources. There are other factors, besides geography, that might affect spatial concentration and that are not captured in my simple model.

One potentially important factor is city- or state-specific economic policies. If an area happens to be located in a state with pro-business laws and regulations, it will have an advantage in terms of job creation relative to other areas.[8] Another factor could be the cost savings from transporting goods from one region to another.[9] For instance,

[6] It's possible that economists may have mismeasured the magnitude of agglomeration economies and congestion costs, thus affecting the values built into my model. However, when I varied the model's magnitude of agglomeration economies and congestion costs within plausible ranges (while ensuring that the model exactly reproduced the Lorenz curve in Figure 1), the drop in spatial concentration from elimination of agglomeration economies rarely exceeded 50 percent. Therefore, even with generous allowances for mismeasurement, agglomeration economies do not appear to account for the bulk of spatial concentration.

[7] See the article by Jordan Rappaport and Jeffrey Sachs.

[8] The article by Thomas Holmes presents evidence that state policies affect the location of industry.

[9] The cost savings from shipping goods within metro areas are captured in the estimates of agglomeration economies used in my model.

TABLE

Metropolitan Areas	Percentage of Employment Due to Agglomeration Economies
Los Angeles-Riverside-Orange County	79
Phoenix-Mesa	48
Dallas-Fort Worth	32
Washington-Baltimore	29
Houston-Galveston-Brazoria	28
Denver-Boulder-Greeley	27
Seattle-Tacoma-Bremerton	25
Detroit-Ann Arbor-Flint	23
San Francisco-Oakland-San Jose	23
Atlanta	22
Boston-Worcester-Lawrence-Lowell-Brockton	22
Minneapolis-St. Paul	22
St. Louis	22
Chicago-Gary-Kenosha	20
Philadelphia-Wilmington-Atlantic City	**20**
New York-Northern New Jersey-Long Island	19
Portland-Salem	18
San Diego	13
Cleveland-Akron	12
Pittsburgh	11

part of Philadelphia's attraction as a business location is its proximity to two other large metro areas: Washington, D.C. and New York City. Philadelphia's proximity to these two places means that businesses in Philadelphia can ship goods relatively cheaply to two other large metro areas, thus giving them relatively cheap access to a very large customer base.[10] A third factor could be that some benefits of spatial concentration go beyond reducing the costs of producing goods and services. It's well known, for instance, that most inventive activities take place in cities. Just as spatial concentration can reduce the costs of producing goods and services, it may also reduce the costs of producing new knowledge through better utilization of knowledge workers.[11]

SUMMARY

Economists have generally pointed to agglomeration economies as the principal reason a country's employment tends to get concentrated in a relatively small number of geographic areas. Agglomeration economies refer to the reduction in business costs that results from a concentration of businesses and workers in the same geographic area. This reduction in business costs provides incentives for workers and firms to cluster together, despite the costs associated with increased congestion. Several empirical studies have found evidence of significant agglomeration economies in U.S. metro areas.

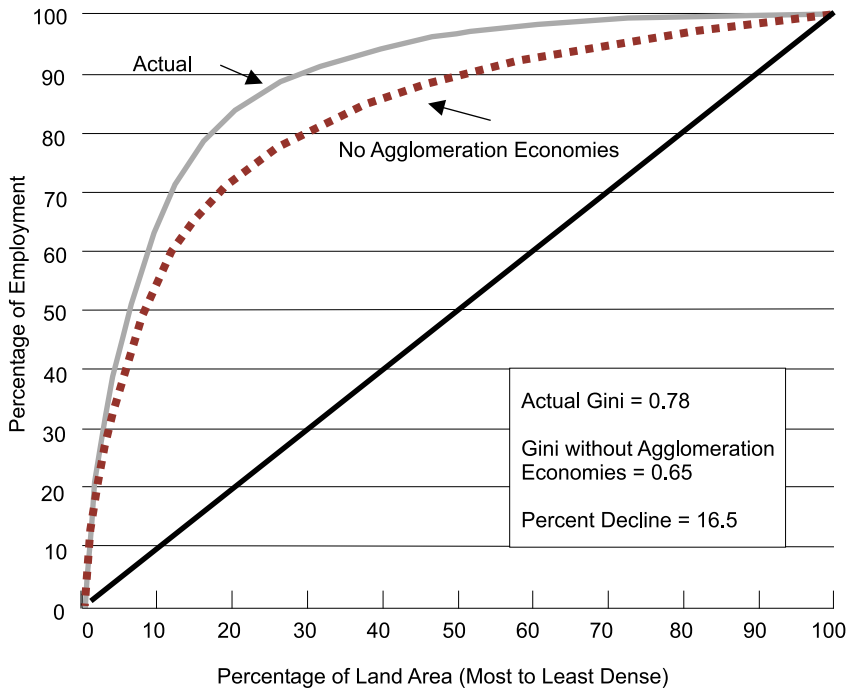
However, the mere existence of agglomeration economies does not settle the question of whether these effects are the primary cause of the spatial concentration of employment. To settle that point, we need to determine if agglomeration economies, as measured, are powerful enough to give rise to the degree of spatial concentration we see in the real world. This article highlighted research that seeks to make this determination. Contrary to expectations, I found that the bulk of the spatial concentration of employment results from factors other than agglomeration economies.

Just as spatial concentration can reduce the costs of producing goods and services, it may also reduce the costs of producing new knowledge through better utilization of knowledge workers.

The flip side of my finding is that some set of other factors accounts for the bulk of spatial concentration. Although my research cannot shed light on the contribution of these other factors, it's possible to hazard a guess (based on the work that other economists have done) as to what these other factors might be: natural resources, state and local economic policies, proximity to other metro areas, and spatial concentration's benefits in creating new knowledge. Whatever the case is, my research suggests that agglomeration economies are one of several important factors, but not the principal factor, affecting spatial concentration of employment.

FIGURE 3

Lorenz Curves With and Without Agglomeration Economies



[10] See the article by Gordon Hanson for evidence in favor of this point.

[11] The article by Adam Jaffe, Manuel Trajtenberg, and Rebecca Henderson and my article with Gerald Carlino present evidence that proximity may help in the communication of new knowledge.

REFERENCES

Carlino, Gerald A., Satyajit Chatterjee, and Robert Hunt. "Knowledge Spillovers and the New Economy of Cities," Working Paper 01-14, Federal Reserve Bank of Philadelphia, September 2001.

Chatterjee, Satyajit, and Gerald A. Carlino. "Aggregate Metropolitan Employment Growth and the Deconcentration of Metropolitan Employment," *Journal of Monetary Economics,* 48, 2001, pp. 549-83.

Hanson, Gordon H. "Market Potential, Increasing Returns, and Geographic Concentration," Graduate School of International Relations and Pacific Studies, University of California, San Diego, December 2001.

Holmes, Thomas. "The Effects of State Policies on the Location of Industry: Evidence from State Borders," *Journal of Political Economy,* 106, 1998, pp. 667-705.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108, 1993, pp. 577-98.

Rappaport, Jordan, and Jeffrey D. Sachs. "The U.S. as a Coastal Nation," Federal Reserve Bank of Kansas City, Working Paper 01-11, revised October 2002.

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

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