CHAPTER 1

The Resurgence of Regional Economies, 1
10 Years Later

THE REDISCOVERY OF THE REGION AND ITS CRITICS

Something funny happened in the early 1980s. The region, long considered an interesting topic to historians and geographers, but not considered to have any interest for mainstream western social science, was rediscovered by a group of political economists, sociologists, political scientists, and geographers. Not that no attention had been paid to regions by social scientists before that: in regional economics, development economics, and economic geography, such topics as regional growth and decline, patterns of location of economic activity, and regional economic structure were well-developed domains of inquiry. But such work treated the region as an outcome of deeper political–economic processes, not as a fundamental unit of social life in contemporary capitalism equivalent to, say, markets, states or families, nor as a fundamental motor process in social life, on the same level as technology, stratification, or interest-seeking behavior. Economic geography was thus considered to be a second-order empirical topic for social science.

In the early 1980s, in contrast, it was asserted that the region might be a fundamental basis of economic and social life “after mass production.” That is, since new successful forms of production—different from the canonical mass production systems of the postwar period—were emerging in some regions and not others, and since they seemed to involve both localization and regional differences and specificities (institutional, technological), it followed that there might be something fundamental that linked late 20th-century capitalism to regionalism and regionalization.

Certain images piqued the interest of social scientists: the dense vertically disintegrated industrial districts of Northeast–Central Italy, Toyota City, Silicon Valley, Orange County, Route 128, the cité scientifique of Paris, Toulouse, Baden–Württemberg, and Bavaria; and even such lesser-known and less high-tech cases as the London and New York financial districts, Los Angeles’ garment district, the Hollywood entertainment industry, the craft industries of Jut-
land, the metalcutters of the Haute Savoie in France, Sakaki in Japan, as well as hundreds of others. All of these were said to be manifestations of a resurgence of the region as the center of "post-Fordist," "flexible," "learning-based," production systems—the emerging face of capitalist industry in this *fin de siècle*. On a larger scale, it became evident that even with increasing intensity of global trade and investment flows, national specificities in terms of products traded and technologies produced were increasing: in certain respects, integration was not bringing similarity, but specialization, a form of regionalization.

But how important is regionalization? Is the region somehow a necessary source of the *dynamism of these production systems and, hence, of the developmental dynamics of contemporary capitalism itself*? Or is regionalization merely an expression of another interesting empirical dimension of technological and organizational changes in successful production systems?

Surprisingly, a fairly large number of social scientists, and not just those whose professional specialty is the region, began to respond in the early 1980s that regionalization was very important, and that it might be more than merely another localization pattern: it might actually be central to the coordination of the most advanced forms of economic life today. And so a lively debate was joined. Over a period of more than ten years, the initial propositions have been reexamined theoretically and empirically, and new propositions have emerged. The debate over regionalization in contemporary capitalism continues to generate fascinating propositions, and for the first time has been taken seriously by social scientists interested in such central topics as technological and organizational innovation and national competitive advantage in a world economy. The subnational region (including such variants as cross-border regions, which include parts of more than one country) is also increasingly regarded as a level of economic policy making for these reasons (and not simply as a way to buy political calm, the traditional reason in much of Europe for regional policy). The stakes of these debates are big in both theoretical and practical terms.

Three main "schools" have participated in the debate: those interested in institutions, those focusing on industrial organization and transactions, and those who concentrate their attention on technological change and learning. Each has made strong claims about the bases of the new competition and the role of the region, and each has inspired empirical research and theoretical critique. A reprise of the debate will show that there is good reason for including the region as an essential level of economic coordination in capitalism. But none of the main schools in the resurgence of regional economies debate has come up with the correct formulation of why this is the case. The critics of each of these schools, while in no way dismantling the case for regional economies, have shown their attempts at formulating the basis for the resurgence of regional economies to be partial, although often very rich in insight. The purpose of this book is to argue that the most general, and necessary, role of the region is as the locus of what economists are beginning to call "untraded interdependencies," which take the form of conventions, informal rules, and habits that coordinate economic actors under conditions of uncertainty; these relations constitute region-specific assets in production. These assets are a central form of scarcity in contemporary capitalism, and hence a central form of geographical differentiation in what is done, how it is done, and in the resulting wealth levels and growth rates of regions.

**INSTITUTIONS AND INDUSTRIAL DIVIDES: THE ILL-FATED DEBATE OVER SMALL FIRMS**

From the mid-1970s, Italian scholars called attention to the different development model that characterized the Northeast–Center of their country, dubbed the "Third Italy" by Bagnasco in 1977. In the English-speaking world, the industrial systems of that region were made famous by Piore and Sabel in 1984. They were the first to capture them in the form of a model; flexibility plus specialization. Generalizing from Italy to certain other cases (notably German), they then placed the success of such forms of production in macroeconomic and historical context and postulated the possibility of an "industrial divide" separating a putative era of flexible specialization from that of postwar mass production.

Their account was both empirically rich and theoretically powerful. It incited a debate that centered mostly on the empirical material they adduced to support the theory. Yet it was the theory, and at that, the elements that were least picked up in the subsequent debate that makes, in my view, the lasting contribution to our understanding of capitalist development in general and the status of the region in particular. Piore and Sabel echoed and paralleled the work that had been going on in Florence under the direction of Giacomo Becattini, who had become one of the major contemporary students of Alfred Marshall. Piore and Sabel made an analogy between what was happening in Italy and Marshall’s notion of "industrial districts" in late 19th-century England. Becattini’s Florentine group engaged in a systematic elaboration of the concept of a “Marshallian industrial district,” centering both on its economic characteristic (externalities lodged in a division of labor) and on the sociocultural supports to interfirm interaction within an industrial district, this latter both through theoretical work and through very detailed studies of the history and structure of Tuscan industrial districts, especially that of Prato, the woolen goods district next to Florence.

There are several fundamental contributions of the Italian school, and the formulations of Piore and Sabel and others, to the study of contemporary capitalist development. These contributions consist in four points that remain unchallenged.
First, technologies of production and divisions of labor in production are not dictated by a movement toward a globally optimal, foreseeable "best practice" for each sector. They are, rather, the outcomes of institutional pressures and choices made at critical points in the histories of products and their markets, and the direction of development is thus not necessarily toward greater scale and integration, but can be the reverse. It is thus possible that rather different, but still efficient, forms of organization will coexist in the same sector.

Second, the flexible specialization school got something basically right in identifying flexibility and specialization as fundamental alternatives to mass production. From IBM to Modena, these principles have not been challenged, though we are learning that the specific ways they are organized into industrial systems frequently take us far from the ideal types of Italy and southern Germany. A consequence of both these observations is that the "industrial divide" postulated by Piore and Sabel has probably been crossed, in the sense that the postwar mass production economy is being replaced by one characterized by greater flexibility and specialization. The strength of the analytical insight regarding flexibility-plus-specialization is suggested by the fact that it has made its way into virtually every important theory of the firm and the production system in "post-Fordism," including the high-volume, lean production models of Japan, and that the language of flexibility-plus-specialization is now employed openly by corporate managers themselves when they restructure (as in the case of IBM). Moreover, the statistical evidence supports the overall notion that post-Fordism has arrested the long-term trend toward increased vertical integration and, in some cases, reversed it. In the American economy, for example, total vertical integration declined from 30% to 21% between 1977 and 1987, and even manufacturing vertical integration decreased from 30% to 27%. In most of the specific sectors said to be increasingly flexible specialized, these declines were even more dramatic.

Third, though the original examples of regionalism were much too pure, it now seems clear that some of the most dynamic forces in contemporary capitalist development—especially advanced forms of technological learning—are both localized and territorially specific. This insight, too, has become common currency not just among regionalists, but among economists and students of technology and trade. The contribution of the flexible specialization school was to assert that this seemed to have something to do with regionalized and territorially specific institutions, though this was not stated as clearly as would have been desired ten years ago.

Fourth, the key texts of this school emphasize that appropriately institutionalized networks are essential to successful ongoing adaptation of a regional economy in the face of uncertainty (technological, market, etc). Something approaching a new orthodoxy has arisen in the academic business economics and management literature in recent years: a network paradigm for organizing production systems. Participation in such networks is akin to a new best business practice, in much of this literature, much in the way that mass production was best practice three decades ago. There are now detailed microeconomic analytics for such a production paradigm, from single firms to whole "chains" (commodity chains). They differ essentially from the main points of the flexible-specialization school, where there is little emphasis on a single best practice or on the notion of optimum.

The critics of the empirical account centered on the lessons of the "Third Italy." They took two different routes. The first and least powerful attempt simply to deny the characteristics attributed by Bagnasco, Becattini, Bellandi, Brusco, Russo, Sforzi, Solinas, Dei Ottati, Trigilia, Regini, and Piore and Sabel, to the industrial systems of NEC Italy. These failed attempts at criticism rested simply on either a misunderstanding of the Italian systems, in empirical terms or of an overreach of the model, reducing it incorrectly to a model about small firms, rather than a model about systems or networks of flexible and specialized firms. Not all the criticisms of the Third Italy/industrial districtsflexible specialization model were spurious however; much of the problem stems from Piore and Sabel's attempt to root their deep theoretical claims about institutions, the division of labor, and the possibility of industrial divide in a very specific notion of flexible-specialization, drawn from Italian and southern German examples: the generalizability of those cases is limited. Some of the criticisms include:

1. Production systems dominated by small firms, especially the exceptionally small firms of the Third Italy, are few and far between in this world; so it gives a wrong impression to base the possibility of an industrial divide on small firm examples;

2. A model of an industrial divide needs to cover a wide sectoral composition, and the Italian and German examples concentrated on traditional non-durables or on specialized supplier (e.g., metalworking) industries, or on luxury versions of mass production (e.g., German cars);

3. There are deep historical roots to the Italian and German examples, although none spring directly and continuously from history or tradition still, the critics asked how it could be expected that in other regions similar forms of industrial skill and coordination could be built de novo. If a balance of cooperation and competition was a particularity of the conventions of certain regions, what about other regions with more orthodox (i.e., Anglo-American) competitive norms?

4. Were these just "niche" producers, enjoying favorable market positions while mass producers would occupy the central terrain of their respective sectors? The critics said yes, the flexibilists said that the center, too, was giving way to flexible, high volume, "lean" production systems. Here, the valid critique of the initial "flexible specialization" thesis concerned its empirical cases; but to hold firm, the critics would have had to show that there was
no significant big-firm version of flexibilization. Some attempted to do just that, claiming that Toyota was not really much different from any other mass producer and that big firms like IBM were really where the action was. But the bulk of the analyses—from Ohno himself to both Japanese and foreign observers of Toyota—claimed the opposite, and the dramatic restructuring of IBM for greater internal decentralization and flexibility in the early 1990s threw a heavy shadow on assertions that flexibility is an affair at the margin of markets. Still, though Piore and Sabel did assert clearly that there was a convergence between big and small firms, they conflated a general flexibilization of production with the particular model of flexible specialization within vertically disintegrated, small firm industrial systems. To this day, one cannot have a debate about flexibility as a general characteristic of contemporary industrial systems without it being confused with the specific model of flexible specialization in the sense of Italy and Germany. And this is true even with attempts, such as that of Hirst and Zeitlin, to distance flexible specialization from the empirical contours of Italy and Germany, by rebaptizing it an “ideal type,” for the ideal type retains the strong image of small firm, cooperative, and regionalized input–output relations.

5. In drawing on the extreme case of localization, that of the Third Italy, the flexible-specialization school opened itself up to the obvious criticism that most competitively successful production systems do not approach that level of regional closure. Toyota City is, considered in this view, to be merely another exception.

6. The most important criticism, but one that was least precisely formulated by the critics, is that the flexible-specialization model did not define, in analytical terms, precisely what it was that distinguished a technologically dynamic, regionally rooted system of firms from those systems of firms that did not share these characteristics, but still appeared to be flexible and specialized. They correctly understood that the flexible-specialization model was only interesting to the extent that such production systems were technologically dynamic and not highly territorially mobile, but the words flexibility and specialization did not necessarily correspond to these characteristics.

Likewise, though it developed certain building blocks of a model of performance, having to do with the importance of collaboration, competition, trust, and confidence, the flexible specialization school itself did not develop an analytical language about ongoing industrial adjustment—or what is now called “learning”—sufficient to capture in a generalizable way the nature of flexible and specialized industrial systems that have long-term survival capacity without wage–price reductions, from those do not.

This book will attempt to redress these deficiencies, but in so doing it will need to reframe certain debates about industrial change and its geography in recent years.

INDUSTRIAL ORGANIZATION, TRANSACTIONS, AGGLOMERATION: THE CALIFORNIA SCHOOL OF EXTERNAL ECONOMIES

What we might call, for lack of a better term, the “California school” came at the problem of new production paradigms and the region from the perspective of different industries, and a different political–institutional setting, from those described above; as such, it is not surprising that they also chose a different theoretical route. In the early 1980s, Allen Scott was already theorizing the relationship between the division of labor, transactions costs, and agglomeration, in his initial studies of the women’s clothing industry in Los Angeles. Just shortly thereafter, with no real prior theoretical disposition, Susan Christopherson and I, in studying Hollywood’s film and television industries, observed a strong process of vertical disintegration; and Piore and Sabel’s book appeared while we were interpreting our empirical results. Other investigations, many carried out by Scott and his students, followed, and both Scott and I continued this work with our own investigations in France and Italy in the mid- and late 1980s.

Other geographers and regionalists, such as John Holmes, took an early interest in the division of labor as well.

The argument that emerged rooted flexibility in the division of labor in production, and linked to agglomeration via an analysis of the transactions costs associated with interfirm linkages. In essence, it took what seemed to be fact in the Italian cases and created an economic model of the agglomeration process. It assumed that certain exogenous or endogenous market conditions gave rise to uncertainty—shifts in market conditions, or movements along a technological trajectory, for example. This uncertainty is met by externalization (disintegration), in one case to minimize exposure to risks of overcapacity (both production and labor force hoarding), in another to maximize the benefits of specialization and minimize the danger of technological lock-in.

Disintegration of production, caeteris paribus, in turn raises the transactions costs of input–output relations. There are more transactions external to the firm, and these transactions are, in a number of the empirical circumstances identified, more frequent, less predictable, and more complex. This raises their costs with geographical distance, and the feasibility of carrying out substantively complex transactions drops for certain kinds of complexity (especially noncodifiable or tacit knowledge or where trust is required and full contingent contracting impossible). So agglomeration is an outcome of the minimization of these transactions costs, where such minimization outweighs other geographically dependent production cost differentials.

This organization and cost-related explanation holds that agglomeration is one element in the external economies that attach to interdependent production systems. Under the circumstances specified above, in the presence of
agglomeration the advantages of interdependence—flexibility, risk minimization, specialization—are increased. In the extreme case, without geographical proximity, these advantages would be cancelled out by the increased costs or difficulties of the intensified transacting. The advantages of agglomeration are external economies because flexibility lowers input costs (by minimizing factor hoarding) and increases throughput to each firm in the sense that a greater number of external input-output interconnections raises the probability of successful sale or purchase, all other things being equal. (A technical issue debated by economists is whether these economies are truly external, since we don't really know whether the production functions of firms are truly interdependent and whether their returns to investment are uneven; all we do know, with certainty, is that the ensemble of firms has cost savings due to interconnections and these cost savings are enhanced by geographical proximity because it lowers the resulting transactions costs.)

This analysis thus partakes of a major trend in the business economics literature, shared by much economic sociology: the economics of network forms of production.25 This analysis seemed to have several advantages over the institutionally inspired flexible-specialization school. First, it did not seem to depend on thick and historical institutional contexts. Indeed, one of the main claims it made had to do with the establishment of new industrial spaces. It argued that new industries—those that emerge after technological branching points—have input structures independent of older industries—and hence enjoy what we labeled “windows of locational opportunity,” in the sense that they are not attached to old stocks of external economies.26 But once a group of firms begins to get ahead, the proliferation of external linkages gives them advantages that rapidly attract new entrants and hence, only a few major new agglomerations can form in a given new industry. So it offered an explanation for new flexible production agglomerations such as Silicon Valley.

Second, it argued that the reagglomeration of older industries, analogous to those found in the European cases, could be accounted for via the process of externalization and interlinkage of firms—the story of Hollywood (told in Chapter 4), going from its own version of “mass” production and spatial diffusion toward vertical disintegration and reagglomeration, was a case in point. It averred that there could be many reasons for such switches in the organizational and geographical pathways of development—in Hollywood, it was regulatory and technological changes that set the process in motion;27 in the Third Italy, it was a combination of long-standing civic cultures and the events of the postwar period28; in the Los Angeles fashion industry, it was endogenous changes in fashion and the possibility of making distribution more attentive to consumer demand, enhancing the number of collections per year.29 Another major case would be consumer durables industries, where technological changes in production and distribution made possible more rapid changeovers; once these possibilities were realized by the Japanese, all world competitors had to follow suit. The list could go on, ad infinitum.

The coverage of the model, in other words, was meant to be greater than the initial version of the flexible specialization model, in that it allowed any mix of firm sizes, any sector, any mixture of linkages. Indeed, the model was defined around three groups of sectors, which account for increasing proportions of employment and output in industry—high technology, revitalized craft production, and producer and financial services—thus extending its reach.

As the debate proceeded, however, we realized that the linkage-transactions cost model was incomplete. Toward the end of the 1980s, we became more sensitive to the possibility that the agglomeration was itself a source of industrial dynamics. We held that agglomerations, once in place, constituted industrial communities where endogenous dynamics of knowledge and technology development occurred, drawing on the example of technology development in Sassuolo documented by Russo,30 and extending this to our own California cases. This account of technological innovation turns explicitly on user-producer relations—that is, on information transactions in an input-output system—and holds that localization and appropriate communication rules are important to innovation in some industries.

This latter brought us to the question of institutions. Though agglomerations could be theoretically accounted for as the way that potential external economies were realized, there was no assurance that markets alone, nor even various forms of contracts, could successfully coordinate the nexus of transactions in an industrial agglomeration.31 Such transactions—in labor markets, in interfirm relations, in innovation and knowledge development—tended to have points of failure in the absence of appropriate institutions. In these two respects—evolution and institutions—we attempted to go beyond the initial Williamsonian framework to argue that the “institutional arrangements” of agglomerations—that is, the nexus of transactions and their economic performance—were themselves outcomes of broader institutional environments, and themselves generators of future choices for pathways of development. So we came “full circle” to rejoin the initial authors of the flexible-specialization thesis (and the Marshallian theme of “atmosphere”), albeit with somewhat different perspective on the role of institutions in development as a whole. And, as we shall see shortly, we came to realize the central importance of the new economics of technological change and its core notions of evolution and path dependency, for our problem of the role of the region in late 20th-century capitalism.

As with the flexible-specialization school, there were spurious criticisms and serious critiques. Among the latter we may count the following.

It seems as if the California school’s analysis does apply to certain modern craft or traditional industries as well as certain labor-intensive manu-
manufacturing and service sectors: in clothing, furniture, jewelry, cinema and television, some financial services and banking industries, as well as certain segments of the mechanical engineering industries, the density of local linkages is consistent with the account of agglomeration. In other words, the California school's successfully analyzed cases overlap very strongly with those of the flexible-specialization school. The critique does not attempt to deny existence of agglomerations, nor of their recent growth, in other industries, but observes that dense local input-output relations are not present in them in sufficient quantity to account for the existence of the agglomeration. The direct local input-output relationships between firms are not dense enough to account for either the size of the agglomeration or for a high proportion of what goes on in the sector. The examples where this seems most applicable are parts of high technology and certain parts of supplier-intensive sectors, such as mechanical engineering—that is, the capital-intensive, high-wage examples of flexible production.33

The critics have attempted a sort of counter theory, which has as its point of departure not the input-output system, but the firm.34 The post-Fordist firm is, for them, a nexus for the management of vast flows of resources, the principal node in a set of shifting property and production networks. The means for the firm to manage these networks is a mixture of ownership, contract, and alliance,35 and in general, this new flexibility of networks carries out all the functions that disintegration does in the flexible-production school's analysis. Rather than an economy of direct cost reduction, even for the most innovative activities, we are in an "economy of organization,"36 where scale, over long distances, can reduce the time and cost of flexible adjustments of capacity and shifts in product mix. Flexibility is retained as a key, but now linked to scale and geographical dispersion. Much of the causality of the flexible-production school's analysis with respect to regions is reversed.

Their description of the large firm as a nexus of shifting relations, and of the industrial system as a nexus of nexuses, is at the heart of much of the contemporary empirical economics of the firm.37 It is not clear, however, that this stands in as an explanation of innovation, agglomeration, or the geography of input-output systems. If the flexible economy were really contained in such dispersed large-firm networks, for example, why would such firms allow significant parts of their activity to be "trapped" in specific countries and regions? Pavitt and Patel,38 and Dunning,39 all show empirically that the core technological activities of the biggest firms are principally rooted in their home countries. Why, indeed, would a firm like IBM bother with the cost and inconvenience of Silicon Valley at all? Here, the economy of organizations school responds that big urban areas are general "basing points" for advanced (knowledge-intensive) activities with high levels of risk.40 We are back to urbanization economies. There are two problems with this. The first is that such

an explanation of big firm–big city interdependence is necessarily a transactions-based explanation, insofar as the reason invoked for big city location is proximity to factor markets due to the need for high levels of factor (read: labor) turnover in the presence of uncertainty. In countries with primate cities, such as France and Japan, what appear in the form of urbanization economies cannot be distinguished from multiple and overlapping localization economies present in Paris and Tokyo. The two ways of interpreting the problem would lead to radically different policy prescriptions. The second is that the economies of even the biggest of the big cities, but especially middle-sized cities, are not only urbanized, but have strong sectoral specificities. So urbanization economies are underpinned by localization. This localization involves parts of firms, those that are in certain technological or economic "spaces," in the sense defined by Perroux.

One suspects that the model of the firm they have evolved is applicable to the flexibilization of rather routine productive activities—precisely those where market uncertainty is manageable (for example in the mass-production segments of the clothing industry) or those where innovation is also gradual and manageable: this is the spatially extended economy of organization. The parts of these same big firms that are involved in those productive activities are not free of agglomeration nor free of uncertainty in the relevant parts of their input–output chains, and it seems unlikely that transportation, telecommunication, and formal institutional arrangements that is, strategic alliances and contracts) are sufficient to obviate the need for proximity in these cases.

Another approach that takes on some of the same questions as the California school is the new geographical economics of Paul Krugman and various international trade theorists. They explain geographical concentration of productive activity via returns to scale. These scale economies affect both intermediate and final outputs, giving rise to an uneven pattern of market dominance—a geometry of imperfect competition—and hence to specialization and trade (especially interindustry trade). The existence of pervasive scale-based market power reflects itself in place-based pecuniary externalities. This line of explanation intersects with that of the California school in that it focuses on why intermediate-goods producers might be found clustered in particular places.

But it seems unlikely that all clusters of intermediate-output producers reduce to market size; even if they did, we would need a theory to explain why and how they become so specialized that they must cluster near their few, specialized customers, and in Krugman's model this is not discussed. As we have noted, the interfirm division of labor, and its evolutionary dynamics, seem to be the drivers of specialization and, hence, of possible scale effects at the intermediate level. In terms of final outputs, the new geographical economics seems to take for granted the results of a long process of concentration of assets, but tells us little about how industries get there, or what space might have
to do with this. For example, if big final-output producers concentrate in an agglomeration to be near their specialist suppliers, then the primary issue for a theory would be how the structure of specialist suppliers and the demand for proximity is created, not simply the decision to optimize in the face of these circumstances. Where the issue is access to final markets, scale seems hardly an explanation for detailed patterns of location and trade where output scale is very high, market linkages are the most standardized, and distance the least costly element in them. Finally, there is almost nothing in the new geographical economics about the potential geographical foundations of economic performance. Everything reduces rather axiomatically to fully efficient, though imperfectly competitive, clusters of producers who enjoy pecuniary externalities. There is, thus, no question of the good versus bad agglomerations, nor of a potential link between agglomeration and performance through the possible effects of agglomeration on innovation and knowledge development.

In any case, the problem staked out by the California school remains: what are the sufficient conditions for the existence of the observed agglomerations of productive activity, which grew so strongly in the 1980s? The California school came up with an explanation, but it was partial, and it sensitized us as to the importance, complexity, and geography of input–output relations. But the localization of such input–output relations, that is, the localization of traded interdependencies, is inadequate to the task of explaining the link between flexible production and the resurgence of regional economies in contemporary capitalism.

The California school’s explanation also suffered from the same problem as that of the flexible-specialization school: the central aspects of its theory could not, in the end, distinguish between good regional economies and bad ones. Vertical disintegration, high transactions costs, and agglomeration could be found in both high-wage, technologically dynamic industries and in low-wage, technologically stagnant ones. Adding in institutions helps, in the case of traditional industries, to distinguish good versions from bad. But in technologically dynamic industries, agglomerations are often found without overwhelmingly dense local input–output linkages and without the kinds of explicit institutional coordination found in many European industrial districts. A different explanation was needed.

INNOVATION, HIGH TECHNOLOGY, AND REGIONAL DEVELOPMENT

The subject of innovation is prefigured in the schools of thought reviewed above, though innovation is seen in both as a consequence of the institutional or organizational frameworks of production. From the late 1970s on, students of regional development investigated the regionally uneven distribution of high technology industries and the apparently better propensity of some regions to develop “high-tech” economic bases than others. Defined as such, their problem was not the same as that of the two schools of thought reviewed above: rather than a broad problem of a change in production paradigms—flexibility and its consequences—they defined their subject as that of the incidence of industries based on new technologies. The empirical subject matter nonetheless overlapped with the preceding schools in certain cases; Saxenian’s work on Silicon Valley falls securely into the institutionalist flexible-specialization school, and Scott and Storper and others argued that high-technology development could be understood from the standpoint of their theory of agglomeration and the division of labor. What distinguishes the work reviewed here is that their point of departure was technology itself, mostly high technology.

In using high technology as the focal point of explanation, they attempted to isolate cases of “advanced” regional development. They assumed that by studying areas that had become centers of production for advanced technologies, there would be lessons for policy making (how to imitate these places), and—for some of these scholars—there would be Schumpeterian lessons about advanced economic development in general: what occurs to economic activities affected by waves of radical technological change? Two branches of this work can be reviewed here. The American school of high technology regional development sought the conditions for growth in Silicon Valley and Route 128. What was it that had set off and sustained growth in these places? The work identified many different factors said to have contributed to these processes, among which the single most discussed is the research university—spin off process. Drawing on the American interest in entrepreneurship (and a strongly American reading of Schumpeter), and on the key experiences of Stanford University–Silicon Valley, MIT–Route 128, an abundant literature on R&D and regional development was generated. To this was added a list of secondary factors said to be present in successful high-tech regions, among which were a “high quality of life,” good infrastructure, and even climate. It does appear, from the American historical record, that there were decisive links between universities and founders of firms in the Massachusetts and California cases. The problems come when this is taken to be a universal logic of new technology-based infant industry development. The logic works only when innovation is strongly formal-science based, as in the early years of semiconductors. It was not true, for example, in the case of airplanes in the 1920s and 1930s, where no local research universities were strongly present in the leap forward of Los Angeles. To this the high tech school responds: “that was then, this is now,” for organized science has become more and more important in the development of new technologies. So they predict that the university–production link will be critical in future technology-based industries, such as biotechnology.
Assume (for the sake of argument) that this is the case. The question is how to theorize what it is, specifically, that leads from research to the establishment of a regional production base. For there are many research universities, and even many that have generated lots of knowledge in semiconductors, but there is a much smaller number of Silicon Valleys and Route 128s. That which actually makes the interaction virtuous remains unexplained. The California school discussed above offered a partial response to this, by noting that what distinguishes nascent agglomerations from other early centers is the rate at which their external economies grow due to the proliferation of input–output linkages and this could, in principle, include university–firm linkages.¹⁷ But this observation, while essential, is insufficient: it does not plug up the gap in knowing what it is about the kind of knowledge or its transfer to producers that makes R&D presence effective sometimes and not effective other times and places.

A second branch of the American school is what we might call the "regional politics" approach.⁴⁸ It arose in part as a critical reaction to the first branch. In the place of the lists of factors (which it systematically showed to be irrelevant), it holds that regional coalitions secure resources that push for the transfer of high technology resources: thus, Silicon Valley got ahead partially because of the Stanford connection, but also because its early industrialists were clever enough to commandeer resources from the military–industrial complex. Indeed, much of the "gunbelt" developed because of politics, especially its key complexes in New England and Southern California. Yet these important observations fall far short of a complete explanation. For one, the Southern California aerospace complex was incipient before the growth of the military–industrial complex: its roots were in the success of Douglas Aircraft's DC3 in the 1930s. The military followed that installed technological competence; it did not place it there in Los Angeles. For another, there are many places where politically motivated investments have taken place (Texas, Georgia, Toulouse, Nice) but nothing approaching a diverse and dynamic high technology agglomeration has been created. Politics alone is insufficient.

The problem, then, is that the second American school has not come up with a coherent theory of high technology development or any necessary link to regional development.

An alternative, European, approach has been developed by the GREMI group (Groupement de Recherche Européen sur les Milieux Innovateurs), principally Franco–Italian–Swiss regional economists. Their central theoretical notion is that of the milieu. There are many different branches of this approach, many of them very rich; I will attempt simply to summarize what I see as the central theme that encompasses them all, with apologies to those who feel that their work is inadequately distinguished from the whole.⁴⁹

The milieu is essentially a context for development, which empowers and guides innovative agents to be able to innovate and to coordinate with other innovating agents. The milieu is something like a territorial version of what the American economic sociologist Mark Granovetter has labeled the "embeddedness" of social and economic processes.⁵⁰ The milieu is described, variously, as a system of regional institutions, rules, and practices that lead to innovation. Many of the milieu theorists use the "network" as their principal organizational metaphor. For some, the milieu is itself a network of actors—producers, researchers, politicians—and so on—in a region. For others, the network concerns the input–output system; it is this network that is embedded in a milieu, and the milieu provides members of the network what they need for coordination, adjustment, and successful innovation.

Milieu is suggestive of something interesting, which rejoins a key theme of the Marshallian school: that there is something intangible, "in the air" as Marshall would have it, that permits innovativeness to proceed in some places and not in others.³¹ The GREMI group, however, has never been able to identify the economic logic by which milieux foster innovation. There is a circularity: innovation occurs because of a milieu, and a milieu is what exists in regions where there is innovation. The following definition is exemplary, not exceptional:

A territory is not a defined space of resources. It is the mode of establishment of a group, in the natural environment, which through the organization and localization of activities, generates prevalent conditions of communication-language and collective learning (the forms of cooperation that create technological and organizational rationalities).

and

The milieu appears as the socioeconomic formation that, at one and the same time, generates the economic dynamic and constitutes itself in setting this dynamic into motion. In other words, milieux take form in organizing themselves and they do so even better insofar as they are territorialized. The emergence of organizational dynamism is correlative to the dynamism of local milieux.⁵²

The milieu school returns, again and again, to the properties of milieux, but they do not specify the potential mechanisms and processes by which such milieux function, nor precisely what the economic logic of a milieu would be—why localization and territorial specificity should make technological and organizational dynamics better. Thus, though they attempt to go beyond the input–output based models of the Californian school, they cannot seem to specify the logic or content of the intangible they are after. As such, they do not reveal what it is about regions in innovation that is essential to contemporary capitalist development.

Nonetheless, the GREMI group, and in particular the work of Perrin
(1993) has successfully reformulated the problem of what regional "science" should be all about, in calling for abandonment of regional analysis based on the two fundamental precepts of neoclassical economic science, that is, comparative statics (equilibrium), and the rational action paradigm for human behavior. Instead, they argue, the economic process is fundamentally about creation of knowledge and resources, and this "Schumpeterian" (and Marxian as well) process cannot be derived from the calculations of the rational actor on the margin. How economic actors reason and interact is, they argue, in large part a product of their context, and this context is likely to have—at least in part—territorial boundaries and specificities. So the process of economic creation by such actors should depend on their milieu.

How, then, to get beyond this impasse to define what this territorial context could consist of in generating the dynamic of economic creation—organizational and technological evolution—in capitalism?

TECHNOLOGY, PATH DEPENDENCY, AND UNTRADED INTERDEPENDENCIES

It was left to a group of nonorthodox economists and sympathizers in other disciplines—a group that was not principally interested in territoriality or regions—to develop the analytical tools which, appropriately adapted, now permit us to identify (at least theoretically, if not empirically) the intangible aspect of a territorial or regional economy that underlies innovative, flexible, agglomerations, of both the high- and low-tech variety.

The first insight came from the evolutionary economics pioneered by Nelson and Winter,53 and refined for the case of technology by Dosi, Arthur, Soete,54 and others. They claimed, essentially, that technologies develop along pathways or trajectories, which describe choice sets that are totally different from those of orthodox economics. In contrast to the standard model's key mechanisms of substitutability and reversibility in choice, where investments and returns can always be adjusted well to each other, they show that choices are characterized by strong irreversibilities; unlike the orthodox model, it is virtually impossible to predict outcomes from a starting point, even if actors are rational. The outcomes reflect no single optimum, but, at best, optima, which are continually redefined as choices are made and other choices are foreclosed. In evolutionary economics, what we do is path-dependent, that is, truly historical; it is not the result of a series of actions on spot markets, where the long-term can be reduced to a series of disconnected instants.

All this is the case because technologies are the products of interdependent choices. For one thing, technologies are subject to a variety of user-producer and user-user interactions: every technology made by a producer must have a user, and as the number of users of a given technology rises, it tends to cut off the possibility of different patterns of use (and hence, production) for other users. This is an external economy that benefits those who follow suit in both cost and feasibility terms (the story told by Arthur is that we all drive on either the right or the left, even though there's no efficiency reason, because everybody else does and it would be costly and hazardous to change); or David's story that we use the inferior QWERTY keyboard because it got ahead of other models early. Interdependency means uncertainty, since we cannot determine exactly what others upon whom our choices depend will do.

Beyond these external economies-as-accidents-of-history, there are reasons why producers tend to follow certain pathways. There are significant technological spillovers in the economy: knowing how to do one thing is frequently consequent upon knowing how to do another, or key to doing certain other things.57 This idea draws on the seminal work of François Perroux in the 1950s, who noted that an economy consists of "spaces" or fields of endeavor, in part having to do with the density of nontraded technological connections between them (for example, common types of knowledge or similar types of machines, or knowledge of how to work similar types of basic materials or inputs).58 In some cases, these nontraded connections overlap with traded input-output relations, that is, using similar inputs such as raw materials or capital goods, but often—and this is critical—the connections are untraded. The new economics of technological change has suggested, in the 1980s, that there are knowledge or "common practice" spillovers such that technological excellence comes in packages or ensembles.59 Since such excellence relies frequently on knowledge or practices that are not fully codifiable, the particular firms who master it are tied into various kinds of networks with other firms, through formal exchanges and untraded interdependencies. The latter include labor markets, public institutions, and locally or nationally derived rules of action, customs, understandings, and values.

Technological trajectories of a given industry are, in certain cases, parallel or complementary to technologies in other industries. Firms thus depend on decisions made outside their borders by other firms or decisionmakers in the technological space.61 When things are going well, external economies—interdependent production functions—create increasing returns for the ensemble of interdependent actors traveling down these pathways (and there is, in technical terms, no ambiguity about these effects being defined as true external economies).

Orthodox economics is mainly concerned with the way the market allocates resources and assumes that structural change and development follow from either the additive forces of this allocational process or from exogenous influences such as technological or demographic change. Evolutionary economics holds that technological change is an endogenous property of eco-
nomic systems and that it is not principally the result of allocational adjustments but of interdependent actions in which signaling, knowledge development, and doing the best one can are central.62

The evolutionists reinterpret the effects of competition in capitalism via this distinction. Comparative advantages are had when the position of a firm or of a group of interdependent firms in a nation or region find a more favorable location along a given production function: they are allocational. This is familiar to everyone who has ever taken an economics class. But, claim the evolutionists, this isn't what drives the distribution of market shares or the composition of economic activities in many cases. Absolute advantages exist when a firm, nation, or region possesses superior technologies such that virtually no set of alternative factor prices (production costs, in essence) would incite geographical redistribution of shares or activities. And this, they claim, is an aspect of the dynamic of capitalist competition wholly overlooked by the orthodox theory of competition-as-allocation.

The evolutionary economists working on technological change, not surprisingly, discovered territory—nations and regions. In theoretical terms, they began to reason that the technological spillovers and their untraded interdependencies would be territorialized under certain conditions, notably where the technological trajectories were particularly open, that is, had wide margins of potential variation. This would increase the uncodifiability and tactfulness of knowledge development and heighten the importance of communicational clarity and common interpretation in understanding information.53

In other words, the territorial specificity and differentiation of certain untraded interdependencies and spillovers, or the territorial differentiation of the same, by permitting actors to travel along superior technological trajectories (or to do so more rapidly than others) can confer on them absolute advantages that shelter them, at least temporarily, from Ricardian competition (that is, competition based on prices for factor inputs). This would show up as territorial specialization and differentiation in trade, whether between regions or at the international level.64 Virtually all the systematic empirical investigations of the technological performance of nations suggest that high, and increasing, differentiation or specialization characterize the western economies since the late 1970s. The main studies include Dosi et al.'s study of international trade,65 Patel and Pavitt's66 many investigations over the last decade, Guerrieri's67 interpretation of OECD statistics, and Dunning's study of multinational firms.68 Both Dunning and Patel and Pavitt note, in contrast to the prevailing assumption that multinational firms are indifferent to territory or to local context, that the major multinational firms of the world locate virtually all their most advanced technological capacities in their home countries and that where there are exceptions, they are almost always explained by investment in a technological core capacity of the host country; that is, they are attracted to the local technological tissue of another country. Jaffe, in the United States, has shown empirically that technological overspills correspond strongly to the patenting activities of firms;69 and Antonelli70 in Italy has indicated, in a very preliminary way, that overspills are often geographically bounded (although the measurement problems remain considerable). Evolutionary economics has not, however, provided a secure analytical basis, via the theory of localization and agglomeration, for why such untraded interdependencies and the technological trajectories they underlie should be territorially specific and differentiated.

Evolutionary economics alludes only indirectly to the counterforces to the territorialization process. Among them are technological imitation and selection, and the ongoing effort to transcend geographical distance in both the untraded interdependencies and input–output relations that are critical to technological learning. But those forces are very far from triumphing, or so it would appear from the empirical evidence. In sum, the evolutionary approach is, if not incompatible with the transactions cost approach to agglomeration, then at least wedded to fundamentally different theoretical commitments about dynamic processes of capitalist development. The transactions-cost-based theory of the California school is about allocation through cost-minimization,71 and this is because transactions cost economics is about traded input–output relations. The evolutionary approach is fundamentally not about cost-minimization; it is about the forces that allow the parameters of cost minimization to be altered and which get in the way of optimizing.72 And in terms of the mechanism of territorialization, evolutionary economics opens up the field of untraded interdependencies, which does not figure prominently in transactions-based approaches. Even though one can argue that untraded interdependencies are rooted in transactions—though perhaps not input–output transactions and market or contract exchanges—the analysis of such transactions cannot be easily accommodated within transactions-cost-based theories.

An answer to the principal dilemma of contemporary economic geography—the resurgence of regional economies and of territorial specialization in an age of increasing ease in transportation and communication of inputs and outputs and of increasingly scientific organizational rationalities of managing complex systems of inputs and outputs—must be sought in two lines of reasoning. One is the tension between respecialization and standardization of inputs and outputs, which, ceteris paribus, raises transactions costs associated with them.73 The other is the association of organizational and technological learning with agglomeration, which in turn has two roots. The first, and more limited case, is that of localized input–output relations, which constitute webs of user–producer relations essential to information development and exchange, hence to learning.74 The second and more general case is the untraded interdependencies that attach to the process of economic and organizational learning and coordination.75 Where these input–output relations or
untraded interdependencies are localized, and this is, as we have suggested, quite frequent in cases of technological or organizational dynamism, then we can say that the region is a key, necessary element in the “supply architecture” for learning and innovation. Theoretical predictions that globalization means the end to economies of proximity are therefore likely to be wrong, because they are deduced only from consideration of hard, traded input–output relations, those which are most prone to geographical dispersion at some point.

CONCLUSION

We have now traced the history of an intellectual debate over the resurgence of regional economies. The next step is to go at the subject in a more theoretically organized fashion: to reconstruct a theoretical framework for understanding the persistence of regional economic specificity in a global economy.

NOTES

1. This title is drawn, deliberately, from Sabel (1988).
2. See, inter alia, Sabel (1988); Scott (1988); Scott and Storper (1986); Becattini (1987).
3. Bianchi (1992); Bianchi et al. (1988); Best (1990); see also papers in Research Policy, special issue on “Networks of Innovators” (1990) and more recently, Tolomelli (1992); Lassini (1985); Bursi (1989).
4. I first became aware of this term in the work of Dosi (1984) and that of Lundvall (1988, 1990), though there are echoes of it in the writings of François Perroux and Thor Scitovsky in the 1950s.
5. Bagnasco (1977); Becattini (1987); Brusco (1982); Fiore and Sabel (1984); Belzandi (1986, 1989); De Ottati (1987); Russo (1986); Trigiulia (1986).
8. Even proponents of the thesis of a revival of certain principles of mass production, in the context of “lean production,” readily admit this as a lesson of recent industrial history; see, for example, Coriat (1991); the best statements, in my opinion, are Dosi (1987) and Dosi and Orsenigo (1985).
10. The new “theory of the firm literature” is so vast that we can offer but a tiny sampling Aoki (1989); Williamson (1990); Coriat (1991); Dosi et al. (1988); Dore (1987); Cohen et al. and Llerena (1989); Hakansson (1989); Best (1990); Porter (1990); Mariti and Smiley (1983).

13. The point that the districts draw some of their “social capital” from deep historical roots and traditions, but that they are not simple continuations of industrial history there, having reinvented themselves several times since World War II is found in virtually all the work of the Florentine school, such as Becattini (1978); De Ottati (1987, 1990, 1991); and, for the Emilian cases, Cappecchi (1990a, b). For a literature review, see de Maria and Scarpetta (1992). And for a clear theoretical argument about social capital, see Putnam (1993). Civic capital is treated in a theoretical form by Coleman (1988, 1990).
15. Sabel et al. (1989); but also suggested by Best (1990); Aoki (1989); and much of the business economics literature, which is more attentive to change in the business environment for both small and big firms, implying that size is not the principal axis of the analysis. See, for example, Hakansson (1987).
19. On the theme of collaboration and competition: Fiore and Sabel (1984); De Ottati (1987); Lorenz (1988); Sabel and Zeitlin (1985); Scranon (1985); Sabel (1993); Sabel et al. (1989); and on a theoretical level, see Axelrod (1984).
23. As collected in Scott and Storper (1986); see, especially, Holmes, in that collection.
24. Worked out in its fullest analytical form in Scott (1988a); see more recently, Scott (1993).
25. There is a vast literature on this subject to which I can do no justice here. See, for example, Powell (1990); Hakansson (1989); Aoki (1990); Foray (1990). See also Johansen and Mattson (1987).
27. Storper (1989); see also Chapter 4 of this book.
28. Becattini (1978); Brusco (1982); Cappecchi (1990a, b); Ritaine (1989).
32. See references at note 31.
34. See references at note 9. This is a spatial or geographical outgrowth of the strategic management literature’s twist on the network thesis, that is, convergence, optimality, and determinationalism.
41. Malecki (1984); see also Brechtny and McQuaid (1988); Glasmeier (1986).
44. Markusen et al. (1986).
45. Gauthier (1993); Scott (1990); and chapters on aerospace in Storper (1982).
46. See my critique in Storper (1985).
47. Scott and Storper (1987).
48. Markusen et al. (1986); Markusen et al. (1991).
49. Aydalot (1986); Aydalot and Keeble (1988); Camagni (1991); Mailat et al. (1990, 1993); Camagni (1992).
52. Both quotes are from an excellent paper by Perrin (1993). My translations attempt to get the sense of the message. They read, in the original, a bit differently:

Un territoire, ce n’est pas un espace délimité de ressources. C’est le mode d’établissement d’un groupe, dans l’environnement naturel, qui, dans l’organisation des localisations des activités, instaure et fait prévaloir les conditions de la communication–langage et de l’apprentissage collectif (des coopérations créatrices de rationalités techniques et organisationnelles).

Le milieu apparaît comme la formation socio-économique qui, à la fois, génère la dynamique et qui se constitue en la mettant en oeuvre. En d’autres termes, les milieux prennent corps en s’auto-organisant et ils y parviennent d’autant mieux qu’ils se territorialisent. L’émergence de la dynamique organisationnelle a été corrélative de celle des milieux locaux.

For another sophisticated treatment, see Dupuy and Gilly (1992).
54. Dosi (1987); Dosi and Orsenigo (1985); Dosi et al. (1990); Arthur (1989).
58. Perroux (1950a,b).
60. Dosi and Orsenigo (1985).
61. Young (1928) and Kaldor (1972) both tried to make these points early on; I tried to use their insights about interconnectedness in Storper (1989).
64. As noted in Porter (1990); Amendola et al. (1992); Patel and Pavitt (1992); and Storper (1992). See also Chapter 8 of this book.
72. Clearly brought out in Dosi et al. (1990); See also Johansen and Mattson (1987). For another sophisticated treatment, see Dupuy and Gilly (1992).
73. Scott and I have also argued that the ongoing process of product differentiation in capitalism destandardizes outputs (and hence inputs); this means that new inputs are necessary. The uncertainty attached to these new input-output relationships recreates the need for proximity in input-output transactions. We used the recent growth of financial services agglomerations as an example, that is, an industry where transport costs for the “product” are practically zero. See Storper and Scott (1989).
74. Lundvall (1990); Russo (1986); von Hippel (1987).
75. Our discussions of the substantive nature of untraded interdependencies in following chapters will refer to them as conventions (Lewis, 1969), and the inspiration for this is the French school of économie des conventions. References include: Eymard-Duvernay (1987); Thévenot (1986); Revue Economique (1989); Salais and Storper (1992, 1993); Storper and Salais (1997). This is by no means a complete list of what is a rapidly evolving literature.