Regions, Globalization, Development

ALLEN J. SCOTT* and MICHAEL STORPER†

*Departments of Geography and Policy Studies, University of California Los Angeles, Los Angeles, CA 90095–1656, USA. Email: ajscott@geog.ucla.edu

†Department of Urban Planning, University of California Los Angeles/Institut d’Etudes Politiques de Paris/London School of Economics, UCLA, Los Angeles, CA 90095–1656, USA. Email: storper@ucla.edu

(Received October 2002; in revised form February 2003)

SCOTT A. J. and STORPER M. (2003) Regions, globalization, development, Reg. Studies 37, 579–593. Regional economies are synergy-laden systems of physical and relational assets, and intensifying globalization is making this situation more and not less the case. As such, regions are an essential dimension of the development process, not just in the more advanced countries but also in less-developed parts of the world. Development theorists have hitherto largely tended to overlook this critical issue in favour of an emphasis on macroeconomic considerations. At the same time, conventional theories of the relationship between urbanization and economic development have favoured the view that the former is simply an effect of the latter. To be fully general, the theory of development must incorporate the role of cities and regions as active and causal elements in the economic growth process. This argument has consequences for development policy, especially in regard to the promotion of positive agglomeration economies and the initiation of growth in poorer regions. A related policy problem concerns ways of dealing with the increase in interregional inequalities associated with contemporary globalization. Issues of economic geography are thus of major significance to development theory and practice.

Economic development Agglomeration World economy Development policy Income inequalities Regional dynamics


Développement économique Agglomération Wirtschaftliche Entwicklung Ballung
Economie mondiale Politique du développement Weltwirtschaft Entwicklungspolitik
Inégalités des revenus Dynamique régionale Einkommensungleichheiten Regionale Dynamik

DOI: 10.1080/0034340032000108697
http://www.regional-studies-associ.ac.uk
INTRODUCTION: THE MISSING ELEMENT IN DEVELOPMENT THEORY

The theory of economic development has had a long and tangled history extending from the classics of eighteenth and nineteenth century political economy, through the German historical school of the early twentieth century (above all, Schumpeter, 1912), to the many different streams of developmental ideas that were in circulation in the immediate post-war decades. Analysts in this latter period focused above all on what were then called Third World countries, and posed the development question largely in terms of the vicious circles of poverty and economic backwardness that seemed to afflict so many parts of Africa, Asia and Latin America (Prebisch, 1982; UNCTAD, 1986). As selected parts in the Third World advanced in the 1970s and 1980s, development theorists began to recognize that at least some of these areas were susceptible to significant industrialization, and they duly added the notion of ‘newly industrializing countries’ (NICs) to their theoretical repertoire. More recently, the development question has shifted in ways that reincorporate the most advanced economies into its purview. This trend is evinced most especially in the writings of the new growth theorists, with their emphasis on positive externalities as a major source of economic development (Romer, 1986, 1990; Lucas, 1988).

Notwithstanding the complexity and diversity of existing approaches to development, the vast majority of them tend to concentrate on macroeconomic variables and processes. Among more orthodox theorists, in particular, strongly recurrent themes are the virtues of economy-wide fiscal and monetary responsibility, market-opening measures, secure property rights, political stability, investments in education, and nominally democratic principles of government (see Bauer and Yamey, 1957; Balassa, 1981; Little, 1982; Krueger, 1993). A currently prominent version of the latter approach is represented by the neoliberal Washington Consensus which has decisively shaped the policies of the principal international economic institutions for the last two decades (Stiglitz, 2002).

Macroeconomic considerations are, of course, critical in any real economic development process, and we have no intention of suggesting otherwise. Nevertheless, our purpose in this paper is to point out and to deal with a silence that – with just a few exceptions – has characterized much of the development literature from the beginning. This concerns the role of selected regions as springboards of the development process in general, and as sites of the most advanced forms of economic development and innovation in particular, where by the term ‘region’ we mean any area of subnational extent that is functionally organized around some internal central pole. Development does not depend on macroeconomic phenomena alone but is also strongly shaped by processes that occur on the ground, in specific regions of the type we have just defined. As a result, development in any given country is always characterized by significant variations in the intensity and character of economic order from one place to another. Here, we are not simply calling attention to an obvious empirical state of affairs; we are also putting forward a significant clue about a complex theoretical question focused on the geographical foundations of economic growth. Any answer to this question, we argue, must consider the locational interdependencies that underpin the persistence of efficiency and innovation enhancing clusters of capital and labour in economic development. Cities and regions, in other words, are critical foundations of the development process as a whole.

Mainstream theorists have often and correctly observed that economic development has major impacts on urbanization in both rich and poor countries, though they have less frequently acknowledged causalities running in the other direction (Kuznets, 1955; Henderson, 1988). One widely-accepted view in this regard is that as development occurs, population and economic activity first become intensely polarized in any given national space, with polarization reversal then setting in as development proceeds further (Richardsen, 1980; Townroe and Keen, 1984). An associated idea is that developing countries urbanize too much and too fast, generating ‘macrocephalic’ urban systems consisting of a few abnormally large cities in each country. These cities are said to have excessively high urban densities, and their size and rapid growth result in a panopoly of economic, social and environmental problems (Lipton, 1977). Many analysts therefore hold that successful development involves the sharing out of economic activity among a greater number of smaller and more manageable urban centres (El Shaks, 1972), though this notion was more commonly ascribed a decade or two ago than it is today. By the same token, the more dispersed pattern of urbanization typical of North America and Western Europe is often taken as an essential index of the higher stages of economic development – a view, as implied by our later discussion, that is unduly restrictive in its implications.

More heterodox forms of development theory, by contrast, have long claimed that processes of development are invariably associated with uneven spatial patterns, and that this condition is actually part and parcel of the mechanism of growth. The most prominent version of this approach, represented by Hirschman, 1958, and Myrdal 1959, and their disciples, is based on the concept of circular and cumulative causation in geographic space. Extensions of this theory led to early path-breaking work on growth poles and their geographical expression as regional growth centres (Boudeville and Antoine, 1968; Perroux, 1961). Growth poles and growth centres were in turn
widely invoked in the formulation of development policies in the post-war period. In less-developed parts of the world, the same policies were often used to underpin import-substitution strategies.

In spite of the initial influence of these early heterodox claims about a positive relationship between agglomeration and development, they have tended to be relegated to the background by subsequent development theorists and the large international agencies that promote their ideas. One possible reason for this state of affairs is the justifiable recognition that urbanization in developing countries is associated with severe social costs and many kinds of technical diseconomies. A second reason is that in the past the theoretical foundations of agglomeration economies were only very partially worked out, so that their overall role in pushing development forward was widely underappreciated. A wealth of new scholarship in economic geography over the 1980s and 1990s, however, has helped to revitalize and improve upon the older heterodox approach by means of a thorough-going reconstruction of the theory of regional economic growth. This recent work makes it possible to claim effectively that agglomeration is a fundamental and ubiquitous constituent of successful development in economic systems at many different levels of GNP per capita (Bairoch, 1988; Henderson, 1988; Riveria-Batiz, 1988; Krugman, 1991; Nadvi and Schmitz, 1994; Eaton and Eckstein, 1997; Fujita et al., 1999; Scott, 2002; Storper and Venables, 2002; Fan and Scott, 2003). Accordingly, the theory that we shall seek to elaborate here puts considerable emphasis on the role of the region as a source of critical developmental assets in the form of increasing returns effects and positive externalities. In addition, we aver that because agglomeration is a principal source of these productivity-enhancing outcomes, urbanization is less to be regarded as a problem to be reversed than as an essential condition of durable development.

REGIONS IN TODAY’S WORLD ECONOMY

These questions about the geographic foundations of development and growth are made yet more urgent by the current empirical realities of globalization. It is fundamentally mistaken to equate globalization with the notion that development today involves a simple spreading out of economic activity, or the transformation of the economic order into a liquefied space of flows. On the contrary, globalization has been accompanied by the assertion and reassertion of agglomerative tendencies in many different areas of the world, in part because of the very openness and competitiveness that it ushers in (Scott, 1998; Puga and Venables, 1999). Dense regional agglomerations of economic activity are major sources of growth in economies at virtually every stage of development today, as suggested by the worldwide expansion and spread of industrial clusters.

Thus, for example, 40% of US employment is currently located in counties constituting just 1.5% of its land area. Equally, the geographical density of employment in many sectors has been on the increase of late years (Kim, 2002). It has been suggested, as well, that 380 separate clusters of firms in the US employ 57% of the total workforce and generate 61% of the nation’s output and fully 78% of its exports (Rosenfeld, 1996; OECD, 1999). Other researchers, using more conservative measures, still find that 30% of the US workforce is accounted for by globally-oriented local employment clusters (Porter, 2001). The OECD, for its part, concludes that local industrial districts account for 30% of total employment in Italy (and 43% of that country’s exports) and 30% of total employment in Holland.

The most striking forms of agglomeration in evidence today are the super-agglomerations or city-regions that have come into being all over the world in the last few decades, with their complex internal structures comprising multiple urban cores, extended suburban appendages and widely-ranging hinterlands, areas, themselves often sites of scattered urban settlements (Courchene, 2001; Hall, 2001; Scott et al., 2001). These city-regions are locomotives of the national economies within which they are situated, in that they are the sites of dense masses of interrelated economic activities that also typically have high levels of productivity by reason of their jointly-generated agglomeration economies and their innovative potentials. In many advanced countries, evidence shows that major metropolitan areas are growing faster than other areas of the national territory, even in those countries where, for a time in the 1970s, there appeared to be a turn toward a dominant pattern of non-metropolitan growth (Frey and Speare, 1988; Forstall, 1993; Summer et al., 1993). In less-developed countries, too, such as Brazil, China, India and South Korea, the effects of agglomeration on productivity are strongly apparent, and economic growth typically proceeds at an especially rapid rate in the large metropolitan regions of those countries. The same metropolitan regions are at once the most important foci of national growth and the places where export-oriented industrialization is most apt to occur (Scott, 1998, 2002).

These findings fit well with the observation that previous rounds of market opening and technological progress have tended to reinforce urbanization, not weaken it (Kim, 1995; Eaton and Eckstein, 1997; Black and Henderson, 1998; Glaeser, 1998; Puga and Venables, 1998). Recent accounts of the formation of an Atlantic economy in the late nineteenth and early twentieth centuries argue that it emerged on the basis of strong agglomeration processes in Europe and America, with the main centres of production maintaining their dominant positions.
through strong increasing returns to scale (Williamson, 1998; Crafts and Venables, 2003). Today’s wave of globalization appears to be similarly anchored in (and is also partially responsible for) an expanding intercontinental patchwork of urban and regional economic systems. In sum, large-scale agglomeration – and its counterpart, regional economic specialization – is a worldwide and historically persistent phenomenon that is intensifying greatly at the present time as a consequence of the forces unleashed by globalization. This leads us to claim that national economic development today is likely not to be less but rather more tied up with processes of geographical concentration compared with the past.

Moreover, as globalization and international economic integration have moved forward, older conceptions of the broad structure of world economic geography as comprising separate blocs (First, Second and Third Worlds), each with its own developmental dynamic, appear to be giving way to another vision. This alternative perspective seeks to build a common theoretical language about the development of regions and countries in all parts of the world, as well as about the broad architecture of the emerging world system of production and exchange. At the same time, it recognizes that territories are arrayed at different points along a vast spectrum of developmental characteristics.

These are strong claims, and they call for extended justification, which we proceed to elaborate in the next section. Meanwhile, it is worth pointing out that a long tradition of econometric analysis dating back to Kawashima, 1975; Sveikauskas, 1975; and Carlino, 1979; provides prima facie evidence in favour of our assertions. This line of work has demonstrated time and again that in the more economically-advanced countries, urban centres persistently exhibit signs of significant and positive productivity effects as a function of their size. A branch of this literature has more recently focused on less developed parts of the world, and draws broadly similar conclusions (see Henderson, 1988; Sueyoshi, 1992; Chen, 1996; Shukla, 1996; Lee and Zang, 1998; Mitra, 2000). This literature as a whole tends to adopt traditional conceptualizations of the problem in terms of so-called urbanization and localization economies, but we consider these concepts to be internally chaotic, and instead we shall seek in the following to replace them with more analytically sustainable categories. There is also a tendency in this literature to underestimate the impacts of urbanization on productivity because the parameters of the econometric models on which it is based are never calibrated over counterfactual cases of economic systems in which dense agglomerations of capital and labour are absent.

We now consider how the empirical realities we have alluded to can be accounted for by contemporary concepts of agglomeration, which we employ in turn as essential components of an updated development theory. This theory seeks to accommodate the cases of both poor and rich countries, and to shed some new light on the phenomenon of uneven spatial and economic development on a world scale.

**THE FUNDAMENTALS OF AGGLOMERATION**

**The analytical decomposition of agglomeration processes**

Cities always appear as privileged sites for economic growth because they economize on capital-intensive infrastructure (which is particularly scarce in developing areas), thus permitting significant economies of scale to be reaped at selected locations. But to this obvious basic factor underlying agglomeration, we must add three further sets of phenomena that complement and intensify its effects, namely: (1) the dynamics of backward and forward inter-linkage of firms in industrial systems; (2) the formation of dense local labour markets around multiple workplaces; and (3) the emergence of localized relational assets promoting learning and innovation effects. Some brief commentary on these points is now in order.

Even though transport and communications costs tend to decline over time, the friction of distance in general continues to have powerful effects on locational outcomes. Improvements in transport and communications processes (e.g. the development of canal systems, railroads, the interstate highway network, the postal service, or the telegraph and the telephone) have rarely if ever slowed down the urbanizing tendencies of modern capitalism, even as they have encouraged its spatial extension. Rather, improvements of these sorts have almost always tended to reinforce the clustering of economic activity both by widening the market range of any given centre and by helping to spark off new rounds of specialization in established urban areas. This state of affairs also seems to be the case in the present period in which internet-based broadband communications technologies have made possible instantaneous transmission of complex messages across the globe at extremely low cost. More accurately, we should say that many inter-firm transactions can be executed cheaply over long distances, while others resist being stretched out by reason of the high linkage costs that they involve, even in a world of rapidly improving transport and communications technologies. Small-scale, non-routine flows with ambiguous information content are notably averse to extension over long distances. This resistance is intensified where firms compete with one another by means of product differentiation, and where markets are characterized by much uncertainty. In circumstances like these, firms find it difficult to stabilize their output profiles, thus necessitating external transactional relations that are constantly shifting in size, form, and origin or destination, and that are therefore expensive per unit
of distance and time. Dense agglomerations containing large numbers of firms allow both suppliers and buyers to compensate for variability and uncertainty by providing ready access to needed resources on short notice. Considerable gains in productivity typically flow to firms from this localized concentration of many different suppliers and buyers. Among the more important of these gains is the ability to maintain low overheads while achieving high flexibility in both internal and external operations. One especially powerful phenomenon is the continuing importance of face-to-face contacts for the transmission of complex and uncertain messages (Leamer and Storper, 2001) and for the establishment of mutual confidence and accurate evaluation of potential partners in constantly changing business relationships (Storper and Venables, 2002).

Comparable dynamics of matching differentiated demands and supplies apply to labour markets. When firms need specialized workers, but are subject to rapid shifts in their product and process designs (as in the case of fashion-oriented or technologically-innovative industries), they usually strive as far as possible to achieve flexibility in their use of labour. At the same time, they seek to avoid the risk of costly delays in finding the various skills on which they depend. To overcome this problem, they need direct access to large and variegated pools of specialized talent. In the same way, if workers are to invest in building up their competencies, but are unable to secure long-term employment contracts, they will prefer to locate where there are many potential employers. In turn, rapid search and rehire processes will compensate them for high turnover. In all of these circumstances, geographical concentration has major productivity-raising effects for firms, and income-raising effects for workers. Firms benefit from the possibility of adjusting their capacity levels as needed, while minimizing the risks of not finding the workers they require for expansion and change. Workers gain by having strong incentives to invest in their own talents and becoming more specialized, but are able to offset the associated risks by being in a place where the existence of multiple employment opportunities raises their chances of finding a job (Jaye, 1983). These search-and-match processes in the local environment are carried out by means of relatively complex transactions, often operating through dense social networks (Granovetter, 1986). Geographical concentration lowers the costs of these transactions and raises the probability of successful matching for all parties.

Regional concentrations of economic activity have another advantage, which is purely dynamic in nature. There is mounting evidence that creativity and learning have a distinctive geography, with regions playing active roles as sites of continuous and informal but cumulatively significant improvements in industrial products and processes (Russo, 1985; Jaffe et al., 1993; Saxenian, 1994; Dunning, 1998; Scott, 1999; Feldman, 2000). Silicon Valley, of course, is the classic reference here, though the phenomenon of localized innovation has been observed in many other industrial clusters. The spatial proximity of large numbers of firms locked into dense networks of interaction provides the essential conditions for many-sided exchanges of information to occur, and out of which new understandings about process and product possibilities are constantly generated. Specialized regional economies are the locus of intense knowledge spillovers, thereby helping to raise the rate of innovation, and to promote long-term growth (Jaffe et al., 1993; Antonelli, 1994; Audretsch and Feldman, 1996; Nooteboom, 1999).

Each of these factors underlying geographic concentration has the effect of creating positive externalities for both firms and workers. Our account thus far actually understates the power of agglomeration in certain ways, for geographical concentrations of the sort we are describing also constitute living human communities with many additional effects on economic performance (Storper, 1997; Temple and Johnson, 1998; Woolcock, 1998). Oftentimes, clusters of firms operate as powerful socialization mechanisms, becoming veritable engines for turning out new talent through workers’ on-the-job experiences and participation in work-related networks (Grabher, 1993). Firms come together, too, in both formal and informal organizations that help to streamline their interactions, to accelerate information transfers, to build trust and reputation effects, and to promote their joint interests (Becattini, 1990; Asheim, 2000). Relationships like these contribute to the stock of collective assets in any given agglomeration. Their effects are thus frequently positive, though they may also on occasions be negative when local conditions and institutional environments induce problems such as rent-seeking behaviour or inter-organizational rivalries.

The agglomeration-development nexus

Cities are a necessary corollary of industrialization because they allow for complex agglomerations of specialized activities to emerge while economizing on infrastructure under conditions of national scarcity. In many developing countries, urban growth is pushed further forward by modernization of the agricultural sector, displacing labour, and generating large-scale migration from countryside to city (Todaro, 1969; Alonzo, 1980; Kelley and Williamson, 1984).

This, however, is at best a partial view of the dynamic properties of the relationship between urbanization and economic development. To begin with, the emphasis on infrastructure (a common theme in many discussions of development) is only one among many reasons for agglomeration. As we saw above, actual agglomerations are characterized by many additional sources of productivity gain through their transactional structures, local
labour markets, learning effects, and so on. These phenomena can sustain the advantages of agglomeration even in the face of rapidly rising costs of urban concentration due to congestion, pollution, escalating land prices, crime, family breakdown, etc. Such costs are especially high in developing countries, but still they fail to arrest urban growth (Azzoni, 1986; Storper, 1991). Another way in which many older arguments underestimate the force of geographical concentration is that they often take large-scale, capital-intensive manufacturing industries to be the privileged motors of development and growth in developing countries. As activities like these relocate outward to other regions, so – it is held – geographical polarization reversal will occur (Townroe and Keen, 1984). We now know, however, that developing countries also move ahead on the basis of many different kinds of sectors, some of which generate strong systems of externalities, and tend to be marked by forceful agglomeration and urbanization tendencies wherever they appear on the landscape. These sectors include small-scale indigenous manufacturing, low-technology industries, craft-based industries, and a wide array of services (Nadvi and Schmitz, 1994; Scott, 2002). Polarization reversal is far from being a universal characteristic of the development process.

The particular patterns of agglomeration that make their appearance in any given instance vary widely depending on local circumstances and the local mix of sectors, and this diversity is further augmented by the role that historical path dependencies play in the evolution of regional economies (Fujita et al., 1999). This is an important reason why, in fact, there are many variations in the character of urban systems in both the developing and developed countries as a whole, rather than convergence toward any single type. What is common to all is the underlying functional link between agglomeration, urbanization and development.

This link, moreover, is susceptible to self-reinforcement over time by the locational dynamics of expanding industrial systems. When a sector first comes into being in a given part of the world (a country, a continent), the firms involved in it are often located in a wide variety of places. This is because young, or emerging, or recently-implanted industries tend to be relatively independent of (or have no opportunity to tap into) pre-existing place-dependent positive externalities, especially where these have developed in relation to older sectors and hence have little specific utility for new industries. However, this first stage of development, characterized as it is by an open ‘window of locational opportunity’, is almost always followed by a second where the large initial number of locations is whittled down as the industry’s local external environment responds to growing demands for inputs of materials, services, labour, and so on, and as geographically-focused increasing returns effects come into being at selected locations (Scott and Storper, 1987). Thus, a few places begin to move ahead as their self-reinforcing concentrations of capital and labour make them progressively more efficient, in both static and dynamic terms. Success breeds success (up to some point of diminishing returns, at least), and the advantages of these places then become locked in, marginalizing competitor locations and effectively crowding them out of the field (Krugman and Obstfeld, 1991). In this manner, what begins initially as a relatively open window of locational opportunity for an industry eventually closes around a small number of core agglomerations.

The frequency and scope of windows of locational opportunity are controlled by many factors, of which internal economies of scale (in production, R&D, transacting, and so on) are especially important. In industries where this feature results in oligopolistic supply structures (e.g. sectors producing commercial aircraft or nuclear-power generators) only a few regions will be able to attract relevant investments and to acquire production capacity. Major shifts in the core locations of these industries can generally occur only when there are important technological changes in products and processes, thereby undermining the advantages of existing producers and, by extension, the regions in which they are concentrated. By contrast, in sectors where optimal scale is achieved at low rates of output (e.g. clothing, shoes, jewellery, and many kinds of electronics industries or business services) there are numerous potential windows of locational opportunity. Sectors of this sort are able to engage in significant forms of product differentiation from one place to another, thus making it possible for latecomers to enter the market and to create distinctive niches for themselves. This point is exemplified by the recent history of the global shoe industry (Gerffi, 1995; Schmitz, 2001). Once agglomeration occurs, (and depending on the nature of further major technological shifts), the locational pattern of these sectors becomes locked in, and local developmental effects intensify.

We have argued that urbanization is one of the major drivers of the process of development in the contemporary world. This argument, to be sure, is by no means novel. However, we have sought to re-express the older Hirschman–Myrdal–Perroux approach in terms of recent advances in the theory of agglomeration and economic geography, and on this basis to redress some of the imbalance that currently appears to exist between macroeconomic approaches to the development question and what we earlier alluded to as development ‘on the ground’. This attempt to achieve a more balanced perspective is not only significant in conceptual terms but also as a practical matter, for it reveals important instruments (as we shall indicate) by means of which policy makers can approach critical tasks of economic development from a bottom-up and hence locally nuanced perspective. One important corollary of our argument is that increasingly uneven
densities of spatial development can actually enhance overall rates of economic growth and are hence not necessarily always undesirable, though we shall argue that this also sometimes gives rise to an increasingly uneven spatial distribution of income, and hence to social and political predicaments that can undermine developmental programmes that fail to pay adequate attention to this circumstance. We have further suggested that the complex nexus of relationships linking urbanization and development operates in countries at every plane of GDP per capita and that economic development can be achieved on the basis of a wide variety of manufacturing and service sectors. These sectors include even simple craft- or small firm-based industries, which were once thought of as the very antithesis of any kind of durable development (Piore and Sabel, 1984; Wade, 1990). It is especially urgent to refocus attention on the developmental potential of cities and regions in the context of globalization, because they are the loci of intense positive externalities in the increasingly borderless global system of economic relationships. In less-developed countries, in particular, agglomeration is critical to development not only because it is a source of enhanced economic productivity, but also because it is a basic condition of specialization within the global division of labour and an essential foundation of export-oriented growth.

DEVELOPMENTAL DISPARITIES IN THE CONTEMPORARY WORLD SYSTEM

Regional divergence or convergence?

The increasing liberalization of economic exchange as globalization has proceeded, combined with steady improvements in technologies of transportation and communication has encouraged the world-wide spread of dense productive agglomerations (see Krugman and Venables, 1993). This effect is complemented by two others. First, agglomerations in different parts of the world find themselves increasingly caught up in relations of competition and complementarity with one another. Inter-agglomeration competition occurs when producers in different places operate on the same markets; complementarity is present when differentially specialized agglomerations are linked together via long-distance commodity chains (Feenstra and Hanson, 1996). Second, agglomerations are also often deeply connected to more peripheral, less densely developed areas, especially where certain types of production units within wider commodity chains find it advantageous to locate at decentralized sites. This phenomenon is especially characteristic of branch plant operations with relatively standardized production activities and hence with low-cost procurement and distribution structures. The net result of the two tendencies noted here is the proliferation of complex trade flows, between different agglomerations and between agglomerations and peripheral areas, at national and international scales, and these flows are expanding with globalization.

Neoclassical theories of development hold that the spatial integration of economic activity in these ways tends progressively to eliminate interregional differences in living standards, by promoting some combination of structural and compositional convergence among participating economies. In fact, the actual record is quite wayward, with convergence occurring in some places at some times, and divergence occurring on other occasions. At the present moment, the play of regional and global economic forces involves many complex cross-currents in which some parts of the world (East Asia and a few metropolitan regions of Latin America) are doing relatively well, while other parts (Africa between the tropics, much of the former Soviet Union, and certain peripheral regions in more developed countries) are falling steadily behind.

The predicaments of uneven spatial development are most dramatically expressed in the observation that 50% of global GDP today is produced by only 15% of the world’s people, most of them concentrated in the Triad nations of the North. Conversely, the poorer half of the world’s population produces just 14% of global GDP. Moreover, world trade has become more concentrated among the Triad nations, to the relative detriment of North–South commercial relations. Most of the developing world has been a relative loser in this process, again with the exception of East Asia. At the same time, much of the world’s most important trading activities (increasingly in the form of intra-firm trade) occur between a relatively limited number of subnational regions or agglomerations (Barnes and Leeburh, 1998; US Department of Commerce, 1998; Fujita et al., 1999; Andersson and Anderson, 2000; Beaverstock et al., 2000). This process accentuates the growth of selected regions, and helps to generate the contemporary phenomenon of large city-regions (as previously defined) scattered across the continents in an integrated worldwide mosaic. Many different parts of the developing world are deeply involved in relationships like these, as exemplified by city-regions such as Mexico, São Paulo, Cairo, Bombay, Kuala Lumpur, Jakarta, and so on. One consequence of this trend, however, is that interregional income inequalities within many developing countries are increasing. Indeed, even in many developed countries, the recent period of intensive globalization combined with a turn to neoliberal policy measures has been accompanied by widening gaps in per capita incomes between sub-national regions. The problem is further accentuated where labour mobility is relatively low, as it is in the UK and most of continental Europe compared to the US (Duranton and Monastiriotis, 2002).

Per capita income differences between countries
Allen J. Scott and Michael Storper

diverged over much of the nineteenth and twentieth centuries, but showed signs of convergence from the 1960s to the 1980s. Over the last decade or so, this tendency toward income convergence between countries has been reversed,2 notwithstanding the dramatic improvements in technologies of spatial interaction that have been occurring (Clark and Feenstra, 2003; Dowrick and De Long, 2003). Similarly, as Pomerantz, 2000, points out, the ‘great divergence’ of income levels in the nineteenth century occurred even though communications and transportation costs were declining rapidly.

The dynamics of differential regional development

Considerable light can be shed on these issues by further analysis of the ways that regional development processes contribute to durable structural and compositional differences between economies. In particular, why do some regions succeed in establishing high-performing economic systems while others remain stillborn, stagnate or decline even as spatial interaction costs fall?

We have already shown in our earlier discussion of windows of locational opportunity how increasing returns effects reinforce growth opportunities for regions that begin (even accidentally) to move ahead as production foci in any given sector, while progressively closing off opportunities for others. Certain endogenous features of agglomerations also have great impacts on local developmental prospects. Economic historians and geographers have shown, for example, that even in industries where best practices diffuse rapidly from country to country (as in the case of cotton mills and railroads in the nineteenth century), factor productivity is often quite uneven over space (Gertler and Digiovanna, 1997; Clark and Feenstra, 2003). What is additionally puzzling is that such differences emerge not only in cases where technologies and managerial practices are similar, but also in industries that tend uniformly to locate in large urban centres (as with much of the electronics industry today). All this implies that there are significant endogenous determinants – local and national – of how well agglomerations function, and hence how much they contribute to economic development in their local and national contexts. By the same token, increasing trade, foreign investment and the international diffusion of technology do not automatically bring about convergence in productivity and development levels (North and Thomas, 1973; Mokyr, 1983; Wade, 1990; Landes, 1998; Clark and Feenstra, 2003).

Many of the endogenous conditions underlying local economic development and facilitating entry into the world economy are cultural or institutional, in the specific sense that they entail the formation of routines of economic behaviour that potentiate and shape activities such as production, entrepreneurship and innovation (Haggard, 1990; Rodrik, 1999). These routines are, in effect, untraded forms of interdependency between economic agents, and hence they collectively constitute the relational assets of the regional economy (Storper, 1997). Standard theories of economic development do not probe adequately into these processes (Uzzi, 1996; Putnam, 2000). Neoclassical theories, including newer augmented versions, assume that successful behaviour will emerge more or less spontaneously out of the wider economic or social context (Mankiw et al., 1992). Others, like the new growth theory, put their faith in the accumulation of stocks of knowledge leading to generalized positive externality effects throughout the economy (Lucas, 1988; Romer, 1990). The latter idea, though it may be useful as a starting point, says little about the concrete habits and relationships through which knowledge and savoir faire are created and deployed in economic action (Rosenberg, 1982; Stiglitz, 1987; Lundvall and Johnson, 1992; Nelson, 1992). Relational assets of this sort are not freely reproducible from one place to another, and access to them is determined at least in part through network membership (Storper and Venables, 2002). This is why untraded interdependencies tend to have a strongly place-bound and culturally-rooted character and often cannot be transferred easily – if at all – from successful to less successful regions (Becattini, 1990; Putnam et al., 1993). Because access to these assets is spatially and organizationally limited, they enhance the economic advantages of their home regions (as well as local business enterprises and network members) and enable them to engage in monopolistic forms of competition à la Chamberlin, 1933 (see also Greenwald and Stiglitz, 1991).

These observations indicate that regional economic development involves a mixture of exogenous constraints, the reorganization and build-up of local asset systems, and political mobilization focused on institutions, socialization and social capital. More generally, the extent to which any region succeeds in creating localized increasing returns effects – which depend importantly on these cultural and institutional foundations – is critical to the entire development process. A direct extension of this point is the claim that the success of national economies (as indicated above all by accession to membership in the global high-income convergence club) is, in significant ways, related to the rise of dynamic and creative agglomerations, as illustrated by the case of the high-performance Asian economies. If this claim is correct, it follows that for countries to join the high-income convergence club in today’s world, they will have to sustain successful agglomerated development processes (though this remark in no way implies that balanced and sustainable rural development is not also an essential ingredient of...
any pathway to national development). Agglomeration is a central concern that can neither be equated to urbanization as a simple demographic phenomenon, nor dissolved away into the realm of macroeconomics.

**THE REGIONAL DIMENSIONS OF DEVELOPMENT POLICY**

In view of these remarks, one of the important tasks of any viable development policy must be to cultivate those multiple and important benefits that flow from regionalized production systems as they encounter the complex intraregional and interregional conditions that govern the logic of agglomeration. A number of negative aspects of interregional competition must also be brought under control if development is to proceed with any degree of smoothness.

Over the last half century, regional development policy has in practice tended to assume the guise of stimulus packages applied to given localities in the attempt to initiate take-off or to counter stagnation. The types of packages selected for these purposes vary greatly from country to country, but they generally comprise programmes such as subsidies to industry, tax breaks, infrastructure provision, governmental schemes to direct new capital investments to lagging areas, labour retraining programmes, and so on (Harrison et al., 1996; Donahue, 1997a). Approaches like these are not necessarily always devoid of positive effects, but in the light of our earlier discussion, they are certainly problematical when they occur in a vacuum, by which we mean a failure to attend to the critical organizational and institutional foundations of regional growth and competitiveness, as discussed above.

Since the 1980s, a burgeoning body of literature and practical experiments has accumulated in which these foundations have indeed been shown to be essential bulwarks of the regional development process (see Bianchi, 1992; Scott, 2001; Schmitz, 2001). More specifically, as we have noted, regional economies are internally tied together through human and organizational interdependencies — often untraded — that have a strong quasi-public goods character, meaning that they are the source of positive externalities that are more or less freely available to local firms but are the property of none. Such positive externalities are observable in diverse domains of regional economic activity, including dense information flows, learning processes, the emergence of craft or design traditions, business network formation, and so on (Storper, 1997; Scott, 2002). In this regard, we can refer to a ‘regional economic commons’ representing the elements of economic advantage that emerge out of the collective order of agglomeration, but that by their nature cannot be reduced to individual ownership and control. These elements are crucial for overall regional success, especially in a globalizing economy.

Concomitantly, new kinds of policy interventions based on the concept of regional economies as aggregates of physical and relational assets need to be identified and refined. This is because the development-enhancing synergies of these assets are subject to two main problems. First, positive externalities tend to be undersupplied where market relations alone prevail (Bator, 1958; Mishan, 1981). For example, skills training, labour market information, technological research, and so on are often severely undersupplied because producers are tempted to engage in free-rider behaviour by poaching these resources from the regional resource pool (Braczyk et al., 1998; Maskell, 1999; Johanssen et al., 2001). The net result is that overall investment in critical social assets is less than optimal, and it is hence essential to devise forms of policy intervention to rectify this problem. Second, significant moral hazards appear (especially in the core network operations of the regional economy) that can generate severe negative externalities if left unattended. These include the emergence of low trust relations between manufacturers and subcontractors, or threats to the reputation of regional product quality due to unscrupulous operators. In addition, regional economies are susceptible to variable developmental pathways, and policy can often help to steer any given system away from undesirable low-level outcomes (such as deteriorating working relationships between manufacturers and their subcontractors) and toward lock in at a higher-level equilibrium.

The many-sided regional economic commons that develops within and around dense industrial agglomerations, then, represents a critical domain of beneficial policy intervention. There are many different frameworks within which such intervention can be undertaken. These include governmental agencies, civic associations, private–public partnerships, or a host of other possible institutional arrangements, depending on local traditions and political sensibilities. Although the need for such intervention exists in regions at every degree of poverty or prosperity, it is probably most difficult to achieve in localities that have high deficits of basic physical and relational assets to begin with. This sort of intervention, in any of its institutional guises, has little resemblance to traditional urban policy, with its emphasis on infrastructure, housing, transportation and urban public finance; it is instead oriented toward the problem of coordination of urban production systems. Moreover, since it depends so greatly on the active consent of many different individuals and groups, it also calls for a high degree of social and political engagement, in which firms, workers and other stakeholders in the local economy are brought into meaningful public debate about what is at issue and about the preferred shape of collective outcomes.

A confirmed anti-dirigiste such as Lai, 1983, would certainly raise objections at this point to the effect that it is always better to live with market failures than with
the ‘inevitable’ gaffes of public agencies and their encouragement of rent-seeking behaviour. No matter how salutary this warning may be, it is tempered to the extent that the notion of a regional economic commons – offering compensating returns to coordination – can be sustained, and accordingly, that the operation of the local economy can enhanced by suitable policy intervention. Our discussion of this matter suggests that there is indeed a role for collective action in promoting regional increasing-returns effects and raising the long-term rate of economic growth, and this claim is entirely consistent with the same point made for the economy as a whole by the new growth theory (Lucas, 1988; Romer, 1990), as well as by the large theoretical and empirical literatures on the social and institutional foundations of successful markets.

Rising levels of local activism in the matter of regional economic development, however, do create some additional risks. These take many different forms ranging from irrational development races, through fiscal wars over subsidies and investments, to the poaching of one region’s talent and resources by another, to locational tournaments for large inward investments (Donahue, 1997a; Bartik, 1991). Interregional competition in pursuit of first-mover advantages is a striking instance of this problem. It is perhaps most clearly evident where several different regions are all striving to become incubators of some critical infant industry, and hence to emerge eventually as the leading centres of that industry as it matures. But in the presence of agglomeration economies, only one or a few champion regions are likely to be successful in any given production niche over the long run, implying that in the absence of informed coordination from the beginning, considerable misallocation of resources will in all likelihood have occurred. In general, securing positive-sum returns to development at the interregional scale – and in a world where many individual regions are actively striving to build internal developmental competencies – appears to demand some additional layer of regulatory oversight. Certain European Union injunctions against ‘social dumping’ are attempts to establish interregional coordination of this type, though they remain insufficiently specific to be fully operational.

At the same time, the formerly widespread policies of central governments to promote regional income equalization have been considerably diluted in recent years in both developed and developing countries, thereby helping to accentuate the processes of interregional income divergence already noted. Globalization as it is being moulded by the ideology of the Washington Consensus may encourage further watering down of these policies, especially where this is accompanied by the enforcement of contractionary monetary policies and fiscal austerity programmes in developing countries (Stiglitz, 2002). As a corollary, many countries are also concentrating their public expenditures on their most dynamic, globally-linked agglomerations at the expense of basic equity issues both within these agglomerations and between them and other areas of the national territory (ACS, 2000; Phillips, 2002). The concomitant tensions built into the contemporary development process – at all geographical scales, whether intrametropolitan, interregional or international – can lead to political backlash in which even the potentially positive aspects of development and globalization may not be recognized because of a failure to deal with their more egregiously negative effects, including the exacerbation of social and interregional inequalities.

All of this suggests that the regional components of economic development policy under contemporary conditions pose a knife-edge dilemma. On the one hand, policy needs to be designed in ways that strengthen agglomeration economies. On the other hand, isolated attempts to strengthen agglomeration economies may intensify disparities in per capita incomes along many different lines of cleavage (Wagner, 2001a, 2001b). These two aspects of the question are in constant tension with one another in today’s world, as exemplified by current debates in which some analysts hold that development policy is best focused on productivity improvements in dynamic agglomerations, (thereby maximizing national growth rates but increasing social tensions), while other analysts suggest that limiting inequality through appropriate forms of income redistribution (social and/or interregional) can lead to more viable long-run development programmes (Amsden, 1989; Aghion, 1998). In any case, for virtually every country, there is today a serious and much neglected policy issue involving the achievement of more effective forms of central/regional coordination and a more appropriate spatial distribution of political power (Bolton et al., 1996; Donahue, 1997b; Inman and Rubinfeld, 1997; Cheshire and Gordon, 1998; Wagner, 2001b).

Analogous tensions over development disparities recur at every level of geographic scale in the world economy (Held et al., 1999), and especially at the global scale itself. In the current regime of intensified globalization in which market imperatives consistently outrun existing institutional capacities for effective regulation, the balance appears to be strongly tilted to increasing inequalities. The discussion laid out here presents the case for an explicit consideration of the economic geography of globalization in relation to its regional foundations, and this issue needs henceforth to figure prominently in any reorganization of the institutions comprising a new multi-scalar system of governance. The competing claims of growth and equity remain firmly on the agenda, even if the current pressures of spatial economic reorganization make it necessary to re-think the ways in which we can best achieve balanced responses to them.
CONCLUSION: DEVELOPMENT THEORY THROUGH THE LENS OF ECONOMIC GEOGRAPHY

Conventional economic theories of development and trade have by and large ignored questions of economic geography. Today some of this neglect is being rectified by economists with an interest in agglomeration economies and regional dynamics (see, for example, Fujita et al., 1999). In our view, however, this perspective can be taken further. The existence of pervasive agglomeration economies based on externalities and increasing returns effects calls for a full recognition of the region as an organic unit of economic reality. This is because agglomeration economies represent a potent, immobile and quasi-public goods—a highly-problematical element of the entire development process. As such, regions exist as keystones of economic organization just as firms, sectors and nations do. Development theory needs now to recognize this point and take it into account.

As we indicated at the beginning of this paper, economists have tended to privilege macroeconomic variables as the best possible line of attack on the problem of development. But this level of observation, though obviously important, is no longer (if it ever was) the uniquely privileged point of entry to an understanding of development, and all the more so today given that the barriers between national economies are in certain respects breaking down, thus enhancing tendencies to agglomeration at selected locations all over the world. Moreover, while development theories directed at poorer countries have at times recognized the fundamental two-way connection between industrialization and urbanization, they have tended to focus on the problem of hyperurbanization and its negative social repercussions, rather than on the region as a locus of high-productivity outcomes. Our point is that one of the most fundamental issues for developing countries today is how to create and sustain the kinds of agglomerations without which they can never hope for entry into the highest ranks of the global economy, while ensuring that income disparities remain well within the limits of the socially just and politically tolerable.

This state of affairs poses many new questions for development theory and policy at the regional, national and international scales. We have sought in the present paper to move beyond elements of development theory that impede a fuller recognition of the geographical realities of the globalization process and to sketch out the beginnings of some broad responses to the questions raised by this exercise.

NOTES

1. This paper was originally written as a position document for the working group on Globalization, Regions, and Economic Development, sponsored by the Center for Comparative and Global Research, International Institute, University of California Los Angeles.

2. Some analysts suggest that this claim is less tenable where calculations are based on population-weighted measures of per capita income by country. When these measures are unweighted, however, the results show that post-war trends toward greater convergence have been decisively reversed in recent years (Sala-i-Martin, 2002).

3. As Davezies, 2000, notes, this dilution concerns mostly wage-based income equalization policies in the European countries. Certain kinds of transfer payments compensate in part for this dilution.

REFERENCES


Allen J. Scott and Michael Storper


SCOTT A. J. (2001) Industrial revitalization in the ABC municipalities, São Paulo: diagnostic analysis and strategic recommenda-


