2
Localization

In 1895 a teenaged girl named Catherine Evans, living in the small Georgia city of Dalton, made a bedspread as a wedding gift. It was an unusual bedspread for the time, in that it was tufted; the craft of tufting or candlewicking had been common in the eighteenth and early nineteenth centuries but had fallen into disuse by that time. As a direct consequence of that wedding gift, Dalton emerged after World War II as the preeminent carpet manufacturing center of the United States. Six of the top twenty U.S. carpet manufacturing firms are located in Dalton; all but one of the rest are located nearby, and the carpet industry in Dalton and vicinity employs nineteen thousand workers.

I'll come back to Catherine Evans and her story later in this lecture. For now let me simply assert that aside from being particularly charming, the carpet story is actually a fairly typical one. To a remarkable extent, manufacturing industries within the United States are highly localized; and when one tries to understand the reasons for that localization, one finds that it can be traced back to some seemingly trivial historical accident.

I'll tell some stories later. First, let's develop an analytical structure.

Sources of Industry Localization

The observation of high localization of industries is, of course, not a new one. Indeed, it was such a striking feature of the process of industrialization that it attracted a great deal of attention in the later nineteenth century, with a fascinating monograph on the subject contained in the 1900 U.S. Census. The literature on industry localization is much too extensive to cite; notable examples over the years include Hoover 1948, Lichtenberg 1960, and very recently, Porter 1990. There is also considerable overlap between the subject of industry localization and urban economics; theory, anecdotal evidence, and solid empirical work can be found in such works as Bairoch 1988, Jacobs 1969 and 1984, and Henderson 1988.

Let us, however, go back to the source. It was Alfred Marshall who presented the classic economic analysis of the phenomenon. (Actually, it was the observation of industry localization that underlay Marshall's concept of external economies, which makes the modern neglect of the subject even more surprising.)

Marshall (1920) identified three distinct reasons for localization. First, by concentrating a number of firms in an industry in the same place, an industrial center allows a pooled market for workers with specialized skills; this pooled market benefits both workers and firms:

> [A] localized industry gains a great advantage from the fact that it offers a constant market for skill. Employers are apt to resort to any place where they are likely to find a good choice of workers with the special skill which they require; while men seeking employment naturally go to places where there are many employers who need such skill as theirs and where therefore it is likely to find a good market. The owner of an isolated factory, even if he has good access to a plentiful supply of general labor, is often put to great shifts for want of some special skilled labor; and a skilled workman, when thrown out of employment in it, has no easy refuge.

Second, an industrial center allows provision of nontraded inputs specific to an industry in greater variety and at lower cost:

> [S]ubsidiary trades grow up in the neighborhood, supplying it with implements and materials, organizing its traffic, and in many ways conducing to the economy of its material...the economic use of expensive machinery can sometimes be attained in a very high degree in a district in which there is a large aggregate production of the same kind, even though no individual capital employed in the trade be very large. For subsidiary industries devoting themselves each to one small branch of the process of production, and working it for a great many of their neighbors, are able to keep in constant use machinery of the most highly specialized character, and to make it pay its expenses...

Finally, because information flows locally more easily than over greater distances, an industrial center generates what we would now call technological spillovers:

> The mysteries of the trade become no mystery; but are as it were in the air.... Good work is rightly appreciated, inventions and improvements in machinery, in processes and the general organization of the business have their merits promptly discussed: if one man starts a new idea, it is taken up by others and combined with

suggestions of their own; and thus it becomes the source of further new ideas.

(On the whole, I like Marshall's turn of phrase better than the modern one!)

Cut through the archaism of Marshall's language and his lack of formalism, and you will see that he had a pretty sophisticated model in mind. He missed a few tricks that I will try to point out, but on the whole the main purpose of this part of the lecture is to rephrase Marshall in a drier, less felicitous style, and thereby bring it up to date.

Let us therefore consider in turn each of the Marshallian reasons for localization.

Labor Market Pooling

Imagine for a moment that there is some industry that consists of just two firms, each of which can produce in either of only two locations. These firms both use the same distinctive kind of skilled labor. For whatever reason, however, the firms' demands for labor are not perfectly correlated. For example, they may produce differentiated products that face uncertain demand; or they may be subject to firm-specific production shocks. Whatever the reason, the labor demand of the firms is both uncertain and imperfectly correlated.

To make matters more concrete, suppose that each firm may experience either "good times," in which it would like to hire 125 specialized workers at the going wage, or "bad times," in

which it would like to hire only 75. We also suppose that there are 200 of these workers in total so that average demand for labor equals supply. (In this example I take the wage rate for the specialized labor as given, so that there may be either excess demand or excess supply for labor. If you like, imagine a wage bargaining process that sets the wage at an expected market-clearing level before the shocks to labor demand are revealed—not too unrealistic an assumption. This assumption is not, however, essential. I will show in a moment that even if the wage rate is completely flexible and the labor market clears, the basic story remains the same.)

Now we may ask: will firms and workers be better off if the two firms choose different locations—each thus forming a company town with a local labor force of 100—or if the two firms choose the same location, with a pooled labor force of 200 that can work at either firm?

You may immediately ask about the possibility of exploitation: wouldn't each firm prefer to have a captive local labor force? I'll come back to that soon and show that it doesn't work the way you may think. For now, just put it aside and assume that the wage is set at an expected market-clearing level. Then it should immediately be apparent that it is in the interest of both firms and workers that everyone be in the same place.

First, consider the situation from firms' point of view. If each has its own town, with a labor force of 100, then it will be unable to take advantage of its good fortune when labor demand is high: during good times, there will be an unfillable

excess demand of 25 workers. If the two firms are in the same place, however, then at least occasionally one firm's good times will coincide with the other firm's bad times, and there will be additional workers available.

Next consider the situation from workers' point of view. If they live in a company town, then the firm's bad times are their bad times too: whenever the firm has low labor demand, 25 workers will be laid off. If the firms are in the same place, then at least sometimes one firm's bad times will be offset by the other firm's good times, and the average rate of unemployment will correspondingly be lower.

This is a pretty trivial example. Yet I think it is useful as a way to clear up some points that are often misunderstood. First, the example clarifies the nature of the gains from labor market pooling. Because of the word "pooling" some people may be tempted to assume that the incentive to create a pooled labor market is something like portfolio diversification, that is, that it has something to do with risk aversion on the part of workers. No doubt minimizing risk is also an issue, but I didn't mention it in the example. So even if workers were entirely risk-neutral, there would be an efficiency gain from creating a localized industry with a pooled labor market.

Second, if you think about it, the example shows that uncertainty alone won't generate localization. You need increasing returns as well The key point is that in order to make a pooled labor market advantageous, I needed to assume that each firm had to choose one location or the other, not both. If each firm

could produce in both locations, or for that matter if each firm could be split into two identical firms, one in each place, then the full "portfolio" of firms and workers could be replicated in each location, and the motivation for localization would be gone. But the most natural justification for the assumption that firms do not locate in both places is that there are sufficient economies of scale to militate for a single production site.

So it is the *interaction* of increasing returns and uncertainty that makes sense of Marshall's labor pooling argument for localization. But as I have described it so far, it is only an argument for the advantage of concentrated production, not a description of the process that might bring about such concentration. Can we provide such a description?

Consider figure 2.1 (which bears more than a coincidental resemblance to some of the figures from the last lecture). I envision an industry in which there are a fixed number of firms

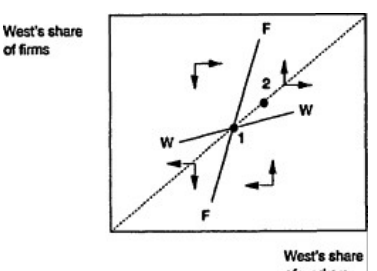

Figure 2.1

and a limited supply of specialized workers. Each firm and each worker must choose one of two locations, East or West. On the horizontal axis I show West's share of the labor force, on the vertical axis its share of the number of firms.

Which location will firms and workers prefer? The curves *FF* and *WW* show which distributions of firms and workers will leave the typical firm and worker respectively indifferent between the two locations. Both curves cross the 45-degree line at point 1, in the middle of the box: if everything is the same, everything is the same.

For any given labor force, firms would rather face less competition for the available workers. Thus an increase in the share of firms in West will make that location less attractive for the typical firm, unless this increase is offset by an increase in the labor force as well. So the set of combinations of share of firms and share of labor force that leaves firms indifferent between the locations, shown by *FF*, is upward sloping.

Workers would prefer to share the demand of any given number of firms with as few other workers as possible; an increase in the labor force in West will therefore make it a less attractive location for workers, unless offset by an increase in the number of Firms. So the set of combinations of share of firms and share of labor force that leaves workers indifferent between East and West, shown by *WW*, is also upward sloping.

As drawn, *FF* is steeper than *WW*. To see why, consider point 2. At that point, the ratio of firms to workers is the same in West

and East; but West has more of both. This means that West will offer better labor market pooling, and hence that at point 2 both firms and workers will prefer West to East. So point 2 is below *FF* (West is preferred by firms) and above *WW* (West is preferred by workers). This is only possible if at point 1 *FF* is steeper than *WW*.

Now ask what happens if we are off these schedules of equal attractiveness. Presumably firms move to the more attractive location, as do workers. The resulting dynamics are illustrated by the arrows in figure 2.1. There are three equilibria; but the one in the middle is knife-edge unstable. So we will converge to concentration of both firms and workers either in East or West, depending on initial conditions.[1]

This is a rough formalization of Marshall's labor pooling argument. It has many loose ends, of course; let me try to tie up a few of them.

Additional Thoughts on Labor Pooling

The biggest loose end in the story just told—a loose end that is flapping free in Marshall as well—is in the description of wage determination. In fact, this is a double issue (a split loose end?). First, how dependent is the argument on the assumption of sticky wages and a non-clearing labor market? Second, what

[1] Obviously there is an issue of self-fulfilling prophecy, of expectations as opposed to history, in models of localization as well as in the more macro model of the previous lecture. I will take it as understood that this issue is lurking in the background, and proceed with ad hoc dynamics that ignore the expectational problem.

about the issue of exploitation—the advantage to firms of a monopsony position in a company town?

A special example can help demonstrate that even with flexible wages it is advantageous to have a pooled labor force. Return to our two-firm example, but now suppose that each firm has a downward-sloping demand for labor rather than a simple number of workers that it wants. Let wages be flexible, and ignore the potential for monopsony. Then in the case where the firms locate in different places there will no longer be unemployment or excess demand for labor; instead, the local wage rate will fluctuate.

If the firms formed a pooled labor market, however, the wage rate would fluctuate less. Risk-averse workers would like this. But beyond this, firms would find that their profits were higher. To see this, take the extreme case in which the labor demands of the two firms are perfectly negatively correlated, so that the wage rate in the pooled case does not fluctuate at all. The average wage rate would be the same[2]; but this does not leave the firms indifferent.

Consider figure 2.2. It shows the "good times" and "bad times" labor demands of one of the two firms. If the firm is isolated, it pays a high good-times wage and a low bad-times wage, while always employing the same number of people; if it is part of a pooled market, it pays a constant wage that is the average of the two. But it now employs more workers in good

2. Actually, to leave the average wage rate unchanged we must assume both linearity of the labor demand schedules and additivity of the shocks; but the basic point goes through regardless, as shown in appendix C.

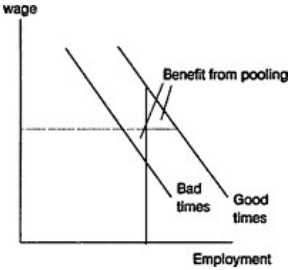

Figure 2.2

times, fewer in bad. As a result, the gain to the firm from the lower wage it pays in good times is greater than the loss it suffers from the higher wage it pays in bad times. In the figure, the net benefit is measured by the sum of the areas of the two indicated triangles.

This is an extreme example, but it makes a more general point. The gains from labor pooling do not rest in any essential way on a failure of labor markets to clear.

What about monopsony power? Won't firms prefer to have a captive labor force that they can exploit? Yes, other things equal. But other things won't be equal—and in fact the monopsony issue actually acts as a further reason for localization.[3]

[3] This point was first made by Julio Rotemberg and Garth Saloner (1990). The style of their model is very different from what I present here, but essentially I am just offering a variant of their analysis.

To see why, now let us eliminate the uncertainty motive for pooling, and suppose that firms have known labor demand schedules. They will ordinarily, however, be off these schedules, because they will restrict their hiring in an effort to keep wages down. The extent of this restriction will depend on the degree of competition among the firms, which will presumably depend on how many firms there are in each location.[4] So an equal proportional increase in the number of firms and workers will benefit workers, hurt firms. So in figure 2.3, which has the same format as figure 2.1, both *FF* and *WW* are shown as flatter than the 45-degree line.

But I show *FF* as steeper than *WW*. That is, an increase in West's share of the labor force, offset by a large enough increase in its share of firms to leave workers indifferent between East and West, will make West a more attractive location for firms.

To see why, compare points 1 and 2 in the figure. At point 1, which is the middle of the box, both wage rates and the profits of a typical firm are the same in East and West. At point 2, where West has a larger labor force offset by a larger number of firms, wage rates are still the same—so workers are still indifferent. However, firms are more profitable in West, even though they pay the same wage, because they engage in less restriction of hiring in order to keep the wage down. Because

4. Formulating the nature of competition in the labor market is tricky. Rotemberg and Saloner assume Bertrand competition, so that the wage goes from the reservation wage to the competitive level when the number of firms goes from one to two. This seems a little extreme; as appendix C points out, however, it is difficult to come up with an equally elegant alternative. For the purposes of this exercise I will simply assume that exploitation goes down as the number of firms rises.

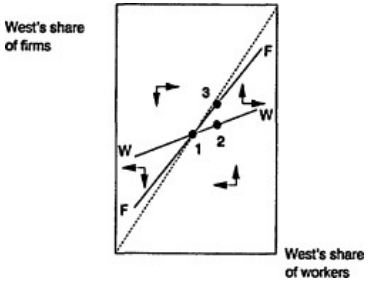

Figure 2.3

the number of firms has not risen in proportion to the number of workers, each firm has larger employment; since workers are paid less than their marginal product, this means higher profits. So in order to leave firms indifferent between West and East, it would be necessary to offset the increase in West's labor force by a larger increase in the number of firms, for example at point 3. Thus we know that *FF*, while flatter than the 45-degree line, is steeper than *WW*.

Once again the dynamics are illustrated by the arrows; and we see that the equilibrium in the middle is unstable, and the industry ends up concentrating either in East or in West. But in this example we have assumed away uncertainty, so that there are no pooling gains from concentrating the industry. Instead, the concentration of the industry is driven entirely by the issue of market power in the labor market. The intuitive argument that says that firms prefer company towns, because of the monopsony power this gives them, is exactly wrong.

Why does it turn out this way? One way to think about it is the following: there is a tug-of-war between firms, who prefer a less competitive labor market and hence production in both locations, and workers, who prefer a more competitive market and hence concentration in one location. Workers win this tug-of-war because a more competitive labor market is also more efficient; given noncollusive choice of location by firms, this efficiency gain decisively tips the scales toward the concentrated solution.

An alternative, and perhaps deeper, way to view the issue is in terms of credibility. Firms would like to convince workers that they will *not* try to exploit their monopsony power, so that they can attract workers to their production location.[5] But the only credible way to do this is to have enough firms in the location that there is an assurance of competition for workers. The commonsense idea that firms would like to have a company town in which workers could be exploited is right; but the point is that workers will shun such towns if they can, so that firms will end up finding it more profitable to locate in agglomerated centers that are not company towns.[6]

---

5. In Rotemberg and Saloner's formulation workers are not mobile, but immobile workers must choose whether or not to invest in industry-specific human capital. The basic principle is the same.

6. Of course company towns do exist, Impressionistically, they seem to happen for one of two reasons. First, there may be specific natural advantages that cause individual factories to be located on scattered sites, as was the case for water-driven New England textile towns. Second, the economies of scale in an industry may be so large that a single firm dominates the industry and agglomerates its plants in order to achieve pooling, like Eastman-Kodak in Rochester or Boeing in Seattle.

So it turns out that the loose ends that Marshall left in his labor market story do not invalidate it. Allowing for the possibility of flexible wage rates, while eroding some of the simplicity of the labor pooling story, does not make it go away; allowing for the possibility that firms may try to exploit monopsony power actually makes the story stronger.

Intermediate Inputs

Marshall's second reason for agglomeration, the availability of specialized inputs and services, seems straightforward enough. A localized industry can support more specialized local suppliers, which in turn makes that industry more efficient and reinforces the localization.

There are two points, however, that could perhaps use some clarification. First, the intermediate inputs story, like the labor pooling story, depends crucially on at least some degree of economies of scale. If there were no economies of scale in the production of intermediate inputs, then even a small-scale center of production could replicate a large one in miniature and still achieve the same level of efficiency. It is only the presence of increasing returns that makes a large center of production able to have more efficient and more diverse suppliers than a small one.

Second, the intermediate inputs story does not depend on some asymmetry in transportation costs between intermediate and final goods. Weber-type stories of transport cost minimization may suggest that localized industrial complexes

will emerge only if it is more costly to transport intermediate inputs than final goods. This is an impression that may have been reinforced by some models in trade, unfortunately including the work of Elhanan Helpman and myself (1985). In our models of international trade we contrasted the case where intermediate goods are tradable but final goods are not—which gave rise in effect to external economies at the level of the world rather than the individual country—with the reverse, which led to the formation of national industrial complexes. The reason for stressing these two extreme cases was convenience: it is easier to model a good that is either perfectly tradable or perfectly nontradable than one that can be traded, but only at a cost. Unfortunately, this approach can convey the impression that localization due to the clustering of suppliers occurs only in the special case in which transportation costs for intermediate goods are particularly high.

This is a misleading impression. In fact, localization will tend to occur unless the costs of transporting intermediates are particularly *low* compared with those of transporting final goods. And a general reduction in transport costs, of both intermediates and final goods, will ordinarily tend to encourage localization rather than discourage it.

To see why, it is useful to consider a tricky sort of model in which intermediates and final goods are the same thing. Imagine a group of products each one of which is demanded both as a final good and as an input into all of the others. For example, suppose that the typical product in the group has total sales of 10, but that 4 of these sales are to manufacturers of other products in the group. And correspondingly we must

suppose that to produce these 10 units requires 4 units of intermediate inputs, which again are drawn from the same industry. Notice that by making each good both a final and an intermediate, I have a fortiori imposed symmetry among intermediates and final goods in terms of their tradability.

Now suppose that there are two possible locations of production, each of which is also the location of half the final demand, that is, 3 units of each product. Where will a firm want to locate? The answer obviously depends on the decisions of other firms. If everyone else is in East, then 7 of the 10 units of total demand will be in East (3 final plus 4 intermediate); this will provide a firm with an incentive to locate its own production in East as well. The incentive will be reinforced by the fact that all of the firm's supplies of intermediate goods will come from East and will therefore be cheaper there. Thus there will be both backward and forward linkages that provide an incentive to concentrate production. Of course there will also be an incentive to move closer to final demand that will pull the other way.

This should be sounding familiar: it sounds an awful lot like the core-periphery model sketched out in the first lecture and derived formally in appendix A. And indeed it is possible to construct a model of intermediate goods and industry localization that is formally exactly analogous to the core-periphery model. This is helpful, because we already know something about that model. In particular, we know that the prospects for formation of a core-periphery pattern depend negatively on transportation costs, positively on the share of "foot-

loose" demand, and positively on the importance of economies of scale. The same should be true here, if these variables are suitably reinterpreted. In particular, the role played by the share of manufactures in the core-periphery model is here taken instead by the share of the industry's output that is used as an intermediate good rather than directly for final demand. Given this, we see that lower transport costs make industrial localization more likely, even if the cost of transporting intermediates falls along with that of transporting final goods.

It is also interesting to note that the same historical forces that in the last lecture I argued gave rise in the nineteenth century to the emergence of geographical concentration at a macro level—falling transportation costs, industrialization, and growing economies of scale—also should have led to increased localization of industry within the manufacturing belt. The only difference is that instead of an increased share of manufactures per se in demand, the necessary change is an increased share of manufactures that are used as inputs into other manufactures. But this did of course happen, as the 'industrial ecology" of the U.S. economy became more diverse and complex. So it is not surprising that localization of industry, the emergence of sharply distinct industrial characters for particular dries or districts, became a striking characteristic of the American economy by the end of the century.

Technological Spillovers

I have saved for last the reason for localization that many economists would put first—namely, the more or less pure externality that results from knowledge spillovers between

nearby firms. The emphasis on high technology in much policy discussion and the fame of such clusters as California's Silicon Valley and Boston's Route 128 have made technological externalities the most obvious thing to mention. Furthermore, economists with a conventional background still have a hankering to preserve perfect competition in their models; purely technological externalities do this.

Yet I have chosen to put pure technological external economies last, not first, for several reasons. First, it is an empirical fact that many of the industries that are highly localized within the United States now or were highly localized in the past are nothing like high technology sectors. Silicon Valley is famous; but equally remarkable concentrations maybe found of carpet producers around Dalton, Georgia; of jewelry producers around Providence, Rhode Island; of financial services in New York; and historically, of such industries as shoes in Massachusetts or rubber in Akron. Evidently forces for localization other than those involving high technology are quite strong.

Second, as a matter of principle I think we should try to focus first on the kinds of external economies that can be modeled other than by assumption. Labor pooling or intermediate goods supply are things that in principle one could examine directly and predict given a knowledge of the technology of the industry. And on the other side, the concreteness of these forces places constraints on what we can assume. Knowledge flows, by contrast, are invisible; they leave no paper trail by which they maybe measured and tracked, and there is nothing to prevent the theorist from assuming anything about them

that she likes. A sociologist might be able to help with survey methods; but I would like to get as far as possible with drab, down-to-earth economic analysis before turning to the other social sciences.

Finally, high technology is fashionable, and I think we are all obliged to make a deliberate effort to light against fashionable ideas. It is all too easy to fall into a kind of facile "megatrends" style of thought in which the wonders of the new are cited and easy assumptions are made that everything is different now. Of course the world has changed—but it was a pretty remarkable place even before the coming of large-scale integrated circuits, and even high technology industries respond to old-fashioned economic forces.

So while I am sure that hue technological spillovers play an important role in the localization of some industries, one should not assume that this is the typical reason—even in the high technology industries themselves.

Some Empirical Evidence

I have alluded to some facts about the degree of localization of industry within the United States. But what sort of facts do we actually have? One kind of evidence, which I regard as very important in spite of its lack of rigor, consists of case studies. Nothing is better at suggesting the kind of model that we ought to use than a collection of examples, particularly with some historical depth, of how particular industries come to be

in particular places. And the stories are often entertaining too. Before turning to stories, however, I want to discuss some very preliminary statistical work that I have undertaken.

The objective of this work was to answer two questions. First, how localized is the "typical" U.S. industry? Are familiar examples like that of the auto concentration around Detroit or high technology in Silicon Valley normal, or are they outliers?

Second, what sorts of industry are highly localized? Are localized industries typically high technology sectors (which would lend support to a technological spillover model), or are they more prosaic? Are they industries that use highly skilled labor or more general-purpose workers?

The technique I have tried to get at these questions is to construct "locational Gini coefficients" for as many U.S. manufacturing industries as possible. A locational Gini curve for an industry is constructed as follows. First, for each of the locational units in our sample we calculate both the share of total national manufacturing employment and the share of national employment in the industry. Then we rank the units by the ratio of these two numbers. Finally, we run down the ranking, keeping a cumulative total of both the sum of total employment share and the sum of employment share in the industry.

Suppose, for example, that there were three regions. Region 1 has 20 percent of total manufacturing employment, but 50 percent of employment in the widget industry. Region 2 has 40 percent of total manufacturing employment, and 40 percent of

widget employment. Region 3 has the remaining 40 percent of manufacturing employment, and the remaining 10 percent of widget employment. Then the resulting curve would be the one illustrated in figure 2.4. Clearly, the more the geographical distribution of the widget industry matches that of overall manufacturing, the closer this curve will lie to the 4S-degree line. So there is an obvious index of localization: the area between the curve and the 45-degree line. An industry that was not localized at all, but simply spread out in proportion to overall employment, would have an index of 0; one that is concentrated almost entirely in a region with small overall employment would have an index dose to 0.5.

In implementing these indices, I (or rather my research assistant) used data on U.S. three-digit industries, classified by states. This is only a preliminary step; although it is illuminating, there are a number of problems that blur the results.

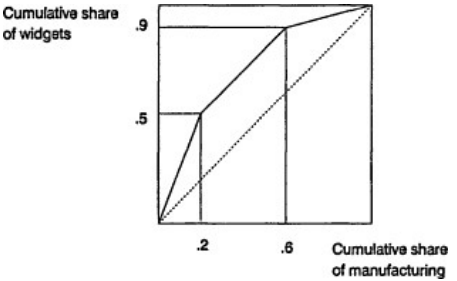

Figure 2.4

First, there are some important holes in the data. The U.S. Bureau of the Census does not reveal information that would compromise the assurance of confidentiality; perhaps surprisingly, this affects data even at the level of large states and big industries. For example, data is withheld on the aircraft industry in Washington (so as not to reveal too much about Boeing) and about the photographic equipment industry in New York (so as to protect Kodak). These are major examples of industry localization, indeed classic ones. Unfortunately, they must simply be dropped from the statistical analysis.

Second, the definition of an industry is problematic. Three-digit industrial classifications are not useless, but they are far from ideal. Old industries, like costume jewelry, rate a three-digit category in spite of very modest employment, while both Silicon Valley and Route 128 are buried inside the huge category of "electronic components." Ideally we would disaggregate and then reaggregate in order to get more meaningful comparisons—but I haven't.

Finally, states aren't really the right geographical units. First of all, they are of very unequal population, which biases industry comparisons: an industry that is completely concentrated around a small town in California will not generate as high a locational Gini as a comparably concentrated industry in Rhode Island. Second, economic regions do not respect state boundaries. The textile complex that shows up very strongly in the data occupies the Piedmont area of the Southeast; state comparisons do not really reveal the compactness (in both the lay and the mathematical senses) of this industry's location.

In spite of all these limitations, we can learn a fair amount from even this preliminary calculation. The results for 106 three-digit industries are reported in appendix D, which also shows total employment in the industry and employment in the three leading states.

The first impression that emerges from the results is that many industries are indeed highly concentrated geographically. The automotive industry offers a useful benchmark. It is a fatuously localized industry. Although there has been some dispersal since the heyday of Motown, half of the employment is still in the traditional automotive district of southern Michigan and neighboring regions of Indiana and Ohio. So we might expect motor vehicles to be an exceptional industry. But it isn't. It is just slightly above the median. The point is not that automotive production is not highly localized—it is. But so are a lot of other industries.

The other feature of the data that stands out is that the most highly concentrated industries by this admittedly crude calculation are not cutting-edge, high technology sectors. Indeed, the thing that leaps out from the table is not the localization of high technology industries but the duster of textile-related industries, all of them in more or less the same place: the Piedmont area of the Carolinas and Georgia. Half of the top twenty industries according to my Gini ranking are Piedmont textile sectors. (As we will see in a moment, there is considerable localization of particular industries in this group within the Piedmont; but leave that aside for a moment.)

There is probably a bias in the way this table was constructed against finding the localization of high-tech industries. For one thing, two high-tech industries that are famously geographically localized had to be excluded from the table because of withheld data: aircraft, dominated by Boeing's huge Seattle facilities, and photographic equipment, with Kodak's Rochester complex. These are, of course, industries that are concentrated in ownership as well as in geography.

More important, the fact that the classification scheme is so antiquated means that quite small traditional industries still rate their own three-digit codes, while advanced sectors are buried in meaningless aggregates. Silicon Valley and Route 128 are real enough, but you just can't find them in the statistics.

So I don't want to be misinterpreted. This evidence does not show that high technology industries are not localized. What it shows is simply that low technology industries are also localized. Whatever drives industries to concentrate in one place, it is not solely a matter of technological spillovers.

This is about as far as I want to take this statistical exercise in this lecture; I'll come back to some U.S. regional data in the next lecture. But now I want to get to the fun part and tell stories.

Case Studies: Some Historical Examples

One of the Piedmont textile industries that I found to be highly concentrated geographically was the carpet industry. My

research assistant did some follow-up work on this industry and uncovered a classic case of the role of historical accident and cumulative processes in generating localization. So let's go back to Catherine Evans and her bedspread.

As I mentioned at the beginning of this lecture, in 1895 the teenaged Miss Evans made a bedspread as a gift. The recipients and their neighbors were delighted with the gift, and over the next few years Miss Evans made a number of tufted items, discovering in 1900 a trick of locking the tufts into the backing. She now began to sell the bedspreads, and she and her friends and neighbors launched a local handicraft industry that began selling items well beyond the immediate vicinity.

The handicraft industry became semimechanized in the 1920s, as tufting was used to satisfy the surging demand for chenille sweaters; but production continued to be done by individual households.

Immediately after World War II, however, a machine was developed for producing tufted carpets. Until that time, ma-chine-made carpets had been woven. Tufting proved to be far cheaper. And guess where one could find people who knew about tufting and were quick to see the potential? In the late 1940s and early 1950s, many small carpet firms sprang up in and around Dalton, together with a duster of supporting firms providing backings, dyeing, etc. Existing carpet manufacturers initially dung to weaving; they eventually either went out of business, driven out by the upstarts from Dalton, or moved their own operations from traditional sites in the Northeast to

Dalton. And so the little Georgia city emerged as America's carpet capital.

It's a lovely story; it's also a very typical one. The whole process of industrialization within the United States was marked by similar stories of small accidents leading to the establishment of one or two persistent centers of production. Anyone who thinks that Silicon Valley is a distinctly modern sort of creation should look at the fascinating monograph, "The Localization of Industries," contained in the 1900 U.S. Census. The monograph identifies fifteen highly localized industries in 1900, including: collars and cuffs, localized in Troy, New York; leather gloves, localized in the two neighboring New York towns of Gloversville (sic) and Johnstown; shoes, in several cities in the northeastern part of Massachusetts; silk goods, in Paterson, New Jersey; jewelry, in and around Providence, Rhode Island; and agricultural machinery, in Chicago.

In each of these cases there is a story similar to, if not quite as charming, as that of Catherine Evans. An accident led to the establishment of the industry in a particular location, and thereafter cumulative processes took over. The Massachusetts shoe industry owed its start to the Welsh cobbler John Adams Dagyr, who set up shop in 1750; the dominance of the Providence jewelry industry (which still makes it onto our list of highly localized industries!) began when a local man invented "filled" gold in 1794; the reign of Troy as the detachable collar and cuff center (alas, fashion!) was inaugurated by a Methodist minister in the 1820s.

What is important to the economist here is, of course, not the initial accident but the nature of the cumulative process that allowed such accidents to have such large and long-lasting effects. What the historical record shows us are two things. First, such cumulative processes are pervasive; Silicon Valley is not at all unique, either in time or space, but is simply a glitzy version of a traditional phenomenon. And second, Marshall's first two reasons for localization, labor pooling and the supply of specialized inputs, play a large role even when pure technological externalities seem unlikely to be important.

Has the basis of localization shifted over time? The authors of the monograph in the twelfth census thought that it would. They noted that many of the highly localized industries were ones that relied heavily on skilled handwork, and speculated that the arbitrary nature of localization would tend to be eroded because "The use of machinery has ... tended to lessen the importance of a specially skilled labor supply. In proportion as an industry becomes automatic, its localization becomes independent of its supply of special labor."

To some extent they were right. First, it is doubtless true that each manufacturing industry, as it matures, tends to become less dependent on the pooled labor market, specialized inputs, and information spillovers that sustain localization. Consider, for example, the U.S. tire industry. Before 1930 this industry was spectacularly localized in the city of Akron, to which financial incentives from the chamber of commerce (industrial policy!) had lured one Benjamin Franklin Goodrich. As Detroit boomed, the rubber center in Akron came to have more than

a hundred firms, generating some of the highest wages in the United States and attracting migrants from all over (including my grandfather). Time was not, however, on Akron's side. As tire production became standardized, it could be delocalized and moved closer to the market. The market itself spread out over time as the auto industry began to establish assembly plants around the United States.

As it turned out, the end was rather sudden. The Depression dealt a devastating blow to Akron's economy and apparently destroyed the critical mass of rubber firms there: when the country emerged, the role of Akron as a rubber center was gone. No major producers of tires are now located in the onetime tire capital of the world.

But while the Akron story shows that localization within an industry tends to fade away, there are always new industries. Detroit fades, but Silicon Valley rises. Indeed, surely there is a kind of product cycle, in which emergent new industries initially flourish in localized industrial districts, then disperse as they mature.

High Technology Clusters

In the last generation, the familiar examples of localization have changed. One rarely now hears about Motown, or Iron City, or the garment district (although they still exist in somewhat attenuated form). Instead it is all high tech: Silicon Valley, Route 128, Research Triangle. How do these new dusters compare with the older localization?

The first thing to say is that on the whole the stories of their founding are less romantic. In general the new high technology dusters were the product less of intrepid individuals than of visionary bureaucrats (if that is not an oxymoron). But otherwise the stories look rather similar.

Silicon Valley was created largely through the initiative of Fred Terman, the vice-president of Stanford University. Through his initiative the university provided an initial stake for Hewlett-Packard, which became the nucleus of the Valley. It also established the famous research park on university land, on which first Hewlett-Packard, then many other firms, began operations. There was a noticeable cumulative process operating through the university itself: the revenues from the research park helped to finance Stanford's ascent to world-class status in science and engineering, and the university's rise helped make Silicon Valley an attractive place for high-tech business.

Route 128 was created, in a more diffuse way, through the initiative of MIT's president, Karl Compton, who encouraged MIT faculty to become entrepreneurs and helped mobilize private venture capital

North Carolina's Research Triangle, finally, was created through state support of a research park, in direct emulation of Silicon Valley and Route 128.

Perhaps the most important thing to emphasize in these high technology stories is the importance of non-high-technology factors in the agglomerative process. Both in Silicon Valley

and around Route 128 a key advantage is the existence of a pool of people with certain skills. In the Boston area, for example, growth companies in the software field can be reasonably sure of being able to find people with esoteric knowledge in a variety of sub-subdisciplines. At the same time, the Boston area has been a good place for people to invest in acquiring these skills, or for those with those skills to live: if a start-up goes bust, as many do, you can find another job without having to relocate. This is just the labor pooling story; the fact that the skill involves high technology, rather than shoemaking or tufting, may be of secondary importance.

An anecdote: when involved with the MIT productivity commission, which eventually produced the best-selling book *Made in America*, I became very unpopular with the engineers for suggesting that some very non-technological-seeming sectors were in economic terms not so different from high tech. For example, consider Milan's fashion industry: a cluster of firms that rely on a highly specialized labor force (designers, models, stitchers, etc.), on specialized suppliers (fabrics, dyes, makeup, etc.), and on early access to information (what's in, what's out). I argued that economically Milan and Route 128 are similar creatures. The engineers thought I was frivolous.

Services

In the late twentieth century the great bulk of our labor force makes services rather than goods. Many of these services are nontradable and simply follow the geographical distribution of the goods-producing population—fast-food outlets, day-

care providers, divorce lawyers surely have locational Ginis pretty close to zero. Some services, however, especially in the financial sector, can be traded. Hartford is an insurance city; Chicago the center of futures trading; Los Angeles the entertainment capital; and so on.

The most spectacular examples of localization in today's world are, in fact, based on services rather than manufacturing. Tokyo and London are not essentially manufacturers. Even Silicon Valley and Route 128 are in effect more nearly centers that supply services to manufacturing than actual physical production sites. And arguably technology is moving in a direction that will promote more localization of services. Transportation of goods has not gotten much cheaper in the past eighty years: the epochal innovations were railroads and steamboats, with everything since representing only modest improvements. But the ability to transmit *information* has grown spectacularly, with telecommunication, computers, fiber optics, etc.

Some trends are visible. One of my students has been looking at the data on the growing concentration of wealth in Southeastern England; he finds that it is the service industries that are concentrating there, while manufacturing is actually shifting the other way.

The important point is that the logic of localization remains similar. Catherine Evanses—that is, small accidental events—start a cumulative process in which the presence of a large number of firms and workers acts as an incentive for still more

firms and workers to congregate at a particular location. The resulting pattern may be determined by underlying resources and technology at some very aggregative level; but at ground level there is a striking role for history and accident.