

Markets, States, and Industrialization

Let us begin with a short journey. We start from the Hyatt Regency, a five-star hotel in Salt Lake, a suburb of Calcutta where many of the gleaming new offices of information technology giants like IBM, TCS, PWC, and Cognizant are located. We drive through the old congested and chaotic city, leaving behind a cluster of printing factories sitting cheek by jowl with a few city blocks populated by wholesalers and retailers of books, through a wholesale commodity market, across the river Hughli to Haora station. We catch a train to Bolpur in Birbhum district, about 150 kilometres away. Soon we are rushing past small stations set up to serve the iron and steel, jute, and cement factories dotting the banks of the river. After about 20 kilometres the factories begin to thin out and eventually disappear, and we are now in rural south Bengal, green with paddy fields and coconut groves. We see mud huts with roofs thatched with coconut fronds and doors so low that even the small undernourished women of the villages have to bend at the waist to pass through. By the time we reach Bolpur, the earth has turned reddish brown; it does not look capable of producing any crop. The town itself is dusty and small. There are no factories, only shops. It is a market town. We take a bus into the heart of the district where the Santhals, a minority group that is classified by the government of India as a Scheduled Tribe, live in abject poverty. Their huts are meagre, their tools few and antiquated; the majority of the women are officially illiterate, many of the men effectively so. Their monthly earnings would not get them one room for one night in the place where we started, at the Hyatt.

There is not much that is surprising about this journey. It could be repeated in countless other cities and regions of the developing world. One could start in Bombay or Bangkok, in north India or

northeast Brazil. The starting points may differ in terms of average income. But such sub-national disparities in economic performance and living standards within and between regions are common, large, sustained, and, often growing. With empirical regularity, we observe high degrees of spatial concentration where a few cities account for much of the national employment and investment.

This is evident if we look at a map of Brazil, or Mexico, or Indonesia, or China, or India, using broad regional definitions (that is, at the state level) or using more fine-grained geographical demarcations (that is, at the district level). This uneven spatial distribution, whereby high income and productivity are concentrated mainly in large agglomerations, translates into differentials in economic performance across cities and regions within a country. Naturally, it raises questions about the growth potential for secondary cities, particularly those in lagging regions, and villages that are far less prosperous than even the secondary cities. This issue is especially important for developing countries as they have relatively lower levels of overall investment. Thus, regions that do not attract dynamic industries are not only characterized by low productivity, but also by lower relative incomes, standards of living, and development indicators such as rates of infant mortality, female literacy, and longevity.

Now, these indicators of development are also predictors of future growth. More educated populations are also more healthy and productive populations. In other words, there is a circular and cumulative relationship between current development and future development. Therefore, these large, persistent, and often growing differences in development standards across sub-national regions raise serious concerns on the development potential of lagging regions and a need to better understand the critical factors that influence the spatial distribution of economic activity within countries. While there has been much academic and policy interest in examining convergence or divergence of inter-regional economic performance over the past two decades, the findings from this body of work are contradictory, narrowly analyzed, and usually do not provide, in the same pages, satisfactory answers on why we observe persistent economic disparities between geographic units and what role is played by policy and political economy to deepen or mitigate these economic differences?

WHY MARKETS AND STATES MATTER

We begin by making an obvious point: the story of modern economic growth is a story of industrialization. Economic growth is a feature of cities and regions which have industrialized, just as the absence of economic growth is a feature of cities and regions which have not industrialized. Notwithstanding isolated cases of cities that are entirely based on a local resource such as a natural attraction or a specialized agricultural product, in general, there is little doubt that the modern city is an industrial city. Even post-industrial growth, which is characteristic of the more developed nations today, is based on the foundations created by industrial growth. These differential growth patterns between more and less industrialized regions have created widely differentiated sub-national spaces leading to significant spatial inequalities at different scales. For instance, in India, the average income ratio of the richest state to the poorest has grown from about under 2.5 to over 3.5 within fifty years of independence.

However, success at industrialization or the failure to industrialize is not determined entirely in the local region in question. Success, failure, and intermediate outcomes are the result of some factors of economic geography such as proximity to resources and markets, transportation costs, localization and urbanization economies, etc. (that is, factors that are market-based) and some factors of political economy such as historical path dependence, policy decisions on infrastructure, exchange rates, land use, globalization, etc. (that is, factors that are strongly influenced by the state). Hence, the fact that the extent and type of industrialization varies over space is best understood and explained in terms of economic geography and political economy. The forces of 'pure' economic geography are market forces, but they are mediated by political economy, or the state. Therefore, industrialization processes have to be understood in terms of the interaction of markets and states.

This is the approach we bring to this analysis of industrialization in India. From independence in 1947, the Indian state has emphasized the role of the manufacturing industry in 'catching up' to the industrial economies of the west. Today, India is a major industrial nation. Its output includes aircrafts, ships, cars, locomotives, heavy electrical machinery, construction equipment, power generation and transmission equipment, chemicals and petrochemicals, pharmaceuticals,

precision instruments, communication equipment and computers. Its industrial output has grown by over 7 per cent per year over the preceding two decades, its GDP has grown by 6.8 per cent per year since 1994, and its economy (in terms of purchasing power parity) is estimated to be the fourth largest in the world.¹ The term 'Made in India' does not yet have global cachet, but for the large domestic economy of one billion consumers, it is an ubiquitous emblem.

This book then is the definitive account of the geography of industrialization in India. Our objectives are threefold:

- (1) To describe and analyse the processes of industrial formation over the long run (especially over the last decade and half), at multiple spatial scales (region, state, metropolis, district, and pin code), and the concomitant rising tide of spatial inequality.
- (2) To show how market forces (economic geography) and state actions (political economy) have contributed to this condition.
- (3) To interrogate some of the most fundamental questions in economic geography—specifically the issues surrounding market access and external economies—and make significant contributions to its theoretical foundations.

To start with, it is necessary to resolve several contradictions in theory and policy. For instance, in the economic literature, typical regional convergence analyses (Barro & Sala-i-Martin 1992, 1995 and followers) are built on the assumption of *decreasing returns to reproducible factors*, where income disparities arising from differences in regional capital/labour ratios diminish over time. Thus, these models predict *convergence* between regions as both trade and factor flows tend to equalize factor prices. However, there is another large literature which argues that increasing returns (often seen as arising from technological or pecuniary agglomeration externalities) are more likely than decreasing returns. These generate a process of circular and cumulative causation leading to decreasing costs of production and continuing concentration. Coupled with the fact that there are significant barriers to labour mobility, these cumulative processes result in conditions in which a significant number of people continue to live in lagging regions. We do not deny the existence of forces that may lead to inter-regional convergence in some cases, but we are persuaded that increasing returns and cumulative and circular processes

are dominant for very long periods. Throughout this book, in different ways, we keep returning to this conclusion.

What does the state do in response to these large and persistent regional inequalities? Most states have opted for interventions to offset some of these market pressures and to promote relatively balanced regional development. This has created a policy tension or contradiction between the market solution of migration or labour flows, that is, 'moving people to jobs', and the interventionist solution of 'moving jobs to people' or promoting capital flows (including a variety of fiscal transfers whether to subsidize credit creation, job creation, employment of local people, or income support) as well as by providing public goods in lagging regions. These policy contradictions are quite apparent. Less obvious are the contradictions between what have been called 'explicit' spatial policies (designed to favour lagging regions) and 'implicit' spatial policies (on exchange rates, import substitution, energy prices, land use, and other arenas) whose impact often overwhelms the influence of explicit policies. Add to these contradictions the dilemma faced by developing nation states in trying to respond to globalization: cities, especially big cities, are the places where products can be made for the global marketplace. Hence, these cities have to be promoted and upgraded, which is typically done at the cost of promoting lagging regions. In short, there are tensions between spatial efficiency and equity.

State policies have an impact that perpetuates the processes of divergence for very long periods. Cumulative causation works in two ways—upward and downward—creating spirals of growth in some regions and either decline or stagnation in others. These are processes that operate over a very long run, partly because they set into motion actions by individuals, groups, and sub-national states which often tend to perpetuate the divergent growth paths. For instance, migrants with human and financial capital leave lagging regions for advanced ones, leaving the poor region poorer and contributing to the growth of the relatively rich region. One of the results is that, growing regions pay more attention to growth, whereas lagging regions focus on the politics of identity and class. It is possible to show, for instance, that the method of revenue collection used by the British colonizers has strongly influenced the terms of the discourse on identity and class in specific regions; this, in turn, has influenced the political economy of regions in the post-independence period.

It is in this context of contradictions (in political economy and economic geography) that we set our examination of industrialization in India. In the remainder of this chapter we spell out the principal theoretical issues which concern us. These issues are relevant not only in India, but for spatial development anywhere. Our review of the theoretical issues is organized as follows:

- (1) First, we turn to the micro-foundations of economic geography and review the main findings and predictions from recent analytic and empirical work in the 'new economic geography' literature as well as the more traditional regional science literature to identify factors that influence the location and growth of economic activity across regions.
- (2) Next, we review the principal strands of the literature on the political economy of regional development by examining the relationship between markets and states as manifested in spatial terms.

SPATIAL CONCENTRATION OF ECONOMIC ACTIVITY

In this section, we review the relevant literature on spatial concentration and growth of economic activity. This review is by no means exhaustive but is aimed at identifying the key issues that emerge from existing analytic and empirical work. Research on location and concentration of economic activity has long been of interest to economists, geographers, planners, and regional scientists (Greenhut and Greenhut 1975; Hotelling 1929; Isard 1956; Lösch 1956; von Thunen 1966; Weber 1929). However, analytic difficulties in modelling increasing returns to scale, marginalized the analysis of geographic aspects in mainstream economic analysis (Krugman 1991a). Recent research on externalities, increasing returns to scale, and imperfect spatial competition (Dixit and Stiglitz 1977; Fujita, *et al.* 1999, Krugman 1991b) has led to a renewed interest in analyzing the spatial organization of economic activity. This is especially true in the case of geographic concentration or clustering.

Models in the 'New Economic Geography' (NEG) literature (see review in Fujita, Krugman, and Venables 1999) allow us to move from the question 'Where will industry concentrate (if it does)?' to the question 'What industry will concentrate where?' These insightful

theoretical models provide, for the most part, renewed analytical support for the ‘cumulative causation’ arguments made in earlier decades on the core-periphery relationship, on agglomeration economies, and on industrial clustering. The main findings from the economic geography literature can be organized in two categories:

- (1) Market access and transport costs
- (2) Agglomeration economies

We review both these factors in this section.

MARKET ACCESS AND THE COSTS OF REMOTENESS

In traditional location models, *production is assumed to take place under conditions of constant or diminishing returns to scale*. Under these conditions, firm location decisions are based on the fact that transportation has costs associated with it. One implication is that industry is likely to spread out to minimize the costs of reaching consumers in different parts of the country. The ‘folk theorem’ of spatial economics (Fujita and Thisse 1996) says *that under conditions of constant or diminishing returns to scale there will be very many small plants supplying local markets*. However, in the presence of increasing returns to scale, firms are able to concentrate production in relatively few locations, and make choices on where to operate (Henderson, Shalizi, and Venables 2001). These models of location choice with increasing returns and imperfectly competitive market structures are developed in the NEG literature.

Krugman’s (1991b) seminal paper shows that increasing returns activities are pulled disproportionately towards locations with good market access. For example, if there are nine locations, in eight of which the share of final expenditure is 10 per cent, and one for which that share is 20 per cent, then other things being equal, more than 20 per cent of manufacturing supply will be met from this larger location. The reason is simply the benefit of having low transport-cost access to this large market in comparison to more expensive access to other markets. This immediately creates a force for the agglomeration of activity. As a disproportionate share of manufacturing is attracted to a location, either the wage rate in the location will increase or labour will be attracted to immigrate—either of which

will tend to increase this location's share of total expenditure still further. The market access effect is sometimes called the 'home market effect', and this combined with labour mobility is the basis of Krugman's thesis.

The extent to which market access enters into the location decision, depends on the level of transport costs. If transport costs are very high, then activity is dispersed. In the extreme case, under autarky, every location must have its own industry to meet final demand. On the other hand, if transport costs are negligible, firms may be randomly distributed, as proximity to markets or intermediate suppliers will not matter. It is only at intermediate levels of transport costs that agglomeration would occur, especially when the spatial mobility of labour is low (Fujita and Thisse 1996). We therefore expect a bell shaped (or inverted U shaped) relationship between the extent of spatial concentration and transport costs (see Figure 1.1).

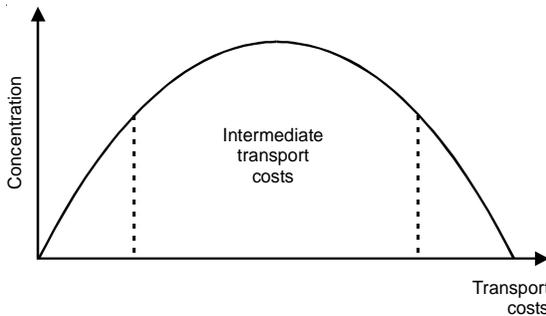


Figure 1.1: Transport costs and concentration of economic activity

Source: Adapted from Fujita, Krugman, and Venables (1999).

In principle, improved access to consumer markets and intermediate buyers and suppliers will increase the demand for a firm's products, thereby providing the incentive to increase scale and invest in cost-reducing technologies. With a decline in transport costs, firms have an incentive to concentrate production in a few locations to reduce fixed costs. Transport costs can be reduced by locating in areas with good access to input and output markets. Thus, access to markets is a strong driver of agglomeration towards locations where transport costs are low enough that it is relatively cheap to supply markets due to availability of quality transport networks (Henderson, *et al.* 2001). In

addition to the pure benefits of minimizing transport costs, the availability of high-quality infrastructure, linking firms to urban market centres increases the probability of technology diffusion through interaction and knowledge spillovers between firms, as well as between firms and research centres, and also increases the potential for input diversity (Lall, Shalizi, & Deichmann 2004). As a result, improved accessibility has the effect of reducing geographic barriers to interaction, which increases specialized labour supply and facilitates information exchange, technology diffusion, and other beneficial spillovers that have a self-reinforcing effect.

AGGLOMERATION ECONOMIES—LOCALIZATION

In addition to market access, firms tend to concentrate production to benefit from localization economies, which are externalities that enhance productivity of all firms in that industry. At the industry level, scale economies accrue to firms due to the size of the industry in a particular location. These economies are external to the firm but internal to the industry. There is considerable theorizing on localization economies in the works of Alfred Marshall (1890), Kenneth Arrow (1962), and Paul Romer (1986); these are often called MAR externalities (from the initials of the primary contributors). They argue that cost-saving externalities are maximized when a local industry is specialized, and their models predict that externalities predominantly occur within the same industry. Therefore, if an industry is subject to MAR externalities, firms are likely to locate in a few cities where producers of that industry are already concentrated. Examples of highly localized industries are ubiquitous. Semiconductor and software in Silicon Valley and automobile in Detroit are classic cases in point. In India, we have had textiles in Mumbai and jute in Calcutta. Later, Michael Porter (1990) emphasized the importance of dynamic externalities created in specialized and geographically concentrated industries.

Benefits from localization include sharing of sector-specific skilled labour, sharing of tacit and codified knowledge, intra-industry linkages, and opportunities for efficient subcontracting. Further, the presence of a disproportionately high concentration of firms within the same industry increases the possibilities for collective action to lobby

regulators or bid down prices of intermediate products (Lall, Shalizi, & Deichmann 2004). These location-based externalities imply that firms are likely to benefit from locating near large concentrations of other firms in their own industry.

In addition to the supply-side linkages discussed above, localization economies are also realized on the demand side. These include reduction of information asymmetries for consumers as well as the ability to attract price and quality comparison shoppers. The existence of auto malls, jewellers rows, bookstore and restaurant enclaves in urban areas are examples of this phenomenon. These so-called 'thick-market externalities' benefit all firms in an industry located in close geographic proximity and can occur in relative isolation from other industries. There is an extensive empirical literature supporting the positive effects of localization economies on economic performance (Henderson 1988; Ciccone & Hall 1995). The benefits of own-industry concentration can, however, be offset by costs such as increased competition between firms for labour and land causing wages and rents to rise, as well as increased transport costs due to congestion effects. Firms in industry sectors which predominantly use standardized technologies and low skilled workers for production may not benefit enough from intra-industry externalities to offset costs from increased own-industry concentration.

In a study of Korean industry, Henderson, *et al.* (2001) estimate scale economies using city level industry data for 1983, 1989, and 1991–93, and find localization economies of about 6 to 8 per cent. Lall and Mengistae (2005) use firm survey data from India and find that own industry concentration has a significant bearing on firm-location decisions across cities, and this effect is the highest for technology-intensive sectors. In another recent study using firm-level data from Indonesia, Deichmann, *et al.* (2005) also find evidence that localization economies are higher for high technology (office computing) and natural resource-based industries (wood and rubber and plastic) and lower for footloose industries such as garments and textiles. Their policy experiments suggest that in the presence of agglomeration economies, increases in infrastructure endowment for lagging regions may have only limited pay-offs in terms of attracting firms from other more established leading regions, particularly in mainstream sectors that have already concentrated in other leading regions.

AGGLOMERATION ECONOMIES—INTER-INDUSTRY LINKAGES

The third force (following market access and localization economies) comes from combining own-industry concentration with the production of intermediate goods. Demand for manufacturing comes not just from final consumers but also from intermediate demand or inter-industry linkages. Therefore, a location with a high share of firms will have a high demand for intermediates, which further increases its attractiveness for manufacturing firms. In addition to these demand effects there are cost benefits; as a large number of intermediate suppliers are attracted into the location, firms using intermediate goods can save on transport costs, making the location still more attractive.

The importance of inter-industry linkages as a major agglomerative force was first recognized by Marshall (1890, 1919). Venables (1996) demonstrated that agglomeration could occur through the combination of firm-location decisions and buyer-supplier linkages even without high factor mobility. The presence of local suppliers can reduce transaction costs and therefore increase productivity. Inter-industry linkages can also serve as a channel for vital information transfers. Firms that are linked through stable buyer-supplier chains often exchange ideas on how to improve the quality of their products or on how to save production costs. It is such on-going interactions that make the dynamics of inter-industry externalities so vibrant. Therefore, if the performance of an industry is highly dependent upon the supply of high-quality intermediate goods (for example in automobile manufacturing), firms are likely to locate in regions with a strong presence of local suppliers. The presence of local supplier linkages makes buyer industries more efficient and reinforces the localization process.

The empirical evidence from developing countries on the importance of intermediate suppliers or inter-industry linkages in influencing location decisions and industrial performance is still in its infancy. Recent work by Amiti and Cameron (2004) shows that externalities which arise from inter-industry linkages are highly localized and have a significant impact on manufacturing performance (measured by wages) in Indonesia. Location models in Deichmann, *et al.* (2005) also find that access to suppliers influences location decisions of firms in several industry sectors (food and beverages,

garments, chemicals, rubber). However, firm-profit models estimated in Lall, Funderburg, and Yepes (2004) for Brazil do not find significant gains from supplier access, when they control for market access and other sources of agglomeration economies. For China, Amiti and Javorcik (2005) find that market and supplier access in the province of entry are the most important factors affecting foreign entry, which is consistent with market fragmentation due to underdeveloped transport infrastructure and informal trade barriers.

AGGLOMERATION ECONOMIES—URBANIZATION ECONOMIES

Scale economies from urbanization, emanate from the overall size (not only in terms of the number of firms but also in terms of population, income, output or wealth) and diversity of the urban agglomeration. For a firm, benefits from urbanization include access to specialized financial and professional services, inter-industry information transfers, and availability of general infrastructure such as telecommunications and transportation hubs. Size is usually correlated with diversity, as larger urban areas can support a wider range of activities. Small cities are specialized in a few manufacturing activities, or are either administrative centres (such as regional capitals or university towns in some countries), or agricultural market centres providing services for farmers. In comparison, larger cities are more diverse, supporting a variety of manufacturing activities that require buyers and suppliers to be in close spatial proximity (input-output linkages). Further, larger cities are centres of innovative technologies and usually tend to offer business or productive services.

The importance of urbanization economies arising from industrial diversity is linked to the work of Chinitz (1961) and Jacobs (1969). In their representation, diversity provides a summary measure of urbanization economies, which accrue across industry sectors and provide benefits to all firms in the agglomeration. They propose that important knowledge transfers primarily occur across industries and the diversity of the local industry mix is important for these externality benefits. They argue that cities are breeding grounds for new ideas and innovations due to the diversity of knowledge sources concentrated and shared in cities. The diversity of cities facilitates innovative experiments with an array of processes, and therefore new products

are more likely to be developed in diversified cities. Therefore, industries with Jacobs type externalities tend to cluster in more diverse and larger metro areas.

The benefits of locating in a large diverse area go beyond pure technology spillovers. Firms in large cities have relatively better access to business services, such as banking, advertising, and legal services. Particularly important in the diversity argument, is the heterogeneity of economic activity. On the consumption side, the utility level of consumers is enhanced by increasing the range of goods that are available locally. At the same time, on the production side, the output variety in the local economy can affect the level of output (Abdel-Rehman 1988; Fujita 1988; Rivera Batiz 1988). That is, urban diversity can yield external scale economies through the variety of consumer and producer goods. Recent empirical studies by Bostic (1997) and Garcia-Mila and McGuire (1993) show that diversity in economic activity has considerable bearing on the levels of regional economic growth.

There is considerable empirical work which examines the contribution of urbanization economies on productivity. In one of the earliest studies, Sveikauskas (1975) used manufacturing data for the US at the two-digit SIC (Standard Industrial Classification) level and found that a doubling of city size increased labour productivity by six per cent. Using Japanese data, Tabuchi (1986) found that a doubling of population density increases labour productivity by 4.3 per cent. The results from empirical studies on the relative importance of specialization and diversity are mixed. Glaeser, *et al.* (1992) find evidence only in favor of diversity. On the other hand, Miracky (1995) finds little evidence to support the diversity argument. For Indonesia, Henderson, *et al.* (1995) show that the significance of diversity is different for different industrial sectors. They find evidence of specialization externalities in mature capital goods industries and of diversity externalities in new high-tech industries. These findings are consistent with the product cycle theory (Vernon 1966), which predicts that new industries tend to prosper in large and diverse urban areas, but with maturity, their production facilities move to smaller and more specialized cities.

SUMMARY

In the final analysis, localization economies, input-output linkages, and urbanization economies are not mutually exclusive. They may occur individually or in combination. For instance, consider whether it is possible to have localization economies without urbanization economies? Imagine a production centre which is populated by fifty small and large firms that produce only one item, say cricket bats. The region produces willow of a type that is well known for being turned into high quality cricket bats. These fifty firms employ on average sixty employees, so the number of factory workers is 3000. If we assume that each factory job has a multiplier effect of 2.5 in the formal sector (that is, for each factory job there are tertiary jobs in finance, education, health care, retail, etc., which total to 1.5 jobs) and 2.0 in the informal sector (transportation workers, domestic servants, etc.), then the total number of workers in this production centre is 10,500. Let us also assume that each worker supports four other individuals. Then we can have a town of 52,500 people that is based on a single product. This centre is too small to realize general urbanization economies. Hence, we can technically argue that it is possible to realize localization economies without urbanization economies. However, such instances are rare and usually based on a local resource (willow in this case; merino wool and champagne are other examples). In most cases several spatial economies operate together and part of the difficulty faced by analysts of such forces is that it is not easy to separate out the effects of each. One of our main tasks in this book is to separate these effects. We use different methods to do so, and in Chapters 4 and 5 we argue that urbanization economies and industrial diversity may be the most important of the external economies as far as industry location is concerned.

POLITICAL ECONOMY OF REGIONAL DEVELOPMENT

Do these market (or economic geography) forces operate without any state intervention? Of course not. The presence of the state is hinted at in the idealized account presented above. Who builds the roads and rail lines that form the transportation networks that are so critical in reducing costs? With rare exceptions, it is the state.

Therefore, from the beginning of development practice and scholarship, questions on regional development have been framed around the institutions of the state and the market: What are the possible outcomes for comparative regional development under free market conditions (with unrestricted factor mobility)? What kinds of state intervention are possible and/or necessary for achieving balanced growth? What are the effects of state intervention on comparative regional development (the equity issue), and on the nation's development prospects (the efficiency issue)? This literature is very familiar to students of regional development. We divide this literature and the knowledge generated by it into two parts; the point of departure is the break created by liberalization and structural reforms in many nations. This break is also associated with globalization.

REGIONAL DEVELOPMENT THEORY— THE PRE-GLOBALIZATION PHASE

Economic models of regional development are built on the general assumption that increasing returns and diminishing returns are both possible in the long run; and since the latter comes later, inter-regional convergence is the likely long-term outcome. These models tend towards equilibrium and convergence, rest on export-driven growth and the economies of agglomeration in dynamic nodal regions, where most regions derive long-term benefits from modernization and technical change (Borts & Stein 1964; Isard 1975; North 1975; Richardson 1973). However, there are contradictions between the economies of scale and agglomeration on the one hand, and size related congestion diseconomies on the other in metropolitan regions (Petraikos 1992; Wheaton & Shishido 1981). Or, in Krugman's (1991a, 1995) terms: there is tension between centripetal forces (higher labour productivity, larger plant size, access to markets and products, that is, backward and forward linkages, thick labour markets, and knowledge spillovers) and centrifugal forces (higher land rents, commuting costs, congestion and pollution, all leading to higher wages and taxes). For indeterminately long periods after industrial development begins, large cities offer increasing returns to capital investment. Eventually, because of lower transportation costs, the costs of size-related congestion rise

above the benefits of concentration-based externalities, so that higher returns become possible in smaller urban centres.

An early culmination of this approach was Jeffrey Williamson's (1965) famous thesis on the inverted-U curve of regional inequality. Williamson used arguments similar to the ones outlined by Kuznets (1955) to suggest that regional inequality increases during the early stages of development, and declines during the later stages. His cross-sectional analysis of regional inequality in countries at different levels of development appeared to provide solid empirical support for his hypothesis, and ever since, the 'Williamson hypothesis' has become one of the cornerstones of the regional development literature. Alonso (1980) went so far as to suggest that the Williamson curve is one of the five bell-shaped curves that almost invariably characterize the development process (see Chakravorty 1994 for a counter-argument).

By the early 1980s, however, conventional neoclassical growth theory (the Solow model) had, in the view of many scholars, become unrealistic in that it treated technological change and human capital as exogenous factors. As a result, a 'new' growth theory that treats these variables as endogenous to the growth process has gained legitimacy. After the pioneering, non-spatial work of Romer (1986) and Lucas (1988), it has been extended by Barro and Sala-i-Martin (1995), Armstrong (1995), Sala-i-Martin (1996) and others to the regional realm (where, it is presumed, the model assumptions are more reasonably satisfied). There are several competing and overlapping versions of such endogenous models, and as shown by Martin and Sunley (1998) their theoretical expectations, from a spatial or non-spatial perspective, are contradictory.

The models predicting convergence were not, however, the first perspectives on comparative regional development. Some pioneers in development economics—Gunnar Myrdal (1957) and Albert Hirschman (1958)—had been skeptical about the growth prospects of lagging regions. They suggested the *core-periphery* and *cumulative causation models* which Kaldor (1970), Friedmann (1966, 1973), Perroux (1950), Boudeville (1966) and others subsequently extended. In this view, because the natural forces of economic geography tend to favour the existing leading regions (the core), they would tend to grow faster than the existing lagging regions (the periphery). Therefore, if there is no state intervention, regional imbalances are

likely to widen. Because of demands from the periphery, state intervention is not only politically necessary and inevitable; it also improves the distribution of welfare. Myrdal, however, was frankly pessimistic about the prospects of lagging regions. According to him:

It is easy to see how expansion in one locality has 'backwash effects' on other localities. More specifically the movements of labour, capital, goods, and services do not by themselves counteract the natural tendency to regional inequality. By themselves, migration, capital movements, and trade are rather the media through which the cumulative process evolves—upwards in the lucky regions and downwards in the unlucky ones. In general, if they have positive effects for the former, their effects on the latter are negative. (Quoted in Higgins & Savoie 1995: 86).

A more hopeful position was taken by Friedmann and Hirschman—they saw the core as the locus of change, where new ideas, technology, and capital intersect to generate economic and cultural dynamism. Granted, they argued, that the non-metropolitan periphery initially falls behind, but eventually expanding markets and urbanization, the spatial diffusion of innovations and culture, and political demands from the periphery (mediated by state actions) should lead to some narrowing of the core-periphery gap.

Marxist and Neomarxist models of the political economy of regional development have been far more critical of the bourgeois state. Dependency theorists such as Paul Baran (1957), Andre Gunder Frank (1967), Michael Timberlake (1987) and others have argued that deepening class polarization and geographical inequality are the outcomes of modernization and industrialization. In this view, sub-national uneven development is a less important subset of the fact of deepening uneven development between nations. The core-periphery system that exists at the intra-national scale also exists at the international scale. The periphery states are dependent on the core states for technology and demand for goods. The elite in these dependent-periphery states assist the capital owners from the developed core to extract surplus, and the underdevelopment of the regional and international periphery is a necessary condition for the development of the core.

In the more subtle views of Harvey (1982) and Massey (1984), regional change is episodic, where neither divergence nor convergence holds over the long run. More recently, de-industrialization or post-

industrialization in developed nations has been theorized under the perspectives of 'post-Fordism' or 'flexible accumulation' (see Piore & Sabel 1984; Scott 1988; Storper & Walker 1989). Here the emphasis is on transactions costs and 'the character of technological change, the form and organization of firms and industries, (and) the creation and transformation of labour markets' in influencing regional change (Schoenberger 1989, p. 133). These ideas have been extended to the context of third world development (Diniz 1994; Storper 1991), but there has been little systematic or system-wide extension of this approach. The followers of Michael Lipton and his 'urban bias' thesis (Lipton 1977) were just as pessimistic as the Neomarxists. In their view, state policies are designed to benefit urban areas at the cost of rural areas: to extract the rural surplus and invest it in more productive urban enterprise. As a result, lagging rural regions fall even further behind the growing urban regions.

Hence, the role of the state in influencing urban and regional development has been theorized from several perspectives—from 'urban bias' and 'dependency' theorists who have argued that policies are created with the intent to favour urban areas or metropolitan centres, to the view that state interference distorts market-determined spatial distributions (of population and investment) in unintended and unforeseeable ways, or has little influence (Brooks 1987; Henderson 1988). The set of 'explicit' urban and regional policies aimed at industrial decentralization—licensing control, location of public sector projects, promotion of industrial estates, tax and land acquisition incentives, etc., in favour of lagging regions—has often been subsumed by more powerful macro-economic policies (which may be called 'implicit' spatial policies) on development path, subsidies to industry relative to agriculture, monetary, exchange rate, and trade policies, etc. (Henderson 1982; Mills 1987). Despite the overall failure of regional policies (which are subsumed by implicit spatial policies), they doubtless have had some effect: perhaps one only of retardation rather than reversal of metropolitan polarization.

To summarize the confusion: regional theory suggests that regional differences are likely to widen in the absence of state intervention (Myrdal), which is not necessarily a negative outcome (Friedmann), as, in the long run regional differences will decline anyway (Williamson, Barro & Sala-i-Martin). State intervention is a necessary aspect of the political process (Hirschman), but is biased toward urban

areas (Lipton) and developed nations (Baran, Frank), and may lead to inter-regional convergence (Sala-i-Martin, Krugman) or is inefficient or irrelevant or dwarfed by 'implicit' regional policies (Henderson, Richardson).

REGIONAL THEORY AND GLOBALIZATION

There is widespread agreement among analysts that economic globalization has ushered in a new phase in inter-regional development. Economic globalization is marked by increasing international trade, a more rapid and voluminous flow of goods, services, information, and investment across international boundaries and ever longer distances, and a new international division of labour, where routine manufacturing takes place in developing nations, and design and control functions are undertaken in more developed nations. Economic globalization is accompanied by, and largely made possible by, ideological globalization, whereby developing nations adopt similar sets of policies that enable trade (Chakravorty 2003). These policies have been variously termed: the Washington Consensus (Williamson 1990, 2000), liberalization, or structural reforms. We do not need, at this point, to understand the differences between these terms.

This shift to market-friendly liberalism over the last decade-and-a-half signifies an important paradigm shift. Economic and political nationalism was the idea that led to the creation and continuance of the nation-state in the post-colonial world. That idea appears to be losing ground to the twin forces of economic and ideological globalization. The role and nature of the state is also changing as a consequence. It is necessary to focus on the relationship between the new, reformed state and its policies, and the resulting development impact at the sub-national level. Historically, the state has been instrumental in shaping the economic geography of regions in the developing world—starting with the establishment and privileging of port cities for external trade and administration during the colonial period, to the creation of a complex array of rules and regulations that established location incentives and disincentives during the nationalist period.

The critical questions after reform are: what is the role of the state in shaping the economic landscape after the nature of the state

has been altered to a (supposedly) less interventionist version; and, how do domestic and international actors respond to the freer markets, especially in terms of investment location? One assumption generally shared in the regional development literature we have outlined in the previous pages is that of policy continuity, if not regime continuity. It is apparent now that this assumption is invalid in most situations: liberalization or structural reform is a fundamental shift away from the policies of the past. As a result, it is now necessary to formulate a new theoretical framework for the analysis of regional development.

What are the general policy imperatives of the post-reform state? In the old nationalist model, the national state tried to be the principal agent of economic change using an institutional and regulatory structure that emphasized centralization over federalism, state ownership of heavy industry and infrastructure over private ownership, and self-reliance or import substitution over export orientation. In the new liberal model, the state is significantly less involved in the ownership of industry and the regulatory structure affecting new investments; there are lower entry barriers to multinational capital; export orientation is favoured over import substitution; and steps toward some decentralization of power and policy instruments in favour of sub-national states are taken.

But the 'new' state is far from a hands-off, free-trading, *laissez faire*, purely market-enthusiastic entity. One of our basic premises is that the neoliberal nation-state is bipolar in more ways than one—for instance, 'there is a coexistence of liberalizing and protectionist policies' (Leinbach 1995, p. 204)—which leads to inaction in some arenas while simultaneously there is more concerted action in other arenas (see Wade 1990 and Brohman 1995 on this thesis). As far as regional development is concerned, the newly liberal state is both a reduced or spatially disengaged state (as far as the promotion of regional balance is concerned), and a more enlarged state in terms of promoting selected metropolitan regions for receiving investment, especially foreign direct investment (FDI). These changes are concomitant with more active sub-national states competing in asymmetrical spatial structures shaped by colonialism, and subsequently nationalism, to capture new and different markets. The altered geography of investment opportunities suggests altered inter-regional development possibilities. The fact that structural reform is

a discontinuity in the development process is, however, merely one of the complications: regional development theory, which was already beset with contradictions from different ideological perspectives (as shown in the foregoing discussion), has been further complicated with the regional implications of the increasingly dominant new or endogenous growth theory. Let us look at the issues from a pragmatic, political economy perspective.

What are the significant factors affecting regional change in developing nations after liberalizing reforms? Elizondo and Krugman (1992) suggest that post-reform regional development is likely to be more evenly balanced. They argue that the magnitude of internal trade is much larger than foreign trade in inward looking trade regimes; 'this leads to concentration of production and trading activities in large metropolitan cities... an opening up of the economy is likely to break the monopoly power of these highly concentrated production and trading centers, weaken the traditional forward and backward linkages and lead to a more even distribution of economic activities across regions.' (Das & Barua 1996, p. 365). Similarly, according to Gilbert (1993, p. 729), 'the cities which benefited most from the previous development model have suddenly had an important prop to their growth removed' in the new model of liberalization and export orientation. These could be valid arguments, of course, but should be considered in the light of actual experience—much depends on the extent of urban and metropolitan bias before reforms, the size of the domestic market, the nature of the regional hierarchy, the degree of protectionism, the quantity of international trade, the specific liberalization strategies adopted, etc. (see Markusen 1995).

In developing nations like India, perhaps the most important structural factor to consider is the availability of infrastructure. Metropolitan regions have, by far, the highest standards of physical infrastructure (in power, roads, housing, telecommunication, etc.) and social infrastructure, such as schools and hospitals (Kessides 1993). During the stage of state-controlled development, this infrastructure concentration factor was somewhat offset for two reasons. First, the state itself was the primary decision maker on where much of the capital was invested; earlier, significant proportions of this investment had gone to infrastructure and heavy industry near raw material sources (in what have often been called growth centres) in non-

metropolitan locations. Often these centres became enclaves, linked to the national and international economy rather than the local economy, but industry was decentralized somewhat. Second, as indicated earlier, incentives in lagging regions and disincentives in metropolitan regions raised the former's share in total investment beyond what a no-policy situation would have allowed.

After market-oriented reform these two factors are drastically changed. The state moves away from industrial ownership, and therefore location decisions, by allowing private industry to participate in all industrial arenas, and by eventually divesting state-owned industry to the private sector. Egalitarian regional policies are withdrawn or not enforced. At the same time, following the East Asian model, the national state sees foreign investment as the key to spurring economic growth; it also logically sees its metropolises as likely foreign investment destinations. It invests in infrastructure in the leading metropolises, and encourages competition between cities and regions for other investments (Cook & Hulme 1988). Sub-national governments, with more freedom to enact policies, tend to react by further eroding local environmental and egalitarian policies, and by emphasizing the growth prospects of the largest cities. The coastal regions in general, and the coastal cities in particular, become the focal points of directed investments, as in the new, outward-looking regime, access to the outside world from coastal locations assumes greater importance.

Finally, the state takes action to reconfigure the metropolis to accommodate the new growth. Many of the concerns with primacy and the advocacy of decentralization stem from the perception that the existing metropolitan cores are very congested, and, therefore, unmanageable and inefficient. Clearly, growth has to be accommodated elsewhere within the metropolitan area. High-technology sub-centres may emerge in existing or newly created satellite townships, small- to medium-scale manufacturing establishments may locate in the existing industrial suburbs, whereas large scale manufacturing may find new locations on the edge of the metropolis. Here, in a federal structure, the role of the local state becomes important—it identifies or designates industrial or technology parks, and high-tech or export processing zones, and provides basic infrastructure to attract new industry to these locations.

Eventually, as the evidence from mega cities around the world seems to indicate, the traditional monocentric city is replaced by larger polycentric urban regions—at an accelerated pace as a result of the reforms (Diniz 1994; Dökmeci & Berköz 1994, Koppel & McGee 1991, Richardson 1988; Ginsburg,).

THE SPREAD OF INDUSTRY

If unrelenting centralization and concentration of industry was the only reality of economic geography, then convergence between regions would never be possible, and we would not see the rise of new cities after the beginning of industrialization because the initial cities would hold the advantage forever. We know that inter-regional convergence is a reality in several nations, especially in the special cases of the 'new' world: the colonies of Europe in the US, Australia, Canada, etc. Moreover, as we have seen in the preceding sections, policy interventions in favor of lagging regions are implemented almost everywhere. In this section, we provide some examples where industry has relocated—either de-concentrated or re-concentrated—following selected policy interventions.

The unequal division of manufacturing across regions is associated with corresponding wage and land rent differences. Firms considering a move out of the agglomeration would pay lower wages and rents, but forego the benefits of proximity to markets and to intermediate suppliers. Empirical evidence, though scanty, tends to show that spatial decentralization to regions with small employment bases almost never happens. Much more common is de-concentration (what we call 'concentrated decentralization' in Chapter 2) from the core to the periphery of metropolitan areas, and this happens with improvements in the transportation system. For example, Henderson, *et al.* (1996) show that many firms moved out of Jakarta to the peripheral areas of the Greater Jakarta metropolitan region in the mid-1980s. These moves were facilitated by the construction of toll ring-roads around the city, retaining some of agglomeration benefits of the region, but reducing congestion costs (for example, land rents and transport costs), enabling firms to benefit from lower land and labour costs in the periphery, which exceeded the increased costs of transportation for serving the same market. Aggregate transport costs per unit of sales

revenue also dropped as a larger market could be accessed by a better network. Similarly, for Brazil the de-concentration of industry from Grande São Paulo to lower wage hinterland cities followed the transport corridors first through São Paulo state and then into Minas Gerais, the interior state with the main iron ore and other mineral reserves (Henderson, *et al.* 2001).

While the limited evidence shows that inter-regional transport improvements allows firms to relocate from core metro areas to their peripheries, there is no convincing evidence to show relocation to small urban centres in lagging areas. For China, Head and Reis (1996) show that since the implementation of an open door policy in 1978 that ended the prohibition of foreign business and investment, foreign firms preferred to locate in cities with large industrial bases and established foreign investment presence. Their econometric work controls for factors such as provision of fiscal incentives, and availability of infrastructure, which makes some regions inherently more attractive than others. We find a related pattern of re-concentration of manufacturing away from Mexico City to northern cities such as Ciudad Juarez, Monterrey, and Tijuana, which are physically close to the United States following the opening of the Mexican economy to foreign trade and investment (Hanson 1998). Since 1980, industrial activity in Mexico has moved to states on the US–Mexico border, reducing the importance of Mexico City as the nation’s main industrial center. Between 1980 and 1993, the border states increased their share of manufacturing employment from 21 per cent to almost 30 per cent, and Mexico City’s share of manufacturing employment declined from over 44 per cent to under 29 per cent. Meanwhile, there have been few new investments in states like Chiapas, Oaxaca, and Campeche, the lowest income states in Mexico. In Chapter 2, we show that a similar process of concentrated decentralization has taken place in India.

PLAN OF THE BOOK

This is the background against which we locate this examination of industrialization in India. Any serious analysis of industrialization must consider the influence of both the market (that is, economic geography) and the state (that is, political economy). This work

integrates the interplay of the market and the state in creating the evolving economic and industrial geography of India. We emphasize the following features:

- A spatial perspective: We base our arguments and conclusions on the analysis of detailed data collected at several spatial scales—the state, the metropolis, the district (there are about 500 districts in India), and the postal pin code (there are more than 100 pin codes in each metropolis). This allows us to study the issues identified earlier in this chapter at multiple scales, from the neighbourhood, through the metropolitan region, and the state, to the nation.
- The long run: We incorporate data from independence and nationalism into structural reforms and liberalization in the 1990s. For much of the book, we use the structural reforms as the point of departure, but we are always mindful that actions taken today have causes that often go back fifty to 200 years.
- Innovative methods and micro-level data: We use current spatial analytical techniques (spatial autoregressive models, Local Indicators of Spatial Association, etc.) in combination with appropriate modelling methods (linear, logistic, and non-parametric models) to tease out the stories from the data. The data we use are disaggregated and detailed—including multi-scale spatial data and firm-level production data—allowing us to build theory from spatial and economic micro-foundations.
- Policy orientation: One of our major goals is to inform and improve policy. We analyze how specific policies have had consequences on the ground, and show how the policy framework needs to be changed in order for there to be more balanced, equitable, and sustainable industrial growth.

Chapter 2 ‘Patterns of Industrial Investment, Old and New’ describes and quantifies the spatial distribution of industry at multiple scales (states, districts, metropolitan areas) in the pre-reform and post-reform periods. We also present data from the early post-independence period. We identify and map the emerging economic geography of India and highlight the leading edges and lagging pockets. We show the continuing significance of history (marked by investments in existing industrial areas) and the significance of geography (marked

by investments in clusters in and around existing or new industrial areas). We show that the location of post-reform investment favours the coast, advanced regions, and existing metropolises (especially the edge areas); these realities are truer for foreign direct investments than domestic investments (especially the direct investments of the state). The results provide evidence of the return of cumulative causation and divergence.

In Chapter 3 'Determinants of Industrial Location' we identify the factors that influence industrial location decisions. These include variables representing capital, labour, infrastructure, regulation, and geography. We show which of these factors are important and to what extent. We also show that source of capital is the primary cause of spatial divergence in investments. Private capital is profit-oriented and directed towards leading industrial regions, coasts, and metropolises, and away from socialist governments. State industrial investments have some regional equity considerations, and therefore are less biased towards leading regions. The results are established using logistic and OLS (Ordinary Least Squares) regression models on district level data.

We continue the analysis begun in the previous chapter, in Chapter 4 'Economic Geography and the Firm' by looking at what factors influence location decisions at the firm level. More specifically, we ask the question: for the manufacturing industry, what are the externalities that matter, and to what extent? We develop an innovative methodology to analyze the influence of economic geography on the cost and wage structure of manufacturing firms. We analyze eight industrial sectors (food/beverages, textiles, leather, printing and publishing, chemicals, metals, machinery, electrical/electronics) by firm size in India, and find that industrial diversity is the only economic geography variable that has a significant, consistent, and substantial cost-reducing effect for firms, particularly small firms. We discuss the implications of these findings for regional growth and development.

In Chapter 5 'Industrial Clusters Within Metropolitan Regions', we turn our attention to the internal structure of metropolitan areas and examine the distribution of industry at this scale. The large and growing literature on industrial clustering suggests that firms seek locations that provide localization economies (benefits from having

common buyers and suppliers, a specialized/skilled labour pool, and informal knowledge transfers). We show instead that industry location decisions are guided by market imperfections, specifically rigidities in the land market caused by state action (segregationist/environmental policies, the absence of exit policies, and activist industrial promotion policies). We use geographically disaggregated industry location and size data (at the pin code level) from Mumbai, Calcutta, and Chennai, to analyze the eight industry sectors we study in Chapter 4. We test for evidence of global and local clustering, and distinguish between and test for co-clustering and co-location of industries. The results are indicative rather than absolute, and suggest that for location decisions, general urbanization economies are more important than localization economies.

Finally, in Chapter 6 'On Spatial Policy' we summarize the key findings and analyze them in the context of state actions and inactions. We begin by looking at a range of spatial policy initiatives used in international settings. Then we consider the question of why these spatial policies have had such limited effects and identify four categories of answers: (1) that spatial economies are stronger than the fiscal incentives provided by policies; (2) that state policy is fundamentally contradictory—what one hand gives the other takes away; (3) that there is little coordination between policies and that they are ad hoc rather than being based on clear analysis; and (4) that institutions have deep effects that cannot easily be overcome by marginal policies. We end with a discussion on one of the most important questions in contemporary India: can anything be done to bring new investments to Bihar?

NOTE

1. There are many sources that contain basic data on the Indian economy. A reliable one is the portal economywatch.com and its India-related pages. Last accessed 30 August 2006.