Volume Title: Globalization in Historical Perspective

Volume Author/Editor: Michael D. Bordo, Alan M. Taylor and Jeffrey G. Williamson, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-06598-7

Volume URL: http://www.nber.org/books/bord03-1

Conference Date: May 3-6, 2001

Publication Date: January 2003

Title: Globalization in History.A Geographical Perspective

Author: Nicholas Crafts, Anthony Venables

URL: http://www.nber.org/chapters/c9592

7

Globalization in History
A Geographical Perspective

Nicholas Crafts and Anthony J. Venables

7.1 Introduction

Globalization is about the changing costs of economic interactions across distance and the effects of these changes on the geographical distribution of economic activity. Technical change has been driving the costs of interactions steadily downward for many centuries, although policy interventions have sometimes raised them. Changes in the economic geography of the world economy have been more complex. There have been periods when activity has become more unevenly distributed across space, and periods when these spatial differences have narrowed as activity has spread from established centers into other regions and countries.

The mechanisms driving these changes were, among other things, easier movement of people, capital, and goods—"globalization." But why did the location of economic activity evolve in the way it has? Why did the world not develop some quite different economic geography, with different centers of production, or with activity more evenly distributed? Many factors are important, but in this chapter we highlight the role of geography. This includes the "first-nature" geography of oceans, rivers, mountains, and endowments, although our focus will be mainly on the "second-nature" geography of the spatial interaction between economic agents. The essence of globalization is that it changes these spatial interactions.

Most traditional analyses are based on economic models in which there are diminishing returns to most activities. Thus, migration tends to reduce

Nicholas Crafts is professor of economic history at the London School of Economics. Anthony J. Venables is professor of international economics at the London School of Economics.

The authors thank the conference participants, particularly Richard Baldwin and Jeff Williamson.

the wage in the host country, and an increase in manufacturing output encounters increasing costs. We argue in this paper that it is not possible to interpret several of the most important aspects of economic development in such a framework. An alternative is provided by models of "new trade theory" and "new economic geography" in which market imperfections at the micro level can give rise to increasing returns at a more aggregate level. The balance between increasing and decreasing returns in these models depends crucially on spatial interactions (determining, for example, the extent of the market) and changes in these interactions can have major effects. Globalization can trigger cumulative causation processes that cause uneven development to occur at a variety of different spatial levels—urban, regional, and international.

Our objective in this paper is to apply this new approach to several aspects of the historical experience of globalization. We proceed in three stages. First we sketch out some of the facts about the changing location of activity and the way that spatial interactions between economic agents changed over time. There were dramatic falls in the costs of moving goods, people, and information, occurring particularly from the 1870s onward. The falling costs were associated with large increases in trade relative to income, narrowing of international price gaps, and increases in migration flows. Second, we outline theoretical approaches to thinking about the consequences of these changes. One approach is the neoclassical model of production and trade, in which production is determined by factor endowments, technological differences, and the freeness of trade. We contrast this with a new economic geography approach, in which locations derive some of their comparative advantage from scale, and ability to exploit scale is in turn limited by the extent of the market. In this approach firms seeking profitable locations will be drawn to locations with good market access and proximity to clusters of related activities, as well as locations with appropriate factor endowments. We show that this alternative view provides a broad-brush picture that, in many respects, seems consistent with the historical record.

We then turn to look in more detail at several historical episodes. From the nineteenth century we focus on the rise of New World economies and the development of urbanization. We confront the central issue of early twentieth-century economic history, namely how the United States came to overtake other regions, and argue that insights from new economic geography can shed important light on this change. From the late twentieth century we revisit the East Asian "miracle," the most spectacular shift of the center of gravity in the world economy since the rise of the United States.

In pursuing the theme that geography matters for economic development we are consciously swimming against the tide of recent work both in economic history and in growth economics. Economic historians, notably in the new institutional economic history (North 1990), have stressed the im-

pact of incentive structures on investment and innovation and have argued that divergence stems from the path dependency of institutional arrangements. Endogenous growth models also tend to underline the centrality of microeconomic foundations for growth outcomes (Aghion and Howitt 1998), whereas neoclassical growth economists still believe in ultimate (twenty-first-century) convergence, following a post–Industrial Revolution interlude of divergence due to lags in the diffusion of best practice institutions, policies, and technology (Lucas 2000). Our position is that these conventional wisdoms are significantly modified by taking into account the way that changing costs of distance interact with economies of scale to shape the economic geography of the world.

A stylized version of this alternative perspective can be outlined as follows. If trade costs are very high then economic activity must be dispersed, whereas if trade costs are very low then firms will not care whether they are close to markets and suppliers. At intermediate levels of trade costs, however, the likelihood of agglomeration is high. Agglomeration forces operating through linkages across a wide range of activities will cause the world to divide into an industrialized rich center and deindustrialized poor periphery even if there are no differences in institutional quality or economic policy. Over time a number of mechanisms, including falling trade costs and growing world demand for manufactures, will make a new location outside the center become competitive, so industry moves there and it now benefits from agglomeration effects. Following the initial agglomeration phase, development therefore takes the form of enlargement of the set of countries in the center. This is not a process of steady convergence of poor countries to rich ones but rather the rapid transition of selected countries (close to or with good transport links to the center) from the poor to the rich club.

7.2 Location and Trade Costs: The Historical Record

In 1750 more than 50 percent of the world's industrial output was produced in China and India, compared to some 18 percent in Western Europe. The following eighty years saw the Industrial Revolution, with western Europe's industrial output more than doubling and that of the United Kingdom increasing by a factor of 7. Over the same period, industrial production in China and India continued to increase (by around 20 percent). It is not our purpose to analyze the origins of the Industrial Revolution but instead to study the changing economic geography of the world from this point on. The technological changes that resulted from the industrial revolution, notably in the form of the harnessing of steam power, not only raised European industrial output but also facilitated large reductions in both inland and ocean transport costs associated with the coming of the railroad and the steamship.

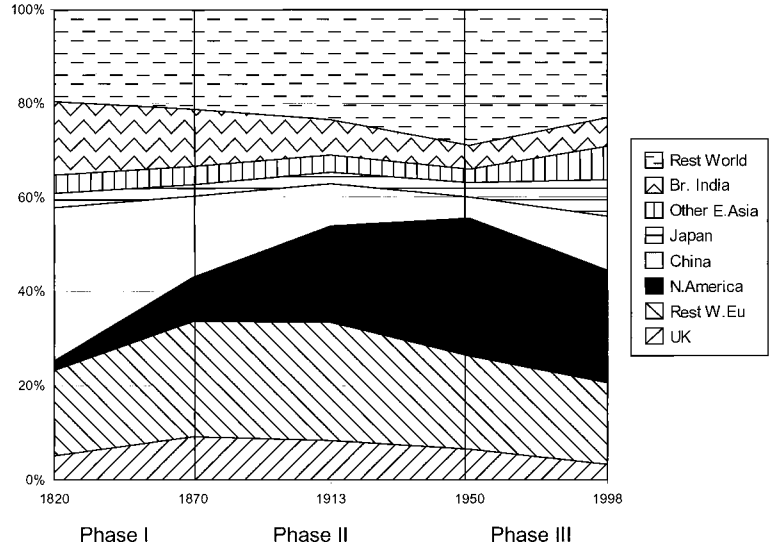

Fig. 7.1 Regions' share of world GDP
Source: Maddison (2001).

7.2.1 Location of Production: The Three Phases

Figure 7.1 shows the shares of world gross domestic product (GDP) attributable to major regions of the world economy at selected dates from 1820 onward, and figure 7.2 gives shares of industrial production for the same regions from 1750 on. Three main phases are apparent in both figures, although they are more pronounced for industrial production than for GDP as a whole. The first phase is the rise of the United Kingdom and western Europe as a whole and the dramatic collapse of China and India from these start dates through to the latter part of the nineteenth century. This period saw not only a decline of industrial production in China and India relative to the rest of the world but also an absolute fall such that 1830s levels were not regained until the 1930s (Bairoch 1982). The second phase is the rise of North America. Its share of world GDP and industrial output increased most rapidly from the American Civil War to the start of the Great Depression, peaking shortly after World War II. The third phase is revealed in the data for 1998 but has its origins in the postwar "golden age" of growth, namely, the large and rapid increase in the shares of Japan, China, and other East Asian countries in world GDP and industrial output.[1]

These phases correspond first to a concentration of activity in the United Kingdom and northwestern Europe (phase I), and then to two different

1. A complementary perspective on geographic aspects of catch-up and convergence is set out in Dowrick and DeLong, chapter 4 in this volume.

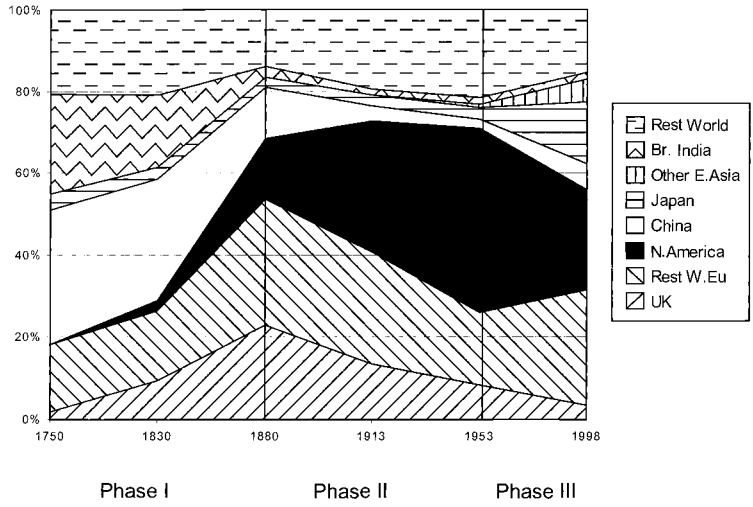

Fig. 7.2 Regions' share in world industrial production
Sources: Bairoch (1982); UN (1965); UNIDO (2001).

phases of dispersion, first to North America (phase II), and then to parts of Asia (phase III). Figure 7.3, which reports shares of world population, underlines the tendencies toward concentration, especially in industrial production, which became apparent during and after the nineteenth century. Whereas in the 1820s China and India accounted for a little over half the world's population and a little under half of world GDP and industrial production, by 1913 western Europe and North America, with about one-fifth of the world's population, produced over half of world GDP and nearly three-quarters of world industrial output. By 1998, with a rather smaller share of world population, these countries still accounted for well over half of world industrial output, whereas China and India, with over 40 percent of world population, produced only about 8 percent of industrial output.

Figure 7.4 reports manufacturing exports (from 1876–80 onward). Here there is evidence of even more concentrated activity. In the late nineteenth century the United Kingdom looms very large with over a third of all exports, even though only representing about 2.5 percent of world population. It was then superseded as the world's leading exporter by the rise of North America, which accounted for over a quarter of manufactured exports in 1955 with only about 6 percent of world population. (Europe looks large in the figure relative to the United States, essentially because intra-European trade is reported, in contrast to intra-U.S. trade). The remarkable feature of the last decades of the twentieth century was the rise of Chinese, Japanese, and other East Asian manufactured exports, representing a real breakthrough for newly industrializing countries.

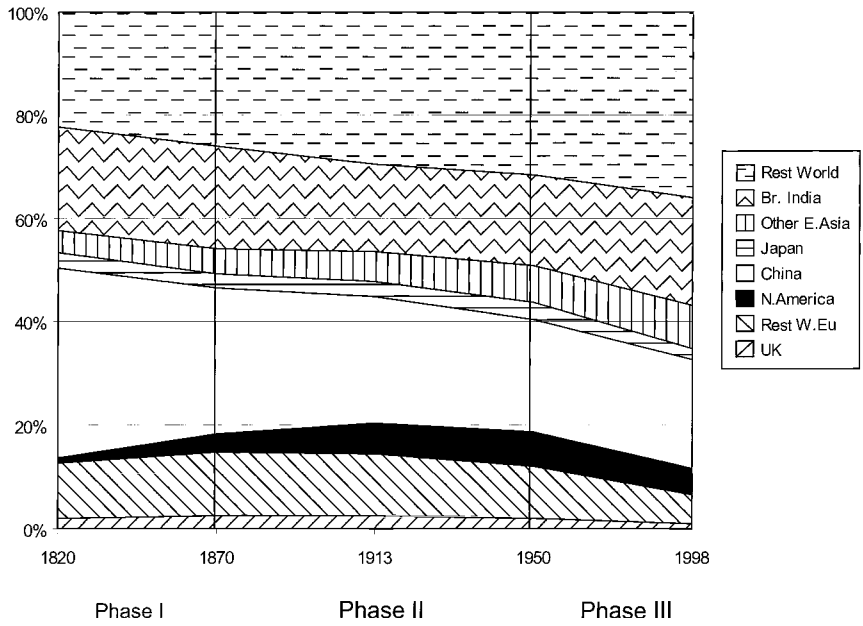

Fig. 7.3 Regions' share in world population
Source: Maddison (2001).

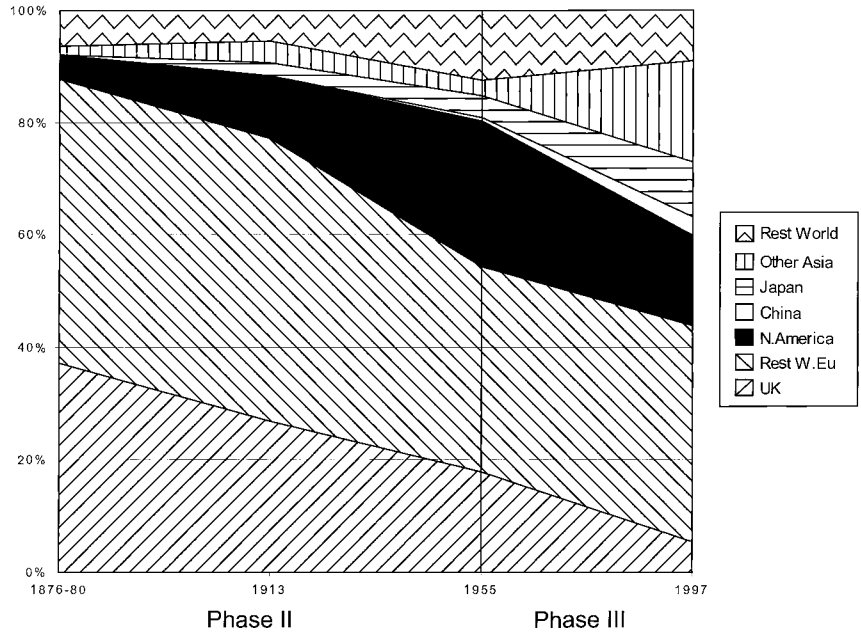

Fig. 7.4 Regions' share in world manufactured exports
Sources: UNCTAD (1983, 2000); Yates (1959).

Table 7.1 **Real Costs of Ocean Shipping (1910 = 100)**

Year	Cost
1750	298
1790	376
1830	287
1870	196
1910	100
1930	107
1960	47
1990	51

Sources: Derived using Dollar (2001), Harley (1988), and Isserlis (1938).

7.2.2 The History of Transport Costs

Although distance remains a barrier even at the start of the twenty-first century, the continuing communications revolution has been one of the most outstanding features of the last 200 years. Table 7.1 reports on the cost of ocean shipping for selected years since 1750. The period between 1830 and 1910 emerges as the era of very substantial decreases, and by the late twentieth century ocean shipping rates in real terms were about one-sixth of the level of the early nineteenth century.[2]

Ocean shipping is only a small part of the story, however, especially for the nineteenth century. This was also a period of spectacular declines in inland transport costs, which between 1800 and 1910 fell by over 90 percent (Bairoch 1990, 142). After World War II, however, new modes of transport became important, and by 1980 the real costs of airfreight had fallen to about a quarter of its level on the eve of World War II (Dollar 2001).

Trends in barriers to trade created by policymakers also need to be taken into account. Here the broad trends are well known even though details are sometimes elusive. The estimate of the unweighted world average tariff rate given by Clemens and Williamson (2001) and illustrated in figure 7.5 rises from about 12 percent in 1865 to 17 percent in 1910. In the interwar period, at a time when transport costs had ceased falling, trade wars pushed the Clemens-Williamson tariff rate up to 25 percent at its 1930s peak, and, in addition, quantitative trade restrictions proliferated, affecting perhaps 50 percent of world trade (Gordon 1941). After World War II, the Clemens-Williamson tariff rate is in the 12–15 percent range, where it remains until the 1970s, after which it falls to a low of 7–8 percent in the late 1990s. The quantitative restrictions of the 1930s and 1940s among the Organization for Economic Cooperation and Development (OECD) countries were largely removed in the postwar liberalization phase, and despite a revival in the era of voluntary export restraints in the 1970s and 1980s, post–Uruguay Round

2. A much more detailed account of this phenomenon can be found in Findlay and O'Rourke, chapter 1 in this volume.

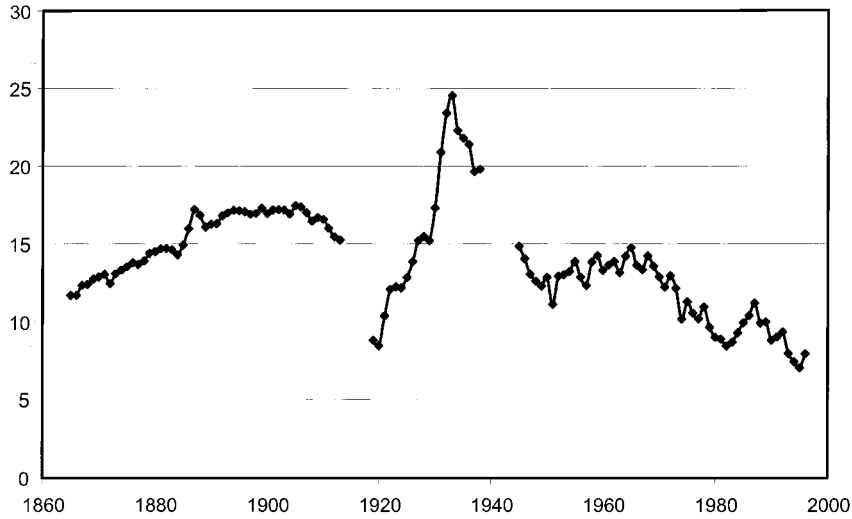

Fig. 7.5 Unweighted world average own tariff, thirty-five countries (percent)
Source: Clemens and Williamson (2001).

Table 7.2 Ratio of World Merchandise Exports to World GDP (%)

Year	Ratio
1820	1.0
1870	4.6
1913	7.9
1929	9.0
1950	5.5
1973	10.5
1998	17.2

Source: Maddison (2001).

these are probably as low as at any time since World War I (Daly and Kuwahara 1998).

Concurrent with these changes in trade costs and tariffs have been changes in ratio of foreign trade to world GDP, reported in table 7.2. In the early nineteenth century trade costs are so high and trade volumes so low (around 1 percent of GDP) that, of necessity, most production is located close to local markets. This constraint becomes relaxed through the nineteenth century, permitting the agglomeration of manufacturing to occur. After the reverses of the interwar period, the growth of trade relative to income resumes, again allowing new economic geographies to develop.

Nevertheless, at the beginning of the twenty-first century distance is still a powerful barrier to economic interaction. Gravity modelling finds that,

Table 7.3 **Economic Interactions and Distance (flows relative their magnitude at 1,000 km)**

	Trade	Equity Flows	FDI	Technology
1,000 km	1.00	1.00	1.00	1.00
2,000 km	0.42	0.55	0.75	0.65
4,000 km	0.18	0.31	0.56	0.28
8,000 km	0.07	0.17	0.42	0.05

Sources: See text.

controlling for the economic mass of the countries concerned, trade between them falls off steeply with distance. The elasticity of trade flows with respect to distance is typically estimated to be between –0.9 and –1.5, and the implications of this for trade volumes are given in the first column of table 7.3, which expresses trade volumes at different distances relative to their value at 1,000 km. With an elasticity of –1.25, trade volumes at 4,000 km are down by 82 percent, and by 8,000 km they are down by 93 percent. Similar methodologies have been used to study other sorts of economic interactions, and some results are summarized in remaining columns of table 7.3. Portes and Rey (1999) study cross-border equity transactions (using data for fourteen countries accounting for around 87 percent of global equity market capitalization, 1989–96), and their baseline specification gives an elasticity of transactions with respect to distance of –0.85, so that flows at 8,000 km are less than one-fifth those at 1,000 km. Foreign direct investment (FDI) flows are studied by Di Mauro (2000), who finds an elasticity with respect to distance of –0.42. The effect of distance on technology flows has been studied by Keller (2000) who looks at the dependence of total factor productivity on research and development (R&D) stocks for twelve industries in the Group of Seven (G7) countries, 1971–95. The R&D stocks include both the own-country stock and foreign country stocks weighted by distance.[3] Both own- and foreign-country stocks are significant determinants of each country's productivity, and so too is the distance effect, with R&D stocks in distant economies having much weaker effects on productivity than do R&D stocks in closer economies, so that the effect at 8,000 km is only 5 percent of its effect at 1,000 km.

7.3 Location and Trade Costs: Theory

How have the changing costs of spatial interactions shaped the geography of world economic activity? In this section we show how theory suggests that declining costs can explain the observed phases of concentration and of dispersion.

3. Distance weighting according to $\exp(-\theta\ \text{distance}_{ij})$.

7.3.1 The Location of Activity

Two sorts of considerations determine the structure of production and level of income of a country or region. One is its internal capacity, its endowment of stocks of factors of production, skills, knowledge, and social infrastructure. The other is its relationship with other countries or regions—its geography, meaning the access that it has to world markets and to external supplies of goods, factors, and knowledge.

Traditional trade theory's analysis of location focuses heavily on the endowments of primary factors of production. Special cases of the approach are the Heckscher-Ohlin trade model (with equal numbers of goods and factors) and the specific factors or Ricardo-Viner model, with more factors than goods. Both models show how, given world prices, the production structure and income of each country are determined, with countries tending to export goods intensive in their abundant factors. What are the predictions of these models about the effects of globalization? The first is that goods trade liberalization allows countries to exploit their comparative advantage more fully that we expect to see land-abundant countries becoming increasingly specialized in agricultural products, and so on. The second prediction derives from the fact that factor mobility and goods trade are, in general, substitutes. This means that goods trade liberalization reduces factor price differences between countries and thereby reduces the incentives for migration and capital movements.[4] Conversely, factor mobility will in general reduce trade flows, as factors flow to countries where they are relatively scarce, and thereby reduce the cross-country endowment differences that are the basis of trade.

The traditional approach is based on constant returns to scale in production, whereas new trade theory and new economic geography are based on increasing returns within the firm, and possibly in the economy more widely. The analysis focuses on the location decisions of firms and workers. Drawing on developments made in trade theory in the 1970s and 1980s, manufacturing production is modeled as distinct increasing returns to scale firms operating in imperfectly competitive markets (usually monopolistically competitive). There is intraindustry trade, as firms—subject to transport costs and trade barriers—sell their products into each market. What determines whether a country is profitable place for a firm to locate? As in traditional theory, factor prices and factor supplies matter. So too does geography, because firms seek to locate close to large markets and to good sources of intermediate input supply. The fact that locations with good market access are particularly attractive means that these locations will typi-

4. See Markusen (1983) and Venables (1999) for discussion of the issue of whether trade liberalization and factor mobility are substitutes or complements.

cally have a disproportionately large share of manufacturing firms and can support substantially higher wages than remote regions.[5]

Two implications follow from this. The first is that size matters: A location with a large market will tend to draw in manufacturing activity, possibly bidding up the wage in the location. The second implication follows from combining this with labor mobility. If labor is mobile between locations, then the higher wage will attract labor inflow, enlarging the market still further. This interaction between firms wanting to locate in large markets and demand from their workers enlarging the market provides the basis for a process of cumulative causation leading to spatial concentration of activity. Krugman (1991b) shows how, if transport costs are low enough, mobile factors will agglomerate in just one location.

Although labor mobility can provide a basis for agglomeration of activity, it is not a necessary condition for it to occur. Much of the demand for firms' output comes not from final consumers but from other firms that purchase intermediate goods and services. Thus, as downstream firms move to a location they enlarge the market for upstream firms, and as upstream firms move they increase the supply and lower the price of intermediate goods. This interaction can create cumulative causation and clustering of linked industrial activities in a location (Venables 1996). The process is no more than the interaction of forward and backward linkages that received so much attention in the development literature of the 1960s, and whose origins date back (at least) to Marshall (1920), in whose words

> Subsidiary trades grow up in the neighbourhood, supplying it with implements and materials, organising its traffic, and in many ways conducing to the economy of its material . . . [T]he economic use of expensive machinery can sometimes be attained in a very high degree in a district in which there is large aggregate production of the same kind. . . . [S]ubsidiary industries devoting themselves each to one small branch of the process of production, and working it for a great many of their neighbours, are able to keep in constant use machinery of the most highly specialised character, and to make it pay its expenses.

The contributions of the new literature are to identify circumstances under which these linkages will lead to clustering of activity and the extent to which they support wage differences between locations. As we will see, outcomes depend critically on the level of trade costs, so clustering occurs at some levels of trade costs, and dispersion at other levels.

Other forces too can give rise to spatial clustering of activity, and we note

5. The implications of good market potential for production are sometimes called the "home market effect." Davis and Weinstein (1997) find considerable empirical support for it. Wage implications of market access are studied in Redding and Venables (2000). The advantages of coastal regions and other geographical factors in developing countries are documented in Gallup, Sachs, and Mellinger (1999).

just two further mechanisms, drawing on Marshall's treatment. His second clustering force is a thick labor market:

> 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 labour, is often put to great shifts for want of some special skilled labour; and a skilled workman, when thrown out of employment in it, has no easy refuge.

This—while undoubtedly important—has received much less attention in the modern literature, although see Krugman (1991a) for a rudimentary model.

The third mechanism is geographically concentrated technological externalities:

> 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 organisation 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.

This idea is applied in much of the regional and urban literature (see, e.g., Henderson 1974), as well as in some older trade literature (Ethier 1979). It is perhaps best viewed as a black box for a variety of difficult-to-model yet important proximity benefits.

7.3.2 History of the World?

Once these clustering forces are put in a full general equilibrium model of trade and location, what happens, and what predictions are derived for the effects of globalization? A sweeping view of world history is provided by the model by Krugman and Venables (1995) that studies the effects of falling trade costs on industrial location and income levels. Their model has just two countries (N and S), endowed with the same quantities of internationally immobile labor. There are two production sectors: perfectly competitive agriculture, and manufacturing. Manufacturing has increasing returns (modeled as monopolistic competition) and forward and backward linkages (modeled as firms using manufactured products as well as labor to produce output for use by other firms as well as for final consumption).

The Krugman and Venables story is summarized in figure 7.6, which has trade costs on the horizontal axis and real wages in N and S on the vertical axis. At very high trade costs the two economies have the same wage rates

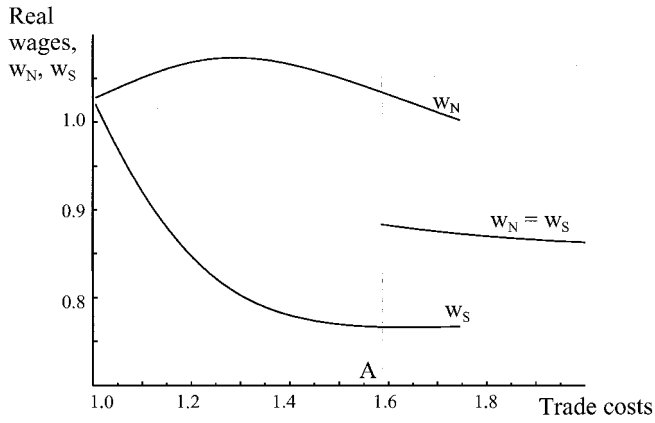

Fig. 7.6 History of the world

($w_N = w_S$), reflecting the fact that they are identical in all respects. The linkages between manufacturing firms create a force for agglomeration, but when trade costs are high these are dominated by the need for firms to operate in each country to supply local consumers. As trade costs fall (moving left on the figure), so the possibility of supplying consumption through trade rather than local production develops, and clustering forces become relatively more important. At point A clustering forces come to dominate, and the equilibrium with equal amounts of manufacturing in each country becomes unstable; if one firm relocates from S to N then it *raises* the profitability of firms in N and *reduces* the profitability of firms in S, causing further firms to follow. Four forces are at work. By moving to N the firm raises wages in N and increases supply to N consumers, both effects tending to reduce profitability in N. But against this it increases the size of the N market (the backward linkage, creating a demand for intermediates) and reduces the costs of intermediates in N (the forward linkage, offering a supply of intermediates). The last two effects come to dominate, and we see agglomeration of industry in one country, which raises wages in the country with industry, as illustrated.[6]

For a range of trade costs below A, the world necessarily has a dichotomous structure. Wages are lower in S, but it does not profit any firm to move to S because to do so would be to forgo the clustering benefits of large markets and the proximity to suppliers that are found in N. However, as trade costs fall it becomes cheaper to ship intermediate goods; linkages matter less, so the location of manufacturing becomes more sensitive to factor

6. There is a range in which agglomeration (with $w_N > w_S$) and dispersion (with $w_N = w_S$) are both stable equilibriums. See Fujita, Krugman, and Venables (1999) for details.

price differences. Manufacturing therefore starts to move to S and the equilibrium wage gap narrows. In this model, wage gap goes all the way to factor price equalization when trade is perfectly free—the "death of distance."

The relationship between the model and our earlier discussion of the changing spatial patterns of industrial location is apparent. Falling trade costs combined with industrial linkages offer an explanation of both the concentration of manufacturing activity and its dispersion. As trade costs go from very high to somewhat lower levels there is deindustrialization of some regions and widening income gaps—the first phase of concentration of activity. At lower levels of trade costs, industry starts to spread out of established centers to some lower-cost regions.

Of course, the model is stylized, and many extensions are needed if it is to be convincingly linked to the historical record. We discuss some of these extensions in detail in following sections, and outline them here. Most obviously, figure 7.6 assumes international immobility of labor. Labor migration can be an additional force for agglomeration—at the city and regional level, as well as internationally—and was clearly important historically, notably in the nineteenth and early twentieth centuries. We return to this in section 7.4. Also, the story needs to be enriched to include many countries, many sectors, and other clustering mechanisms. If there are many countries then the convergence phase is no longer smooth; it involves an increasing number of locations with industry, rather than steady industrialization of them all. Other clustering mechanisms may interact with trade costs in different ways from the linkages described above. For example, the strength of clustering forces arising from labor market skills is likely to be largely unaffected by trade costs. In sectors where this is important, falling trade costs will not bring about the death of distance, and clusters are likely to remain in place. We take up some of these issues in following sections, and formal analysis of them is undertaken in Fujita, Krugman, and Venables (1999).

7.4 The Nineteenth Century

We have already shown the potential that a geographical approach has for the explanation of one of the three phases highlighted in figures 7.1 and 7.2, namely, the decline—absolute as well as relative—of industrial activity outside the emerging core of northwestern Europe. In this section we want to pursue two further aspects of the nineteenth-century experience in greater detail. One is the rise of the New World, and the other is the growth of urbanization.

7.4.1 The Economic Development of the New World

Following the relative and absolute decline of the Asian economies, the other main change in the economic geography of the nineteenth-century world was the rise of the New World, and within this the particular domi-

nance of the United States. By 1913, the United States was a leading industrial producer and a successful manufacturing exporter. Its industrialization, accomplished behind high tariff walls, was concentrated in the "manufacturing belt" of the northeast, a region with the highest GDP per person in the world. This had not seemed at all probable in 1860, when America's role in the world economy was apparently destined to be that of a large primary products exporter based on an abundant endowment of natural resources.

This prompts two obvious, related questions. First, why did the United States rather than Latin America become the area that overtook the United Kingdom and the rest of Europe in real GDP per person? Second, why did the United States also become the only non-European country to establish a position as a net manufactured exporter? In 1913, while Canada, Latin America, and Oceania had net imports of manufactures to the value of $525 million, $828 million, and $361 million, respectively, the United States had net exports of $368 million and already represented the third largest share in world manufactured exports (Yates 1959).

The contrast with the overall experience of Latin America was marked. The nineteenth century can be seen as a period when Latin America fell seriously behind, although by 1913 its most successful economy, Argentina, had experienced several decades of rapid growth and had an income level greater than many European countries. Even so, Argentina had failed to match the United States over the course of the nineteenth century from a position of near parity of incomes per head in 1800 (Coatsworth 1998). Latin America as a whole, which accounted for slightly more of world GDP than the United States in 1820, produced only 4.5 percent in 1913, compared with 19.1 percent for the United States and fell from a level of GDP per person of 52.9 percent that of the United States in 1820 to 28.5 percent in 1913.

Recent interpretations of these developments by economic historians have stressed the differing role of institutions and rent-seeking in North and South America and the political economy configurations from which they emerged. North, Summerhill, and Weingast (2000) pointed to the unfortunate legacy of the ending of Spanish colonialism and an associated failure to establish secure political foundations for economic growth in Latin America; they contrasted this outcome with the aftermath of British rule in the United States, which resulted in a constitution with strong protection of property rights. They see this as the crucial difference: "No deus ex machina translates endowments into political outcomes. If that were so, Argentina would be as rich as the United States" (2000, 19).

Engerman and Sokoloff (1997) also argued that institutions made all the difference to development outcomes between Latin and North America but placed their emphasis on the role of initial factor endowments in creating institutional divergence that exhibited path-dependent tendencies. Inter-

estingly, one of the ways in which their story plays out is through different implications for labor inflow with small family farms in North America conducive to good institutions and greater equality of wealth and political power which underwrote both rapid growth and high immigration.

Similarly, David and Wright (1997) have pointed to several highly favorable aspects of American institutions and policies for the exploitation of abundant resources that led to American primacy in the minerals-based, resource-intensive technology that was central to technological progress in the early twentieth century. These included promoting education and scientific research in relevant disciplines, subsidizing transportation, and organizing geological surveys and sustaining minerals property rights but without claiming government entitlement to royalties. Organized thus, American endowments promoted a technological trajectory that no European country could emulate.

We have no wish to dispute these claims, but we do suggest that it is important that they are placed more firmly in a geographic context. In particular, we believe that size and increasing returns to scale mattered. Table 7.4 displays some information on the size of the United States compared with other leading New World economies and the United Kingdom. In addition, we disaggregate the United States into the northeast and other regions. It is clear that, by 1870, when international transport costs began to fall rapidly, the United States was already a very large economy.

Indeed, at that time, the United States had almost matched the United Kingdom in terms of total GDP, and its population was nearly a third larger. The population of the United States by then already exceeded that of the whole Latin American and Caribbean area defined by Maddison (2001), and its GDP was well over three times larger. Relative to the other individual economies of the New World the United States was in a completely different league in terms of the size of its economy. This was also true, however, for the northeast, which taken separately matched the United Kingdom in terms of GDP per person around 1880 and for population by about 1900. This region already had 29.5 percent of the labor force in manufacturing in 1870, rising to 38.7 percent by 1910 (Perloff et al. 1960), far ahead of any New World country and approaching British levels of industrialization.

The growth of the New World economies was boosted by massive factor flows from the Old World. Declining costs of transport, together with rising incomes in a world relatively free of immigration restrictions, encouraged large international migration. Between 1870 and 1910 this augmented the New World labor force by 40 percent while at the same time reducing the Old World labor force by 13 percent. The impacts on labor force size in some individual countries were much larger—for example, an increase of 86 percent in Argentina and a fall of 45 percent in Ireland—while the U.S. inflow amounted to 24 percent and Great Britain's outflow to 11 percent of the 1910 labor force (Taylor and Williamson 1997). The ratio of foreign as-

Table 7.4 Population, GDP, and GDP per person, 1870 and 1913

	1870	1913
Population (in thousands)		
Argentina	1,796	7,653
Australia	1,770	4,821
Canada	3,781	7,852
United States	40,241	97,606
Northeast	21,609	49,193
Rest of country	18,632	48,413
United Kingdom	31,393	45,649
GDP (in millions of 1990 international dollars)		
Argentina	2,354	29,058
Australia	6,452	27,552
Canada	6,407	34,916
United States	98,374	517,383
Northeast	65,615	320,004
Rest of country	32,759	197,379
United Kingdom	100,179	224,618
GDP per person (in 1990 international dollars)		
Argentina	1,311	3,797
Australia	3,645	5,715
Canada	1,695	4,447
United States	2,445	5,301
Northeast	3,036	6,505
Rest of country	1,758	4,077
United Kingdom	3,191	4,921
Real Wages		
Argentina	86	101
Australia	169	127
Canada	147	200
United States	165	160
United Kingdom	100	100

Source: Maddison (2001); U.S. regional figures approximated using the data in Perloff et al. (1960) and their definition of the northeast, which comprises the New England, Middle Atlantic, and Great Lakes regions. Real wage comparisons from Williamson (1995, 1998).

sets to world GDP grew from 7 percent in 1870 to 18 percent in 1914, about the same level as in 1980 (Obstfeld and Taylor, chapter 3 in this volume). The United Kingdom was the principal capital exporter, and outflows averaged almost 5 percent of GDP; 34 percent of all British foreign investment went to North America, compared with 17 percent to Latin America (Simon 1968). Further discussion of Old World–New World factor flows from a neoclassical perspective can be found in Lindert and Williamson, chapter 5 in this volume.

7.4.2 Modeling Migration and Development

If we take the drivers of change to be falling transport costs of goods and factor mobility (in particular labor migrations, facilitated by falling costs of

moving people), the challenge for a model is to explain the following stylized facts: the continuing wage advantage of North America relative to the United Kingdom and to other New World economies, despite migration flows; the rise of manufacturing in the United States, overturning its apparent comparative advantage in agricultural products; and the failure of manufacturing to develop in other New World economies.

The overtaking of Great Britain by the United States used to be explained in terms of various kinds of market failure in the former. Briefly, these arguments claimed that inefficiencies in the capital market encouraged excessive foreign and inadequate domestic investment, while conservative British firms were slow to adopt new techniques and to diversify into new industries. These claims have, however, failed to stand up to the scrutiny of economic historians using neoclassical economics because it is now recognized that foreign investment was justified in terms of its returns, diversification into new lines of activity was not impeded by the capital market, and choices of technique were rational given British factor costs (Crafts 2002a).

Given the successful use of neoclassical economics to debunk crude claims of British failure it is perhaps not surprising to find that existing studies modeling the development of North-Atlantic economy have been built largely on a comparative advantage trade model. Applications of the Heckscher-Ohlin model are said to have performed well (Hutchinson 2000; Wright 1990). In this tradition, O'Rourke and Williamson (1994) concluded that a calibrated computable general equilibrium (CGE) model of this type allowed a good explanation of trends toward Anglo-American factor price convergence, driven by commodity market integration in the face of falling transport costs and by labor migration. General equilibrium modeling in this tradition by O'Rourke and Williamson (as summarized by O'Rourke 1996) found that over the period 1870–1910 and initial wage gap of 71.2 percent between the United States and the United Kingdom would have been reduced by 34.8 percentage points by migration but raised by 13.7 percentage points by capital flows. The net impact of factor flows would therefore have reduced the gap by 24 percentage points, to which commodity market integration would have added a further 28.5 percentage points. Interestingly, however, instead of narrowing sharply (by 52.5 percentage points) the U.S.-U.K. wage gap decreased by just 5 percentage points (table 7.4).[7]

This last points to American access to sources of productivity improvement not available to the United Kingdom and consequently reveals a serious problem with the neoclassical exoneration of the late Victorian British economy, namely, that it does not have an adequate explanation for American overtaking. Indeed, it might be argued that this is a general difficulty with neoclassical growth economics; in principle, it can readily embrace catching-up and convergence but not changing leadership in a Solovian

7. Or possibly actually widened, depending on data sources used.

world of constant returns to scale and common technology. There are two ways to address this issue in the context of the North-Atlantic economy. The first is to drop the assumption of common technology and argue that the United States developed its own technology (based on cheap raw materials and mass markets) that was not transferable to Europe at this time (Abramovitz and David 1996). This has historical plausibility but may not be the whole story, however. The second, relatively neglected, is to focus on the role that geography played through scale economies and agglomeration benefits. While accepting that localized technical change also mattered, we explore this by undertaking some rather simple formal modeling to draw out the differences between a comparative advantage approach and a new economic geography approach, and to argue that the latter does much better at explaining both the factor-price and the quantity side.

Before developing the models, it is worth recording other voices that have suggested that the traditional neoclassical framework does not encompass an important part of the picture in that economies of scale in manufacturing are ignored. Although this may be entirely reasonable for the pre–Civil War American economy, it is much less appropriate for the later nineteenth century. Both the traditional business history literature (Chandler 1977) and the cliometricians (Cain and Paterson 1986; James 1983) agree that economies of scale in manufacturing between 1870 and 1913 were substantial and pervasive. These were associated with labor-saving and materials-using biases in technological change and were exploited in the context of a large and rapidly expanding domestic market. Moreover, a closer look at trade flows also reveals some limitations of analyses of the Heckscher-Ohlin type. There was already a considerable amount of intraindustry trade prior to World War I, and this was associated with scale economies in labor and materials use (Hutchinson 2000).

Turning to the modeling, let us start with a stylized model of how people and activity relocate between world regions. We suppose that there are three regions, each having the same endowment of land and the same spatial relationship to each other (they are located at vertices of an equilateral triangle). The model is intended to be suggestive of the location of activity between Europe, the United States, and the rest of the New World, but we impose symmetry in order to get to the heart of the economic forces at work. We assume that there are two production sectors, agriculture and manufacturing. The output of both these sectors is tradable, although both are subject to transport costs. Production in agriculture uses labor and land, and manufacturing uses labor and manufactures (as an intermediate good). Sales of agriculture all go to final consumption, but sales of manufacturing go both to final consumption and to meet the derived demand for manufactures from manufacturing industry. The structure of the model is similar to Krugman and Venables (1995) and is set out formally in the appendix.

The experiment that we undertake is to start with an initial position in

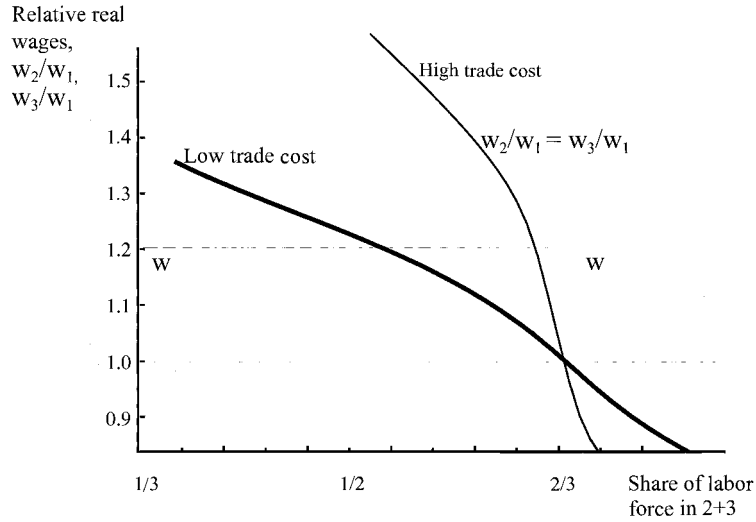

Fig. 7.7 Relative wages in competitive model

which most of the labor force is in region 1 (Europe) and to look at the effects of moving labor out of this region. As it moves we shall assume that it goes to regions 2 and 3 in a way that equalizes real wages in 2 and 3. The idea we seek to capture is that there is out-migration from 1 that is costly, but (in the spirit of our symmetry assumption) the same migration costs are incurred in going to either of the other regions. In the exposition that follows we will talk as if out-migration from region 1 is exogenous. However, we keep track of the real wage gap between region 1 and other regions, so inverting the relationship between the distribution of population and the wage gap shows how a given wage gap (equal to the migration cost) is consistent with a level of population movement.[8]

Globalization and Geography: The Competitive Model

We start with a perfectly competitive variant of the model, in which production in all sectors takes place under constant returns and comparative advantage is determined entirely by factor endowments. The proportion of the world labor force in regions 2 and 3 combined is measured on the horizontal axis of figure 7.7, so that migration is measured by a movement to the right along the figure. On the vertical axis we measure the real wages in regions 2 and 3 relative to region 1 ($w_2/w_1 = w_3/w_1$). The light line is for a case when goods trade costs are high (both agriculture and manufacture face an iceberg transport cost factor of 1.7), and the heavy line corresponds to a lower transport cost factor of 1.25.

8. Migration plays a central role in our story. Further analysis of the economic impact of migration in this period can be found in Chiswick and Hatton, chapter 2 in this volume.

The information contained in the figure is in line with expectations. We see that as long as region 1 is labor abundant (regions 2 and 3 combined have less than two-thirds of the world labor force although each region has one-third of the land), then the wage in regions 2 and 3 exceeds the wage in region 1. Migration narrows the wage gap, as does a reduction in the cost of shipping goods (as in O'Rourke and Williamson 1994). At a given level of migration costs, indicated by the horizontal line *ww*, migration flows are smaller the lower are trade costs, indicating that factor mobility and goods trade are substitutes.

Figure 7.7 is the benchmark case, demonstrating how either factor flows or goods trade liberalization causes factor price convergence. However, in this competitive variant neither region 1 nor region 2 can become a net exporter of manufactures. These regions expand their share of manufacturing only by attracting labor inflow and attract labor inflow only by being land abundant, and hence net importers of manufactures. This means that there is no mechanism in this model by which an economy that initially has a comparative advantage in agriculture can overturn this and become a net exporter of manufacturing. Furthermore, regions 2 and 3 are, in this model, bound to follow identical development paths. Given symmetry in technology, preferences, and endowments, the two regions must have the same outcomes.

Globalization and Geography: The Monopolistic Competition Model

The second variant of the model makes manufacturing monopolistically competitive, containing firms that operate under increasing returns and are subject to forward and backward linkages.[9] The model now predicts a quite different development path, for two main reasons. The first is that market size (as well as factor prices) becomes an important determinant of where manufacturing locates; as we have already seen, if two locations differ only in market size then disproportionately many firms will locate in the larger market. The second reason for the different development path is the propensity of manufacturing to agglomerate, arising because of forward and backward linkages and reinforced by the mobility of labor.[10]

Panels A and B of figure 7.8 give the case when trade costs are relatively high. Like figure 7.7, the horizontal axis measures the combined population of regions 2 and 3. On the vertical axis, panel A of figure 7.8 has relative real wages and panel B has the share of world manufacturing activity in region 2 and in region 3. We see that when the combined population of regions 2 and 3 is small all manufacturing is agglomerated in region 1. The wage in regions 2 and 3 is quite high because of high land-labor ratios, and adding

9. As outlined in the appendix. This is the same structure as in section 7.3.2, except that there are three regions, labor migration is studied, and agriculture, as well as industry, has transport costs and product differentiation.

10. This input-output structure was also present in the perfectly competitive variant of figure 7.6, but the linkages are important only when combined with increasing returns.

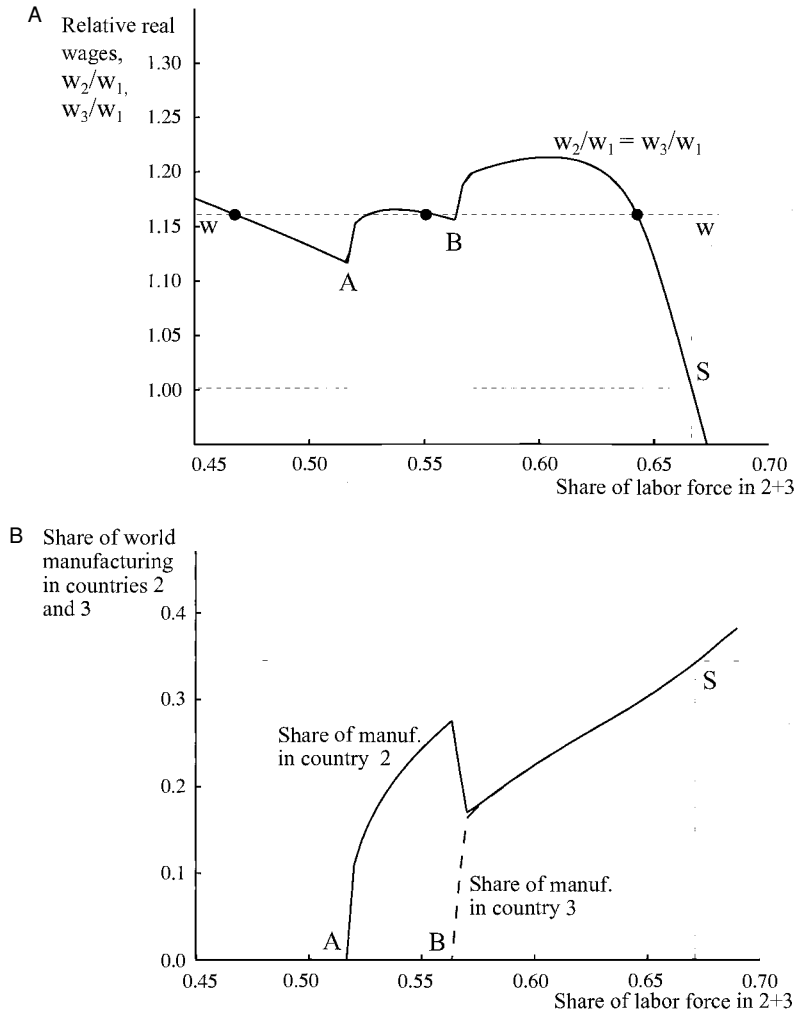

Fig. 7.8 High trade cost: *A*, Relative wages; *B*, Manufacturing shares

more labor reduces the wage gap (like in fig. 7.7). However, as the labor force of regions 2 and 3 increases, the combination of lower wages and larger market size makes it profitable for manufacturing activity to start in these regions (at point A). Industrialization in both simultaneously is, however, unstable, because if one region got just slightly ahead then agglomeration benefits would make it more profitable, attracting more manufacturing and more labor inflow. If the regions are identical it is a matter of chance which one industrializes, and we suppose that it is region 2, as indicated in panel B. The effect of this is to increase wages, as labor in region 2 is drawn off the land and into manufacturing.

Further labor outflow from region 1 will go predominantly to region 2, but after some point the additional labor in these regions starts to reduce wages again, as well as further enlarging market size. Region 3 then industrializes (at point B), catching up with region 2, and resulting in another increase in relative wages. The relative wage path illustrated in panel A of figure 7.8 can then be understood in terms of labor inflow tending to depress wages as land-labor ratios fall, punctuated by industrialization episodes raising labor demand and wages. We can also use panel A of figure 7.8 to analyze the endogenous migration story. If migration were perfectly free to respond to any wage differential, however small, then labor flows would move the world economy to point S, at which all three regions are identical with the same economic structures and factor price equalization. Alternatively, if we contrive migration costs to be just sufficient to support a wage gap illustrated by the line *ww*, then there are three stable migrational equilibriums (as well as two unstable), as marked by the solid circles. Thus, the equilibriums of industry in just region 1, in regions 1 and 2, or in regions 1, 2, and 3 are all stable equilibriums. However, at this level of migration costs, simple dynamics starting with population concentrated in region 1 would leave the world in the first of these equilibriums, with regions 2 and 3 remaining agricultural.

Panels A and B of figure 7.9 are analogous, but computed for a lower value of trade costs. There are three main differences. First, manufacturing commences in region 2 only when more population has moved to regions 2 and 3; this is because a larger market size is required to offset the effects of more intense import competition from region 1. Second, manufacturing never takes off in region 3—again, because of the more intense import competition it faces; essentially, at this level of transport costs world demand for manufactures can be met from just one or two clusters. Region 2 therefore develops a different economic structure from region 3, with a larger population and higher share of world income. Third, region 2 becomes a net exporter of manufactures, and this occurs at the point at which its share of world production of manufactures exceeds its share of world income (see panel B).

Lower trade costs have the effect of decreasing wages in labor-abundant economies (as in the competitive case, fig. 7.7), and the wage path is illustrated by the heavy line *aa* on figure 7.9, panel A. There is a kink in this curve at the point at which industrialization in region 2 commences, but in the case illustrated this kink occurs when real wages are lower in regions 2 and 3 than in region 1, suggesting that migration would not bring about sufficient labor movement to reach the point at which manufacturing develops. Thus, product market integration has the effect of locking the manufacturing agglomeration into an established center and also, because of the labor demand this creates, of reducing the incentive for out-migration from this center.

The higher wage curves in panel A figure 7.9 offer some responses to this dilemma. The first of these, *bb*, is computed allowing transport costs to fall

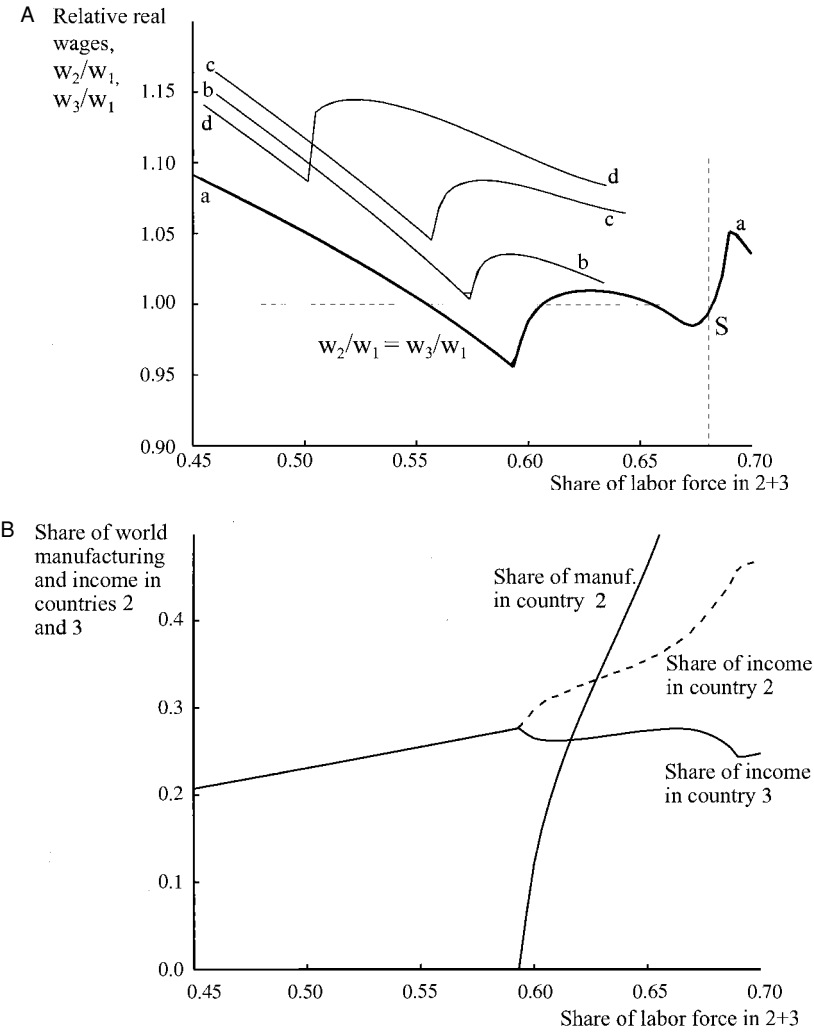

Fig. 7.9 Low trade cost: *A*, Relative wages; *B*, Manufacturing and income shares

concurrently with the movement of labor. Higher trade costs mean higher wages in regions 2 and 3 relative to region 1, thus increasing migration flows and creating a migration path that supports industrialization by region 2 although not by region 3. Thus, the model is able to explain both the asymmetric development of ex ante identical countries and the way in which an economy with initial comparative advantage in agriculture can industrialize and become a net exporter of manufactures.

The next curve, *cc*, introduces (additionally) an asymmetry between regions, letting region 2 have 20 percent greater land endowment than regions

1 and 3. This raises wages and increases the population and market size of region 2, thereby bringing forward the industrialization of region 2, as indicated by the position of the kink in this wage schedule. The final curve, *dd*, illustrates the effect of a region 2 import tariff on manufactures of 10 percent (on top of the different land endowment). When region 2 has no industry this reduces real wages, for the usual reasons of welfare loss associated with tariffs. However, the tariff brings forward industrialization, which in turn raises wages and accelerates the growth of population of the region.

Although these are very stylized exercises, we think that viewing the development of the New World through this lens can offer important insights. As we have seen, it offers an explanation of how one region can industrialize while another does not; of how this region can have its relative wages increase despite population inflow; and of how it can become a net exporter of manufactures despite its initial comparative advantage. The importance of scale effects suggests that open-access migration policies may have mattered much more than is generally acknowledged, and it indicates a potentially more powerful and different role for tariff policies than traditional analyses would allow (Irwin 2000). Given the interrelationship between migration and trade costs, the model also highlights the importance of the timing of the transport improvements that came when migration into the American economy was already substantial and when manufacturing production functions were being transformed. None of this should be taken to negate the insights of those economic historians who have rightly pointed to the role of institutions in growth outcomes, but it might suggest that undue emphasis on property rights is to be avoided.

Finally, in related research, Crafts and Venables (2001) have simulated the development of the North Atlantic economy using a computable general equilibrium model, calibrated to 1870 and 1913 data. This gives results that are consistent with the approach set out here. The competitive variant of the model cannot replicate the large growth of American manufacturing and predicts a large decline in the U.S.-U.K. real wage gap, whereas incorporating increasing returns and linkages in the manufacturing sector largely rectifies these deficiencies. Comparison of results with and without the high American tariffs of the period shows that these had a substantial positive effect on industrialization based on the positive feedbacks associated with the migration that it induced. In nineteenth-century conditions, these results suggest that the United States, starting with an "empty country," gained from employing the opposite of the current OECD policy norm of blocking migration from poorer countries and freeing up trade.

7.4.3 Urbanization

The implications of declining trade costs are felt at the subnational level as well as internationally, and, just as they facilitated concentration of world manufacturing, so they also promoted the development of urban agglom-

Table 7.5 **Urbanization Levels (% population, criterion of 5,000 for urban population)**

	1800	1850	1910	1980
England	23	45	75	79
France	12	19	38	69
Germany	9	15	49	75
Europe	12	19	41	66
United States	5	14	42	65
Australia	—	8	42	80
Latin America	14	18	22	63
Third world	9	9	10	32

Source: Bairoch (1988, tables 13.4, 29.1).
Note: Dash indicates data are not available.

erations. Urbanization is one of the most dramatic changes in economic geography that occurred during the decades before World War I, and its impact on the location of the labor force far outweighed that of international migration. Whereas about 34 million people emigrated from European countries between 1851 and 1910 (Ferenczi and Willcox 1929), the increase in urban population in Europe and North America totalled 145 million in the same period. Moreover, the number of large cities grew disproportionately: Whereas in 1800 there were 24 cities in the developed world with a population over 100,000, by 1914 this had risen to 281 (Bairoch 1988). Table 7.5 reports the rapid increase in urbanization rates in these countries that contrast with an unchanged urban proportion in the third world.

The hypothesis underlying the growth of urban centers is simple: The division of labor is limited by the extent of the market, and improved transport technologies overcome this, enabling production to take place on a larger scale (and with more division of labor) and enabling cities to form and reap the agglomeration benefits outlined above.

Until recently, there has been surprisingly little formal economic analysis of this hypothesis. The central-place theory of Losch (1954) puts forward the trade-off between returns to scale and transport costs, but its focus is on the optimal lattice of market areas, rather than the equilibrium size and structure of cities. Henderson (1974) broke with these traditions and modeled city size on the basis of technological externalities within industries. But at the same time, he took a strangely aspatial approach, saying nothing about where cities are located, the spatial nature of economic interactions, or the role of transport and communications technologies in enabling city formation. Fujita (1989) developed both an explicit geography and microfoundations for returns to scale.[11] The trade-off between transport costs

11. Fujita, Krugman, and Venables (1999) extend this approach. See also Puga (1998) for the interaction of scale economies and transport costs.

and increasing returns means that the real wage that can be paid in a city is a function of its size, and there is in general a unique city size that gives the maximum wage. This wage-maximizing city size depends on transport costs; as transport costs fall, cities will become larger.

In Great Britain, this process became really apparent during the canal era of the late eighteenth century and is epitomized by the growth of Birmingham (Turnbull 1987). Detailed simulation of British experience during the Industrial Revolution reveals that cities were underpinned by an elastic supply of agricultural imports from the rest of the world contingent on the development of an improved commercial and transport infrastructure (Harley and Crafts 2000; Crafts and Harley 2002). For nineteenth-century Europe, regression analysis shows that the major influences on the pace of urbanization across countries were the growth of industrialization, international trade, and agricultural productivity (Bairoch and Goertz 1986). Lowered transport costs and, in particular, new rail facilities facilitated the growth of large cities in nineteenth-century America and gave rise to agglomeration benefits as the costs of moving goods fell. The division of labor was enhanced by increased market size (Ades and Glaeser 1999). Small-scale producers in regions like the midwest were disadvantaged, and manufacturing activities became increasingly spatially concentrated. By 1890, over 25 percent of value added in American manufacturing originated in New York, Philadelphia, and Chicago (Pred 1977).

Although clustering of activities promotes development of cities, there has been debate about the extent to which clustering forces are industry specific or broader. If they are industry specific, then the process of city growth will be accompanied by specialization. High degrees of specialization were indeed an important feature of nineteenth-century cities. In 46 of the largest 100 American cities in 1880, one or two industries accounted for more than 50 percent of manufacturing employment (Kim 2000). In the second half of the nineteenth century large cities were increasingly industrial, and spatial concentration of manufacturing in already large cities was a prominent feature of the industrialization experience.

David's (1989) study of Chicago confirmed that its phenomenal growth was founded on agglomeration effects rather than internal economies of scale. Chicago's success, however, stemmed from a diversified industrial base that suggests that interindustry knowledge spillovers may also have been important, as hypothesized by Jacobs (1969). This is also a strong theme in the account of late nineteenth-century European urbanization by Hohenberg and Lees (1985), who stressed an explosion of knowledge-centered economic growth, and it appears to be borne out by the econometric investigation of English city growth performed by Simon and Nardinelli (1996).

In fact, it appears likely that both industry-specific (Henderson-type) and Jacobs-type external economies of scale were operative in Victorian

cities and could be of substantial importance. This is the conclusion reached by Broadberry and Marrison (2002) in an analysis of the British cotton textiles industry on the eve of World War I. They found that both types of external-scale economies were critical to the industry's ability to withstand foreign competition from relatively low-wage producers.

Thus, although we usually think of globalization as occurring at the international level, its driving forces are also important at the subnational level. They promoted the urbanization of the nineteenth-century world, thereby facilitating the division of labor and exploitation of returns to scale associated with industrialization. In addition, it is clear that the experience of nineteenth-century urbanization bears out the value of the new economic geography approach to explaining the location of production.

7.5 The Late Twentieth Century

The interwar period is well known to have been a period of globalization backlash in which there was disintegration of the world economy. This was an epoch of trade wars and international capital controls, and also a time when transport costs ceased to fall. The reconstruction of the world economy after World War II involved a successful liberalization of international trade in manufactures under the General Agreement on Tariffs and Trade (GATT) and a resurgence in international capital mobility, notably from the breakdown in the Bretton Woods fixed exchange rate system in the early 1970s. As table 7.2 reported, the ratio of world merchandise exports to world GDP, which had fallen to 5.5 percent in 1950, rose to 10.5 percent by 1973 and to 17.2 percent in 1998. Foreign assets as a proportion of world GDP, which had fallen to 5 percent in 1945, regained the 1914 level of 18 percent in 1980 and by 1995 had surged to 57 percent (Obstfeld and Taylor, chapter 3 in this volume).

Falling transport and communications costs continued to be a driver of globalization, as table 7.1 suggests. A decline in shipping costs was augmented by several other important developments. One was the development of new information and communications technologies (ICT), the implications of which we discuss in section 7.5.3. Another was the reduction in transit times associated with the development of air travel (and airfreight) and the development of containerization, bringing both faster port handling and faster ocean shipping. The importance of time in transit has been estimated in recent work by Hummels (2000), who finds that the cost of an extra day's travel is (for imports as a whole) around 0.3 percent of the value shipped. For manufacturing sectors, the number goes up to 0.5 percent, costs that are around 30 times larger than the interest charge on the value of the goods. One implication of these figures is that transport costs have fallen much more through time than is suggested by looking at freight charges alone. The share of U.S. imports going by airfreight rose from zero

to 30 percent between 1950 and 1998, and containerization approximately doubled the speed of ocean shipping. Together these innovations give a reduction in average shipping time of twenty-six days, equivalent to a shipping cost reduction worth 12–13 percent of the value of goods traded.

The growing value of trade only tells part of the story, because there were also new types of trade developing. The growth of international production networks is reflected in growing volumes of trade in parts and components. Yeats (1998) estimates that 30 percent of world trade in manufactures is trade in components rather than final products. Hummels, Ishii, and Yi (2001) chart trade flows that cross borders multiple times, as when a country imports a component and then re-exports it embodied in some downstream product. They find that (for ten OECD countries) the share of imported value added in exports rose by one-third between 1970 and 1990, reaching 21 percent of export value.

Finally, the period saw the growing role of FDI. Although the world FDI stock showed virtually no growth between 1938 and 1960 (Jones and Schroter 1983) after that it rose rapidly from 5.4 percent of world GDP in 1980 to 14.1 percent in 1998 (World Bank 2000). The vast majority of this capital was in Europe and North America—about 68 percent in 1980 and 63 percent in 1999—but East Asia, not including China or Japan, had 10 percent already by 1980, and China's share had grown to over 6 percent by 1999.

7.5.1 Divergence, Big Time

Traditional neoclassical theories of economic growth predict convergence of incomes based on the catch-up of countries with initially low levels of (broad) capital and output per worker in a world of universally available technology. The empirical application of these ideas has usually been phrased in terms of conditional convergence allowing some role for differences in rates of factor accumulation (Barro and Sala-i-Martin 1995). The actual experience of the world in the twentieth century has, however, been described recently as "divergence, big time" (Pritchett 1997) which is not surprising given the trends reported in table 7.6. Whereas in 1870 income per head in Africa was about one-eighth that in the leading country, by 1998 the ratio was about one-twentieth (Maddison 2001). In 1998, as table 7.6 shows, many of the world's population lived in countries where income levels were a lower percentage of the U.S. level than in 1950. Western Europe and East Asia gained ground relatively, while other countries fell back. This is the pattern of "twin peaks" highlighted by Quah (1997).

A variant on the neoclassical perspective is provided in Lucas (2000). He argues that the divergence of the twentieth century will be reversed because sooner or later every country will join the industrial revolution as best-practice policies and institutions are imitated in hitherto unsuccessful countries and thus the Solovian assumption of "universal technology" be-

Table 7.6 **Real GDP per Person Gaps with the United States (United States = 100)**

	1950	1998
Western Europe	48.0	65.6
Eastern Europe	22.2	20.0
China	4.6	11.4
Japan	20.1	74.7
Other East Asia	9.6	20.1
British India	6.4	6.1
Latin America	26.7	21.2
Africa	8.9	5.0

Source: Maddison (2001).

Note: In each year the income level is expressed as a percentage of the U.S. level. Regions defined as in figure 7.1.

comes valid: "The restoration of inter-society income equality will be one of the major economic events of the century to come" (Lucas, 166). He bases his prediction on a simple model in which new entrants to the growth process start at $(2 + 2.5n)$ percent per year, where n is the number of fifty-year periods to have elapsed since 1800; thus, a country experiencing take-off in the early twenty-first century will grow initially at 12 percent per year, compared with 7 percent for the 1900 entrant. All countries have an equal chance of joining the growth club with a hazard rate evolving from .01 to .03 over time. This last assumption is clearly contrary to the predictions of the geography school as well as the new institutional economic history.

The new institutional economic historian's perspective provided by North (1990) sees institutions as the key stumbling block. In this view there are no Coasian bargains available to ensure that bad institutions are replaced; rather, the world is one of path dependency, where network externalities, vested interests spawned by the existing arrangements, and informal constraints, embodied in customs, traditions, and codes of conduct that are impervious to deliberate policy reform, hold sway. The economic geography perspective argues that agglomeration benefits dominate the development process such that size and distance matter, as set out in section 7.3.

The evidence of growth regressions certainly suggests that institutions have a strong effect on growth outcomes (Knack and Keefer 1995) and bad institutions remain unreformed in many countries (Kaufmann, Kraay, and Zoido-Lobaton 1999). But recent experience also shows that institutional reform in the third world has delivered a good deal less than followers of growth regressions might have expected (Easterly 2001). This last study found that, taking into account standard conditioning variables, third world growth is strongly influenced by growth in the country's main OECD trading partner. It is also typically the case either that regional dummy variables (East Asia, positive; Africa and Latin America, negative) show up strongly or that justifying their omission requires the inclusion of explicit

geographic variables (Gallup, Sachs, and Mellinger 1999). And when income levels are related to measures of market and supplier access, about 70 percent of the variance can be explained in this way (Redding and Venables 2000).

Thus, the world may not be quite the level playing field that the Lucas model supposes. Both institutional and geographic variables affect catch-up growth prospects. The chances of joining the fast growth club appear to be quite uneven. East Asia has succeeded, however. According to conventional wisdom this has been the result of good institutions and policy that have underpinned high rates of capital accumulation and strong productivity performance (World Bank 1993). What does a geographic perspective have to add?

7.5.2 The Spread of Industry

At the heart of East Asian success has been prowess in manufacturing. In section 7.2 we established the growing spread of industry out of established centers and pointed out how exceptional East Asian performance had been, especially in growth of manufactured exports and production.

Conventional treatments suggest that Japan and then the Tigers and China established institutions and policies that were conducive to strong investment in both human and physical capital and facilitated technology transfer. In most cases "developmental states" were involved in jump-starting the development process and in creating institutions that lowered transactions costs in imperfect markets, thus implementing a Gerschenkronian escape from economic backwardness (Crafts 2002b). In particular, these economies were committed to an outwardly oriented growth process in which competition to succeed in world markets held rent-seeking in check (World Bank 1993).

These arguments are well taken, but they are not the whole story. As the development process evolved, aspects consistent with what would be expected from the geographical approach previously outlined play an increasingly important role. If we add many countries to the Krugman-Venables story of figure 7.1, the approach predicts that during phase III convergence will not be uniform but will instead take the form of countries, in sequence, making a relatively rapid transit from the "poor club" to the "rich club." For example, Puga and Venables (1996) modeled a situation of a large number of identical countries, with manufacturing initially agglomerated in just one of them. They considered a steady (exogenous) growth in demand for manufactures, which had the effect of bidding up the wage in the country with the agglomeration. At some point the wage gap between this country and others becomes too large to be sustainable, and industry starts to move to other countries. However, moving to all other countries is unstable, as in figures 7.8 and 7.9; if one country gets just slightly ahead, then cumulative causation causes this one to take off and the others to fall

back. The model therefore predicts rapid transit by one country from the poor club to the rich club. Continuing demand growth (as well as falling transport costs) then makes the cycle repeat itself: Industry once again spills out, and another country makes a rapid transit to the rich club. The model was intended to be suggestive of the industrialization experience in Asia, which is illustrated in figure 7.10. The vertical axis of this figure is the share of manufacturing in GDP in selected Asian countries, and the story is very much as predicted by the theory.

In this perspective, the initial success of Japan adds to the development prospects of the rest of East Asia much as the so-called "flying geese" model suggests. As Japanese wage costs rose, particularly from the 1980s, domestic manufacturing investment was discouraged and FDI flowed out to other parts of the region; "hollowing out" of Japanese industry became a noticeable feature (Cowling and Tomlinson 2000). In 1991–95 Japanese investment in Asian manufacturing totalled $22.9 billion, compared with $7.6 billion in 1951–85 (Legewie 1999). East Asian wage costs typically were low relative both to other parts of the world and to the labor productivity gap with the established centers. Thus, even in the mid–1980s labor costs in Korea and Taiwan were only around 10 percent of the American (20 percent of the Japanese level; Jacobs 2000) when manufacturing labor productivity was close to 20 percent of the American level (Timmer 1999) and at a time when African wages were, in most cases, 15 to 25 percent of those in the United States.

The costs of regional transactions fell sharply. As one recent survey put it, both the hardware and software of East Asian linkages improved rapidly; better communications and networks of overseas Chinese both played an important part (Petri 1995). As new economic geography models predict, clustering became apparent (e.g., microcomputers in Taiwan, electronics in Malaysia), and external economies of scale from agglomeration accrued. In a detailed study of Korea, Henderson, Lee, and Lee (2001) found that in the period 1983–93 these were comparable to U.S. experience and that a doubling of an industry's size in a particular city implied a productivity increase of about 6 percent. Japanese FDI has tended to cluster, which also underlines the importance of agglomeration benefits (Head, Ries, and Swenson 1995).

The manufacturing export performance of the newly industrializing Asian countries has been impressive. It cannot be explained, however, within the confines of the traditional Hecksher-Ohlin model, as Lall (1998) showed. The pattern of exports with its strong achievement in high-technology sectors has relied on external economies of scale which accrue from learning and labor market pooling. Undoubtedly, this has been the result in part of good policy, in terms of addressing market failures (for example) through public-private partnerships although not through intersectoral transfers of resources (Crafts 1999); but it seems clear that intra-

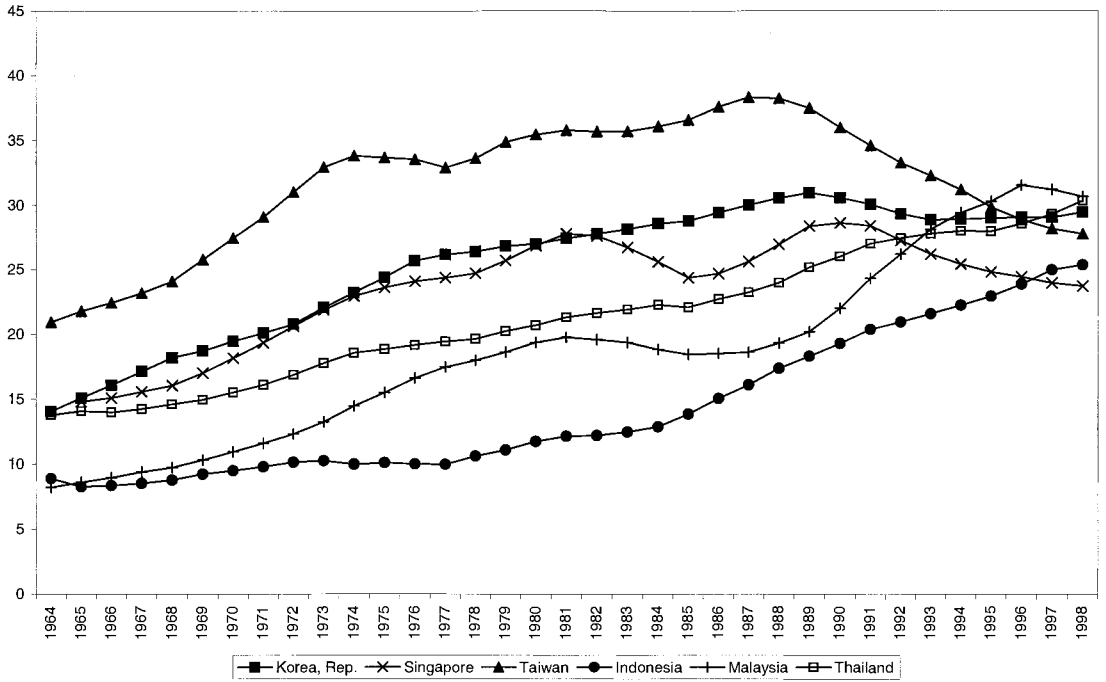

Fig. 7.10 Manufacturing VA share of GDP, five-year moving average
Source: World Bank (2001).

sectoral productivity improvement based on agglomeration benefits has been central to the achievement (Hobday 1995).

7.5.3 New Technologies: The Death of Distance?

We conclude with some remarks about the implications of new technologies for the likely future economic geography of the world. Although the continuing spread of industry—through lower trade costs, easier remote management, and the development of production networks—seems likely, what are the prospects for a more radical change, a "death of distance" whereby technologies will enable suitably qualified countries to operate "as though geography has no meaning" (Cairncross 2001)? Evidently, some activities can now be fully digitized and located and transmitted around the world at essentially zero cost. The best examples are the ICT-enabled services, such as transcription of medical records or cartoon film drawings, that now (with other software services) make up around 10 percent of India's total exports. Although these activities have brought prosperity to some developing regions, it seems unlikely that more than a few percent of world GDP is likely to fall in this category. Activities that become digitizable fall rapidly in price and are likely to be fully automated in time (e.g., by voice recognition software or computer graphics).

Other activities remain tied to markets, or retain a propensity to cluster. Some technical changes seem to increase the value of proximity—for example, the improved stock control and information flows that support just-in-time technologies, and the consequent clustering of suppliers around assembly plants. The importance of access to pools of skilled or specialist labor (one of Marshall's agglomeration forces) is not likely to be significantly diminished, nor the mobility of this labor significantly increased, by new technologies. Indeed, the microfoundations of agglomeration economies for U.S. manufacturing industries now, unlike in the nineteenth century, seem to rest more than any other single factor on labor market pooling (Rosenthal and Strange 2001). The role of cities in reducing transport costs for goods is much less important now, whereas the role that cities play in eliminating the distance between people appears to be growing as time costs increase (Glaeser 1998).

It is quite probable that the growing use of e-commerce will in some cases increase the attractiveness of market transactions relative to vertically integrated production; that is, it will change the optimal boundaries of the firm—for example, by raising the effective number of suppliers of intermediates and reducing the scope for opportunism in the presence of asset specificity. Thus, the classic historical example of vertical integration, the takeover of Fisher Body by General Motors in the 1920s, was reversed in 1999 when the parts division of General Motors became a separate company, Delphi Automotive Systems (Lucking-Reiley and Spulber 2001).

But in an increasingly information-based economy, much information

seems to be too complex to be codifiable, so face-to-face contact remains important. Sometimes this is because of the inherent complexity of the information (as in R&D and coauthorship), and sometimes because of incomplete contracting, which requires face-to-face contact for monitoring and for building trust (Leamer and Storper 2001). Indeed, the possibility of spatially separating these activities from more routine parts of the supply process will likely enhance their concentration. For example, in financial services, once the backroom operations can be separated from the front room, then the agglomeration forces on the latter become overwhelming and the attractions of London, New York, and Tokyo are no longer diluted by the expense of office space for clerical activities.

What this suggests, then, is that agglomeration forces—and consequent inequalities in the location of activity and in income levels—are likely to remain important. However, the basis of these forces might change, from the nineteenth-century model of high transport costs for goods and agglomeration of heavy industries, to the twenty-first-century model of agglomeration of information-based activities—in finance, R&D, and entertainment—in cities with pools of highly specialized labor.

7.6 Conclusions

Our aim in this paper has been to show that placing the economic history of the past two centuries in a geographical perspective can add to our understanding of the past experience of economic development and thus to future prospects for income convergence following a long phase of divergence. We have argued that agglomeration has mattered a great deal and will continue to be important. Modeling of the shifting international location of industry is enhanced by including scale economies and linkage effects.

A historical episode in which this approach offers major new insights is the performance of the United States relative both to Great Britain and to Latin America. This cannot readily be encompassed in a traditional neoclassical framework and is not wholly explained by the quality of American institutions. Our analysis highlights the roles played by migration and tariff policies in promoting the industrialization of the United States.

For today's world, recognizing the importance of agglomeration in economic development implies that size and location will continue strongly to influence future relative income levels. Predictions of the death of distance in the new economy based on ICT are premature. However, there are several distinct sources of agglomeration benefits, and in the twenty-first century these are likely to revolve much more around complexities of information and pools of skilled labor than the costs of transporting manufactured goods.

With regard to the issues of catch-up and convergence in economic

growth, we have stressed that both the neoclassical growth and the new institutional economic history schools are missing an important dimension. We do not share the optimism of Lucas (2000) that the present century will be one in which international income inequality is eliminated, and we believe that current fashions in economic history are in danger of exaggerating the role played by institutional quality in development outcomes. A geographical perspective suggests that in economic development the playing field is far from level and that recognition of this casts a different light on both past performance and future prospects.

Obviously, this paper is no more than a preliminary analysis. Nevertheless, we hope that it may help to establish a research agenda that can enrich the study of comparative economic development in an imperfectly integrated world economy.

Appendix

The Three-Country Model

There are three countries, and country-specific variables are denoted by subscripts. The two sectors are manufacturing and agriculture, indicated by superscripts. Proportion μ of consumers' expenditure goes on manufacturing, the remainder on agriculture. Within each sector there are differentiated products, and demands are derived from a subutility function (or price index) taking the forms

$$G_j^M = \left[\sum_i n_i^M (p_i^M t^M)^{1-\sigma^M}\right]^{1/(1-\sigma^M)}, \qquad G_j^A = \left[\sum_i (p_i^A t^A)^{1-\sigma^A}\right]^{1/(1-\sigma^A)}$$

where n_i^M is the number of industrial products and each country produces a single agricultural variety; p_i^M and p_i^A are the prices; and t^M and t^A the iceberg transport costs. Values of demand for a product produced in country i and sold in country j are

$$p_i^M x_{ij}^M = \left(\frac{p_i^M t^M}{G_j^M}\right)^{1-\sigma^M} E_j^M, \qquad p_i^A x_{ij}^A = \left(\frac{p_i^A t^A}{G_j^A}\right)^{1-\sigma^A} E_j^A$$

Agriculture is produced by a Cobb-Douglas production function using each country's endowment of land (set equal to 1), and an amount of labor. The labor share in the production function is θ. If L_i and L_i^M denote the total labor force and labor employed in manufacturing, respectively, then agricultural output and the wage are

$$x_i^A = (L_i - L_i^M)^\theta, \qquad w_i = \theta(L_i - L_i^M)^{\theta-1}$$

Manufacturing firms use labor and manufacturing to produce output, with manufacturing share γ. They therefore have price equal to

$$p_i^M = (w_i)^{1-\alpha}(G_i^M)^{\alpha}.$$

Input demands can be found by Shephard's lemma. In the perfect competition case n_i^M is exogenous, and quantities of each variety are given by demand. In the monopolistic competition case n_i^M is endogenous and adjusts so that each firm makes zero profits. Given an increasing returns-to-scale technology, this occurs when it reaches a certain level of output, x, so we have the additional equation

$$x = \sum_j x_{ij}^M.$$

Expenditure levels come from income (wage income and agricultural rent) and from derived demands, according to

$$E_i^M = \mu[w_i l_i^M + p_i^A(L_i - L_i^M)^{\theta}] + \alpha n_i^M p_i^M x,$$
$$E_i^A = (1 - \mu)[w_i l_i^M + p_i^A(L_i - L_i^M)^{\theta}].$$

Real wages are nominal wages deflated by the price index, $(G_i^A)^{1-\mu}(G_i^M)^{\mu}$.

Results are presented for $\theta = 0.1$, $\mu = 0.3$ and $\alpha = 0.35$, and $\sigma^A = 20$ and $\sigma^M = 20$ (perfect competition) or $\sigma^M = 6$ (monopolistic competition). Trade costs in the two sectors are equal, and the high level of the iceberg factor is 1.8, and the low level 1.475.

References

Abramovitz, Moses, and Paul David. 1996. Convergence and delayed catch-up: Productivity leadership and the waning of American exceptionalism. In *The mosaic of economic growth,* ed. Ralph Landau, Timothy Taylor, and Gavin Wright, 21–62. Stanford: Stanford University Press.

Ades, Alberto F., and Edward L. Glaeser. 1999. Evidence on growth, increasing returns, and the extent of the market. *Quarterly Journal of Economics* 114:1025–45.

Aghion, Philippe, and Peter Howitt. 1998. *Endogenous growth theory.* Cambridge: MIT Press.

Bairoch, Paul. 1982. International industrialization levels from 1750 to 1980. *Journal of European Economic History* 11:269–331.

———. 1988. *Cities and economic development: From the dawn of history to the present.* Chicago: University of Chicago Press.

———. 1990. The impact of crop yields, agricultural productivity, and transport costs on urban growth between 1800 and 1910. In *Urbanization in history,* ed. A. D. van der Woude, A. Hayami, and J. de Vries, 134–51. Oxford, U.K.: Clarendon Press.

Bairoch, Paul, and Gary Goertz. 1986. Factors of urbanization in the nineteenth century developed countries: A descriptive and econometric analysis. *Urban Studies* 23:285–305.

Barro, Robert J., and Xavier Sala-i-Martin. 1995. *Economic growth.* New York: McGraw-Hill.

Broadberry, Stephen N., and Andrew Marrison. 2002. External economies of scale in the Lancashire cotton industry, 1900–1950. *Economic History Review* 55:51–77.

Cain, Louis P., and Donald G. Paterson. 1986. Biased technical change, scale, and factor substitution in American industry, 1850–1919. *Journal of Economic History* 46:153–64.

Cairncross, Frances. 2001. *The death of distance.* Boston: Harvard Business School Press.

Chandler, Alfred D. 1977. *The visible hand.* Cambridge, Mass.: Belknap.

Chiswick, Barry, and Timothy Hatton. 2001. International migration and the integration of labor markets. Paper presented at the National Bureau of Economic Research conference Globalization in Historical Perspective, 4–5 May, Santa Barbara, Calif.

Clemens, Michael A., and Jeffrey G. Williamson. 2001. A tariff-growth paradox? Protection's impact the world around 1875–1997. NBER Working Paper no. 8459. Cambridge, Mass.: National Bureau of Economic Research, September.

Coatsworth, John H. 1998. Economic and institutional trajectories in nineteenth century Latin America. In *Latin America and the world economy since 1800,* ed. John H. Coatsworth and Alan M. Taylor, 23–54. Cambridge: Harvard University Press.

Cowling, Keith, and Philip R. Tomlinson. 2000. The Japanese crisis: A case of strategic failure. *Economic Journal* 110:F358–F381.

Crafts, Nicholas. 1999. East Asian growth before and after the crisis. *IMF Staff Papers* 46:139–66.

———. 2002a. *Britain's relative economic decline 1870–1999.* London: Institute of Economic Affairs.

———. 2002b. The East Asian escape from backwardness: Retrospect and prospect. In *Economic challenges of the twenty first century in historical perspective,* ed. Paul David and Mark Thomas. Oxford: Oxford University Press (forthcoming).

Crafts, Nicholas F. R., and C. Knick Harley. 2002. Precocious British industrialization: A general equilibrium perspective. Working Papers in Economic History no. 67. London: London School of Economics, August.

Crafts, Nicholas F. R., and Anthony J. Venables. 2001. Globalization and geography in historical perspective. Paper presented at the National Bureau of Economic Research Globalization in Historical Perspective. 4–5 May, Santa Barbara, Calif.

Daly, M., and H. Kuwahara. 1998. The impact of the Uruguay Round on tariff and non-tariff barriers to trade. *The World Economy* 21:207–34.

David, Paul A. 1989. The Marshallian dynamics of industrial localization: Chicago, 1850–1890. Stanford University. Mimeograph.

David, Paul A., and Gavin Wright. 1997. Increasing returns and the genesis of American resource abundance. *Industrial and Corporate Change* 6:203–45.

Davis, Donald R., and David E. Weinstein. 1997. Does economic geography matter for international specialisation? NBER Working Paper no. 5706 (revision). Cambridge, Mass.: National Bureau of Economic Research.

Di Mauro, Francesca. 2000. The impact of economic integration on FDI and exports: A gravity approach. CEPS Working Document no. 156. Brussels: Center for European Policy Studies.

Dollar, David. 2001. Globalization, inequality, and poverty since 1980. Washington, D.C.: World Bank. Mimeograph.

Dowrick, Steven, and J. Bradford DeLong. 2001. Globalization and convergence.

Paper presented at the National Bureau of Economic Research conference Globalization in Historical Perspective, 4–5 May, Santa Barbara, Calif.
Easterly, William. 2001. The lost decades: Developing countries' stagnation in spite of policy reform, 1980–1998. *Journal of Economic Growth* 6:135–67.
Engerman, Stanley L., and Kenneth L. Sokoloff. 1997. Factor endowments, institutions, and differential paths of growth among new world economies. In *How Latin American fell behind,* ed. Stephen Haber, 260–304. Stanford, Calif.: Stanford University Press.
Ethier, Wilfred J. 1979. Internationally decreasing costs and world trade. *Journal of International Economics* 9:1–24.
Ferenczi, Imre, and Walter F. Willcox. 1929. *International migrations.* New York: National Bureau of Economic Research.
Findlay, Ronald, and Kevin O'Rourke. 2001. Commodity market integration, 1500–2000. Paper presented at the National Bureau of Economic Research Globalization in Historical Perspective, 4–5 May, Santa Barbara, Calif.
Fujita, Masahisa. 1989. *Urban economic theory: Land use and city size.* Cambridge: Cambridge University Press.
Fujita, Masahisa, Paul R. Krugman, and Anthony J. Venables. 1999. *The spatial economy: Cities, regions, and international trade.* Cambridge: MIT Press.
Gallup, John L., Jeffrey D. Sachs, and Andrew D. Mellinger. 1999. Geography and economic development. *International Regional Science Review* 22:179–232.
Glaeser, Edward L. 1998. Are cities dying? *Journal of Economic Perspectives* 12(2): 139–60.
Gordon, Margaret. 1941. *Barriers to world trade.* New York: Macmillan.
Harley, C. Knick. 1988. Ocean freight rates and productivity, 1740–1913: The primacy of mechanical invention reaffirmed. *Journal of Economic History* 48:851–76.
Harley, C. Knick, and Nicholas F. R. Crafts. 2000. Simulating the two views of the Industrial Revolution. *Journal of Economic History* 60:819–41.
Head, Keith, John Ries, and Deborah Swenson. 1995. Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. *Journal of International Economics* 38:223–47.
Henderson, J. Vernon. 1974. The sizes and types of cities. *American Economic Review* 64:640–56.
Henderson, Vernon, Todd Lee, and Yung Joon Lee. 2001. Scale externalities in Korea. *Journal of Urban Economics* 49:479–504.
Hobday, Michael. 1995. *Innovation in East Asia.* Cheltenham, U.K.: Edward Elgar.
Hohenberg, Paul M., and Lynn H. Lees. 1985. *The making of urban Europe.* Cambridge: Harvard University Press.
Hummels, David J. 2000. Time as a trade barrier. Purdue University. Mimeograph.
Hummels, David J., Jun Ishii, and Kei-Mu Yi. 2001. The nature and growth of vertical specialization in world trade. *Journal of International Economics* 54:75–96.
Hutchinson, William K. 2000. United States international trade: Proximate causes, 1870–1910. Paper presented at the American Social Science Association meetings, 7–9 January, Boston.
Irwin, Douglas A. 2000. Tariffs and growth in late nineteenth century America. NBER Working Paper no. 7639. Cambridge, Mass.: National Bureau of Economic Research.
Isserlis, Leon. 1938. Tramp shipping cargoes and freights. *Journal of the Royal Statistical Society* 101:53–146.
Jacobs, Eva. 2000. *Handbook of U.S. labor statistics.* Lanham, Md.: Bernan.
Jacobs, Jane. 1969. *The economy of cities.* New York: Vintage Books.

James, John A. 1983. Structural change in American manufacturing, 1850–1890. *Journal of Economic History* 43:433–59.
Jones, Geoffrey, and H. G. Schroter. 1983. Continental European multinationals, 1850–1992. In *The rise of multinationals in continental Europe,* ed. Geoffrey Jones and H. G. Scroter, 3–27. Aldershot, U.K.: Edward Elgar.
Kaufmann, Daniel, Aart Kraay, and Pablo Zoido-Lobaton. 1999. Governance matters. World Bank Policy Research Working Paper no. 2196. Washington, D.C.: World Bank.
Keller, Wolfgang. 2000. Geographic location of international technology diffusion. *American Economic Review* 92:120–42.
Kim, Sukkoo. 2000. Urban development in the United States, 1690–1990. *Southern Economic Journal* 66:855–80.
Knack, Stephen, and Philip Keefer. 1995. Institutions and economic performance: Cross-country tests using alternative institutional measures. *Economics and Politics* 7:207–27.
Krugman, Paul R. 1991a. *Geography and trade.* Cambridge: MIT Press.
———. 1991b. Increasing returns and economic geography. *Journal of Political Economy* 49:137–50.
Krugman, Paul R., and Anthony J. Venables. 1995. Globalization and the inequality of Nations. *Quarterly Journal of Economics* 110:857–80.
Lall, Sanjaya. 1998. Exports of manufactures by developing countries: Emerging patterns of trade and location. *Oxford Review of Economic Policy* 14 (2): 54–73.
Leamer, Edward E., and Michael Storper. 2001. The economic geography of the Internet age. NBER Working Paper no. 8450. Cambridge, Mass.: National Bureau of Economic Research.
Legewie, Jochen. 1999. The impact of FDI on domestic employment: The phenomenon of industrial hollowing out in Japan. In *Foreign direct investment and the global economy,* ed. Nicholas A. Phelps, 179–99. London: Stationery Office.
Lindert, Peter H., and Jeffrey G. Williamson. 2001. Does globalization make the world more unequal? Paper presented at the National Bureau of Economic Research conference Globalization in Historical Perspective, 4–5 May, Santa Barbara, Calif.
Losch, A. 1954. Reprint. *The economics of location.* English translation. New Haven, Conn.: Yale University Press. Original edition, Jena, Germany, 1940.
Lucas, Robert E. 2000. Some macroeconomics for the twenty-first century. *Journal of Economic Perspectives* 14 (1): 159–68.
Lucking-Reiley, David, and Daniel F. Spulber. 2001. Business-to-business electronic commerce. *Journal of Economic Perspectives* 14 (4): 3–22.
Maddison, Angus. 2001. *The world economy: A millennial perspective.* Paris: Organization for Economic Cooperation and Development.
Markusen, James R. 1983. Factor movements and commodity trade as complements. *Journal of International Economics* 13:341–56.
Marshall, Alfred. 1920. *Principles of economics.* 1890. 8th ed., London: Macmillan.
North, Douglass C. 1990. *Institutions, institutional change, and economic performance.* Cambridge: Cambridge University Press.
North, Douglass C., William R. Summerhill, and Barry R. Weingast. 2000. Order, disorder, and economic change: Latin America vs. North America. In *Governing for prosperity,* ed. Bruce Bueno de Mesquita and Hilton L. Root, 17–58. New Haven, Conn.: Yale University Press.
Obstfeld, Maurice, and Alan M. Taylor. 2001. Globalization and capital markets. Paper presented at the National Bureau of Economic Research conference Globalization in Historical Perspective, 4–5 May, Santa Barbara, Calif.
O'Rourke, Kevin. 1996. Trade, migration, and convergence: An historical perspec-

tive. CEPR Discussion Paper no. 1319. London: Center for Economic Policy Research.
O'Rourke, Kevin, and Jeffrey G. Williamson. 1994. Late nineteenth century Anglo-American factor-price convergence: Were Hecksher and Ohlin right? *Journal of Economic History* 54:892–916.
Perloff, Harvey S., Edgar S. Dunn Jr., Eric E. Lampard, and Richard F. Muth. 1960. *Regions, resources, and economic growth.* Baltimore: Johns Hopkins University Press.
Petri, Peter A. 1995. The interdependence of trade and investment in the Pacific. In *Corporate links and foreign direct investment in Asia and the Pacific,* ed. Edward K. Y. Chen and Peter Drysdale, 29–55. New York: Harper Educational.
Portes, Richard, and Helene Rey. 1999. The determinants of cross-border equity flows. NBER Working Paper no. 7336. Cambridge, Mass.: National Bureau of Economic Research, September.
Pred, Allan C. 1977. *City-systems in advanced economies.* New York: Wiley.
Pritchett, Lant. 1997. Divergence, big time. *Journal of Economic Perspectives* 11 (3): 3–17.
Puga, Diego. 1998. Urbanisation patterns: European vs. less developed countries. *Journal of Regional Science* 38:231–52.
Puga, Diego, and Anthony J. Venables. 1996. The spread of industry: Agglomeration in economic development. *Journal of the Japanese and International Economies* 10:440–64.
Quah, Danny T. 1997. Empirics for growth and distribution: Stratification, polarization, and convergence clubs. *Journal of Economic Growth* 2:27–60.
Redding, Stephen J., and Anthony J. Venables. 2000. Economic geography and international inequality. CEPR Discussion Paper no. 2568. London: Center for Economic Policy Research.
Rosenthal, Stuart S., and William C. Strange. 2001. The determinants of agglomeration. *Journal of Urban Economics* 50:191–229.
Simon, Curtis J., and Clark Nardinelli. 1996. The talk of the town: Human capital, information, and the growth of English cities, 1861 to 1961. *Explorations in Economic History* 33:384–413.
Simon, Matthew. 1968. The pattern of New British portfolio foreign investment, 1865–1914. In *The export of capital from Britain 1870–1914,* ed. A. R. Hall, 15–44. London: Methuen.
Taylor, Alan M., and Jeffrey G. Williamson. 1997. Convergence in the age of mass migration. *European Review of Economic History* 1:27–63.
Timmer, Marcel. 1999. *The dynamics of Asian manufacturing.* Eindhoven, The Netherlands: Centre for Innovation Studies.
Turnbull, Gerard L. 1987. Canals, coal, and regional growth during the Industrial Revolution. *Economic History Review* 40:537–60.
United Nations. 1965. *The growth of world industry, 1938–1961.* New York: UN.
United Nations Conference on Trade and Development. 1983. *Handbook of international trade and development statistics.* New York: UN.
———. 2000. *UNCTAD handbook of statistics.* New York: UN.
United Nations Industrial Development Organization. 2001. *International yearbook of industrial statistics.* Cheltenham, U.K.: Edward Elgar.
Venables, Anthony J. 1996. Equilibrium locations of vertically linked industries. *International Economic Review* 37:341–59.
———. 1999. Trade liberalization and factor mobility: An overview. In *Migration: The controversies and the evidence,* ed. Riccardo Faini, Jaime de Melo, and Klaus F. Zimmerman, 23–48. Cambridge: Cambridge University Press.
Williamson, Jeffrey G. 1995. The evolution of global labor markets since 1830:

Background evidence and hypotheses. *Explorations in Economic History* 32:141–96.
———. 1998. Real wages and relative factor prices in the third world before 1940: What do they tell us about the sources of growth? Discussion Paper no. 1855. Harvard Institute of Economic Research.
World Bank. 1993. *The East Asian miracle.* Washington, D.C.: World Bank.
———. 2000. *World investment report.* Washington, D.C.: World Bank.
———. 2001. *World development indicators.* Washington, D.C.: World Bank.
Wright, Gavin. 1990. The origins of American industrial success. *American Economic Review* 80:651–68.
Yates, P. Lamartine. 1959. *Forty years of foreign trade.* London: Allen and Unwin.
Yeats, Alexander J. 1998. Just how big is global production sharing? Policy Research Paper no. 1871. Washington, D.C.: World Bank.

Comment

Richard E. Baldwin

Linking economic geography to the history of globalization is a good and an old idea. Kuznets, Myrdal, and Rostow gave geography and agglomeration forces center stage in their theories, but they lacked tools to formalize the links.[1] These were provided in the early 1990s with the emergence of the so-called new economic geography. This paper by Crafts and Venables applies this new approach to several aspects of the historical experience of globalization. Although this effort is welcome, the outcome falls short of what one could have hoped for given the towering contributions of the authors to economic history on the one hand and economic geography on the other.

In the mid-1990s, Paul Krugman and Tony Venables *formally* showed that "new" economic geography models provided a sweeping account of how falling trade costs and agglomeration forces could explain the broad outline of two centuries of global economic developments (the working title of their paper was "History of the World: Part I"). The basic logic of the Krugman-Venables story sparked great interest, but many found its mapping to historical data disturbingly vague. Crafts and Venables would be the perfect team to shore up its historical underpinnings, but here their paper disappoints; what we get are a few cursory and well-known facts on transport costs and, separately, on global economic activity since the early nineteenth century.

Moreover, the authors focus a good deal of attention on the way economic geography explains how the United States came to overtake other regions in terms of income and output. They note that "following the relative

Richard E. Baldwin is professor of international economics at the Graduate Institute of International Studies, Geneva, and a research associate of the National Bureau of Economic Research.

1. For example, when it comes Rostow's famous "growth take-off," he writes that it "may come about through a technological (including transport) innovation which sets in motion a chain of secondary expansion in modern sectors and has powerful potential external economy effects which the society exploits" (1960, 36).

and absolute decline of the Asian economies, the other main change in the economic geography of the nineteenth-century world was the rise of the New World, and within this the particular dominance of the United States." This focus on how the United States shouldered the United Kingdom out of the number one spot, combined with a resolute disregard of continental Europe, seems to reflect a very British view of the world.

The nineteenth-century rise of the United States' economy was caused by, and indeed could not have happened without, a growth take-off, so the issue the authors are addressing is really the timing of growth take-offs. Geography clearly has a role to play in explaining the spread of modern growth, but the U.S. experience must be seen in a broader context. The emergence of modern growth, which is intimately associated with industrialization, first occurred in Great Britain in the late eighteenth century (Crafts 1995). After a significant delay of a few decades, modern growth began in a series of other economies—all of which are close to the United Kingdom in terms of transport costs. Belgium, France, and the United States experienced industrialization and modern growth in the 1830s and 1840s, with Germany, Sweden, Austria, Italy, Russia, and Canada joining in the mid- to late 1800s. Keeping in mind that sea transport was much more efficient than land transport for much of this period, the centrality of geography in this nineteenth-century saga just leaps off the page. The authors' concentration on why the United States rather than Latin America became the area that overtook the United Kingdom looks distinctly odd. Indeed, Puga and Venables (1996) have a nice model of the spread of industrialization, which could have—with the straightforward addition of knowledge spillovers and endogenous growth—been usefully applied to study the role of geography in this key growth/industrialization question. This is just one example of the contributions that this pair of authors could have made.

Grand Unified Theory of Globalization and Geography

Perhaps the biggest missed opportunity in this paper concerns what might be called the grand unified theory of globalization and geography. All the elements are lying around in the literature; Crafts and Venables would have been the perfect pair to put them together. Before sketching out what this might have looked like, allow me to stylize the main facts of globalization since the mid-nineteenth century.

Globalization's Five Famous Facts

The world has seen two waves of globalization: one from roughly 1850 to 1914, and one from the 1960s to the present. At a high level of abstraction the key facts are as follows:[2]

2. This section draws on Baldwin and Martin (2000), available as an NBER working paper or at [http://heiwww.unige.ch/~baldwin/].

- *Industrialization/deindustrialization.* In the first wave, the "North" (Western Europe and the United States) industrialized while the South (especially India and China) deindustrialized. In the second wave, the South (East Asia) industrializes while the North deindustrializes.
- *Divergence.* The first wave saw North and South incomes diverge massively, whereas the second wave witnesses a convergence, at least between the North and the industrializing South.
- *Trade.* International trade in goods and factors (labor migration and long-term capital flows) exploded in the first wave. After being shut down by two world wars, a surge of protectionism, and the Great Depression, the second wave has been marked by a return of trade and capital flows to levels that have recently topped those seen in Victorian England. Mass international migration, however, remains small by the standards of the first wave.
- *Growth take-off.* Some time before the first globalization wave kicked in, the Industrial Revolution triggered modern growth in the North, but the South continued to stagnate in per capita terms. Modern growth—that is, a self-sustained growth process whereby output per hour worked rises steadily year by year—began in the United Kingdom but spreads to western Europe and the United States around the middle of the nineteenth century. Of course, this is not independent of the income divergence since big differences in income levels come from sustained differences in growth rates—not from one-time shifts of the location of industry. Moreover, the limited income convergence in the second wave is linked to spectacular growth in the industrializing South and a moderate slowdown in the North.
- *Urbanization.* Whereas some of the largest cities in the world were in the South prior to the nineteenth century, the first globalization wave is accompanied by a rapid and historically unprecedented urbanization in the North. Northern urbanization continued during the second wave, but cities grew even more rapidly in the South.

Accounting for the Facts: Economic Geography's Say

The Krugman-Venables history of the world sews together the first three of the five facts. Here it is. In 1750 or so, the world's economic geography was quite homogenous—that is, poor and agrarian. With domestic and international trade costs nearly prohibitive, each village essentially had to make all its own goods; this meant manufactured goods were dear and the available range of varieties limited. As trade costs fell, both inside and between nations, specialization became feasible, and this triggered a process of what Myrdal (1939) called "cumulative causality." Modeling this circular-causality process is the heart of the new economic geography's contribution, so an aside is in order. Migration of firms or workers dehomogenizes the world, turning it into economically big and small regions (markets). When industries are imperfectly competitive and trade is costly,

Krugman's "home market effect" favors the location of industry in large regions, but since industries are marked by increasing returns, getting a disproportionate share of industry means a region's labor is disproportionately productive, and this in turn results in higher real wages, a higher return to capital, or both. The circle is closed by noting that capital and labor are attracted to the region with higher rewards, and their migration makes the big region bigger and the small region smaller.

According to Krugman and Venables, advances in transport technology in the early nineteenth century triggered this dehomogenization of the world's economic geography, and, as history would have it, the North won at the South's expense. This single event is the root cause of the first three facts: Northern industrialization and Southern deindustrialization, the rapid expansion of international trade (England becomes the world's workshop, providing cheap and varied manufactured goods in exchange for raw materials, and this specialization both fosters trade and is fostered by it), and income divergence (due to increasing returns in industry and decreasing returns in other sectors, a high share of industry in GDP means high labor productivity and thus high incomes).

One problem with this story is that the magnitudes just do not fit. One-time concentrations of industry just cannot account for the observed income gaps. Here is the argument. Krugman and Venables ignore endogenous technological progress, assuming that physical technology is identical in the North and South. Thus, in the Krugman-Venables story, the difference in incomes between the United Kingdom and India must be due to the difference in industry's share in the U.K. and Indian output mix and the productivity gap between industry and traditional sectors. If the United Kingdom's per capita income was 100 in 1850, India's was 23 according to Maddison (1995, tables C16 and D1), so the income gap to be explained is 77. Moreover, Crafts (1989, 417) tells us that in 1840, 47 percent of the U.K. workforce was in industry, and Bairoch (1982, table 9) tells us that India was only 4.7 percent as industrialized as the United Kingdom in 1860, so (ignoring the mismatch in dates) we can conclude that the static allocation of industry can only account for the income difference if industrial workers are 171 times (i.e., 17,100 percent) more productive than workers in the traditional sector.

This just cannot be right. Plainly, the real story must lie elsewhere, and growth is the obvious suspect. Indeed, since the headline story in the nineteenth century was the spread of modern growth, the Krugman-Venables story is a bit like Hamlet with the Prince. Clearly, one has to add endogenous growth to the Krugman-Venables story to account for the facts on income divergence and convergence as well as on growth take-offs.

Adding Endogenous Growth to Economic Geography Models

Fortunately, the literature combining economic geography and growth models is fairly well developed (see, e.g., Baldwin and Forslid 1997; Martin

and Ottaviano; 1999; and Baldwin 1999). Indeed, Baldwin, Martin, and Ottaviano (2001) have used this combination of geography and growth to account for the first four of globalization's five famous facts.

The basic idea is quite simple and turns on the fact that transporting ideas—as well as goods—is expensive, so that learning spillovers tend to be localized geographically. Starting in the Krugman-Venables phase I, where transporting goods is expensive and industry is thinly spread, growth does not occur because the dispersion of industry prohibits a virtuous learning and innovation cycle from starting. As the transport cost of goods falls, industry—and thus learning—gets geographically concentrated. Due to localized technological spillovers, industrial agglomeration in the North implies that the South has no incentive to invest and innovate, while the incentive to innovate in the North increases. In this way, industrial agglomeration not only generates industrialization and a growth take-off in the North, but it also produces a massive income divergence.

Now, as we move on to the second wave of globalization, we presume that the cost of transporting goods asymptotes toward some natural limit, but additionally, and importantly, we assume that the cost of trading ideas decreases. At some point, this generates a rapid industrialization in the South because the South is more easily able to benefit from historical innovation in the North and more easily able to access Northern markets. The emergence of southern industry slows global growth somewhat (since it disperses learning) and forces a relative deindustrialization in the North.

The only facts left unaccounted for concern urbanization. To get this into the story, one would have to allow internal geography in the regions considered (Baldwin, Martin, and Ottaviano follow Krugman and Venables in assuming that regions are just points in space), but once the technical difficulties were mastered, the economics would be straightforward. In the first wave of globalization, economic activity characterized by localized spillovers is concentrating in the North. It would not therefore be too surprising that urbanization proceeded faster in the North than in the South during this era. Likewise, in the second wave of globalization, the industrialization of the South (emergence of the Asian tigers, etc.) strengthens the forces that foster within-South concentration of economic activity (i.e., urbanization), while the deindustrialization of the North does the opposite.

These comments have, I hope, illustrated that the Crafts-Venables paper is a small first step in what should prove a very fruitful direction for research.

References

Bairoch, P. 1982. International industrialization levels from 1750 to 1980. *Journal of European Economic History* 2: 268–333.

Baldwin, R. 1999. Agglomeration and endogenous capital. *European Economic Review* 43 (2): 253–80.

Baldwin, R., and R. Forslid. 1997. The core-periphery model and endogenous growth: Stablising and de-stablising integration. *Economica* 67 (3): 307–24.

Baldwin, R., and P. Martin. 2000. Two waves of globalisation: Superficial similarities and fundamental differences. In *Globalisation and labour,* ed. H. Siebert, 3–59. Tubingen, Germany: Mohr.

Baldwin, R., P. Martin, and G. Ottaviano. 2001. Global income divergence, trade, and industrialization: The geography of growth take-offs. *Journal of Economic Growth* 6 (1): 5–37.

Crafts, N. 1989. British industrialization in an international context. *Journal of Interdisciplinary History* 19 (3): 415–28.

———. 1995. Exogenous or endogenous growth? The Industrial Revolution reconsidered. *Journal of Economic History* 55 (4): 745–72.

Krugman, P., and A. Venables. 1995. Globalisation and inequality of nations. *Quarterly Journal of Economics* 60: 857–80.

Maddison, A. 1995. *The world economy: 1820–1992.* Paris: OECD Development Centre.

Martin, P., and G. Ottaviano. 1999. Growing locations: Industry location in a model of endogenous growth. *European Economic Review* 43 (2): 281–302.

———. 2000. Growth and agglomeration. *International Economic Review* 42 (4): 947–68.

Myrdal, G. 1939. *Monetary equilibrium.* London: Hodge.

Puga, D., and A. J. Venables. 1996. The spread of industry: Spatial agglomeration in economic development. *Journal of the Japanese and International Economies* 10 (4): 440–64.

Rostow, W. W. 1960. *The stages of economic growth: A non-communist manifesto.* Cambridge: Cambridge University Press.