Studies in Indian Urban Development

Edwin S. Mills
Charles M. Becker

with a contribution
by Satyendra Verma

Published for The World Bank
Oxford University Press

Oxford University Press

NEW YORK OXFORD LONDON GLASGOW
TORONTO MELBOURNE WELLINGTON HONG KONG
TOKYO KUALA LUMPUR SINGAPORE JAKARTA
DELHI BOMBAY CALCUTTA MADRAS KARACHI
NAIROBI DAR ES SALAAM CAPE TOWN

First Printing July 1986
Cover photo by John Courtney

Library of Congress Cataloging-in-Publication Data

Mills, Edwin S.
Studies in Indian urban development.
Includes bibliographies and index.
1. Cities and towns—India—Growth. 2. Economic development—India. 3. City planning—India.
I. Becker, Charles M., 1954- . II. Title.
HT147.I5M56 1986 307.1'4'0954 86-2489
ISBN 0-19-520507-3

Contents

Preface

The chapters in this book report research sponsored by the World Bank on urban development in India. Their purpose is to provide a broad framework of historical, international, and regional findings within which a wide range of issues related to Indian urbanization can be studied. Much has been published about India's economic problems, but no publication provides a comprehensive and coherent analysis of India's urbanization. We hope this book fills that gap. We hope it will be of interest to urban and regional economists, to students of economic development in India and elsewhere, and to government and international officials concerned with India's economic development.

The book is basically positive economics. We attempt to identify historical trends, to explain regional differences, and to compare India's experience with that of other developing countries. For the most part, we avoid normative analysis. However, no country urbanizes without encountering problems of government policy and social stress. And in a country as poor as India, and in which government's influence is as pervasive as it is in India, such problems are inevitably severe, even traumatizing. In two places, our discussions naturally extend to a consideration of the effects of certain actions of the Indian government: chapter 7 examines government attempts to influence the size and the growth rates of India's largest cities, and chapter 8 takes a look at government attempts to alter the personal income distribution. Even in those discussions, however, our goal is to suggest relevant and constructive ways to think about the problems, not to present specific conclusions about Indian government policy.

Chapter 1 provides an introduction to the Indian economy. It traces the progress of India's economic growth during the twentieth century and presents a brief description of the government's institutional framework for economic planning.

Chapter 2 examines the relation between urbanization and economic development on a worldwide basis. In an elaborate cross-country analysis, we relate urbanization to national economic conditions and other factors. How closely does urbanization follow economic development? What measures of development are most closely associated with urbanization? Is India's experience analogous to the worldwide relationship? We show that urbanization relates closely to widely available indices of economic development and that India's experience has paralleled the worldwide pattern in recent decades. We conclude the chapter with projections of Indian urbanization to 1990.

Chapter 3 shifts from worldwide comparisons to comparisons within India at various times in its history. We trace the course of Indian urbanization during the twentieth century and relate it to India's economic development. Has urbanization correlated closely with economic development during this century? Was the relationship different before and after India gained independence? Before independence, both economic growth and urbanization were slow in India. But urbanization did continue even when economic growth was negligible. Since independence, both economic growth and urbanization have accelerated and have proceeded at similar paces.

Chapter 4 presents a study of Indian city sizes and city growth. Has India's urban population become more concentrated in its large cities as the economy has developed? Was the trend different before and after independence? Is there evidence that government programs to curtail the growth of large cities have been effective? Are there important regional differences in the relative sizes of India's cities? How does one explain the historical trends and regional differences in Indian city sizes? We show that stability and slow change have been the salient characteristics of the size distribution of India's urban areas. There is almost no evidence of any long-term trends in the distribution during the twentieth century. Nor is there evidence of the effects of government policies to reduce the size and the growth rates of large cities. The chapter concludes with a study of the determinants of growth of individual cities. Attention is focused on whether a city's industrial structure is an important determinant of its growth and on whether large cities tend to grow faster than small cities.

In chapter 5, the focus narrows to the state of Madhya Pradesh, where the effects on migration and population growth of government-provided services and of sectoral sizes can be measured. The goal in this chapter

Table 1-1. *Real Income Per Capita, 1868–1945*

Year	*Real income per capita (rupees, 1946–47 prices)*	*Average annual growth rate (percent)*
1868–69	120	—
1872–73	125	1.0
1890	134	0.9
1900	145	0.8
1910	154	0.6
1920	164	0.6
1930	171	0.4
1935	166	−0.6
1940	169	0.4
1945	166	−0.4

— Not applicable.
Note: Incomes deflated by source.
Source: Lal (1981).

were about one-half their 1980–81 level, and the figures in the table imply that living standards in 1868–69 were about three-fourths their 1945 level. Even in 1945, living standards must have been close to the subsistence level for a large percentage of Indians. Yet, if one defines the subsistence level as the living standard at which population growth is zero without constraints imposed by modern birth control, then average Indian living standards were above subsistence throughout the three-fourths of a century covered by table 1-1. Estimates indicate that the population of what is now the Indian Union grew consistently after 1800 (Lal 1981, p. 168). The only decade in which population did not grow was 1911–21 (see table 3-1 in chapter 3), when India was ravaged by the worldwide influenza epidemic that followed World War I.

Real income grew at moderate and decreasing rates from 1868–69 to 1930. From 1930 to 1945, the years of worldwide depression and World War II, there was no growth. Real income per capita was even lower in 1945 than it was in 1930.

Since 1950, the Indian economy has been increasingly well documented. Table 1-2 presents real net national product per capita at decade intervals and its growth rate for the thirty years between 1950–51 and 1980–81. Indian living standards are still among the lowest in the world: in 1980, India's gross national product (GNP) per capita was US$240, close to the average GNP of the twenty-eight countries the World Bank classified as low income in 1980.[1] Indian real incomes, however, grew significantly and steadily during the thirty years covered by table 1-2. The average

Table 1-2. *Real Net National Product Per Capita, 1950–51 to 1980–81*

Year	Real net national product per capita (rupees, 1970–71 prices)	Average annual growth rate (percent)
1950–51	466.0	—
1960–61	558.8	1.8
1970–71	632.8	1.2
1980–81	696.3	1.0

— Not applicable.
Note: Net national product figures deflated by source, using official price index.
Source: Government of India (1982).

annual growth rate during that thirty-year period was 1.3 percent. Comparison of tables 1-1 and 1-2 indicates that India—which gained its independence in 1947—never achieved that growth rate during the colonial period, except, perhaps, in isolated years. Table 1-2 implies that average living standards improved 50 percent between 1950 and 1980. The postindependence growth rate has certainly been four or five times the growth rate during the first half of the century. Although not spectacular by developing-country standards, India's achievement since independence has been substantial and steady. India's 1960–80 growth rate of GNP per capita exceeded that for the average of the countries classified as low income by the World Bank.

The data in table 1-2 offer disturbing evidence of a steadily declining growth rate. The growth rate during the 1970s was barely more than half of what it was during the 1950s. The natural assumption might be that growth would be slower at the very low income levels of the 1950s than at the higher levels of the 1970s. Slower growth at low income levels is indicated by growth models that include various low-income traps. Of course, growth deceleration was common throughout much of the world during the 1970s because of energy crises and other events. But India's growth decelerated during the 1960s, before many other countries—developing and developed—showed signs of economic deceleration.

Social Indicators

By no means all the benefits of economic development are captured in the national income accounts. This section presents a brief review of evidence of improvements in Indian living standards that are not, or not completely, captured in the national income accounts.

There is no widely accepted set of estimates of birth and death rates for India. The birth rate probably changed little from 1900 until independence, but it has dropped steadily since independence. The death rate declined steadily after the 1920s and has dropped rapidly since independence (Mitra 1978, p. 492). Life expectancy fell for both males and females from 1881 to 1921; it was estimated at 20 years in 1921 (Government of India 1978, p. 28). Life expectancy increased steadily after 1921 and has increased rapidly since independence. In 1971 it was 46.4 years for males and 44.7 years for females. The World Bank (1982, p. 150) estimated the combined life expectancy of males and females at 52 years for 1980. This record not only indicates rapid improvement, but also places Indian life expectancy well above the average life expectancy in low-income countries other than India and China.

Comprehensive measures of morbidity are not available for India or for most other countries. Mitra's exhaustive survey (1978), however, makes it impossible to doubt that considerable progress has been made. The incidence of nearly all waterborne and communicable diseases seems to have decreased since independence. Similarly, famine and malnutrition are less common, partly because food-grain production per capita has increased somewhat and partly because the storage and distribution systems have been improved.

India's literacy rate has improved steadily during the twentieth century. It was 5.35 percent in 1901, 16.67 percent in 1951, and 36.17 percent in 1981. Until 1921 the male literacy rate was more than ten times the female rate. By 1981 the female rate was, for the first time, more than half the male rate (Government of India 1981, paper 1, p. 43). India's adult literacy rate exceeds that for the average of low-income countries other than India and China (World Bank 1982, p. 154).

India's social indicators have not improved as rapidly as have those in countries with more rapid economic growth. But by every measure available, India's improvement has been steady and substantial since independence and, in some cases, even before independence.

Industrialization and Investment

In all countries, economic growth is accompanied by shifts of workers and output from agriculture to manufacturing, services, and other predominantly urban sectors. In this section we trace the sectoral shift in India. This shift is related to urbanization in chapter 3.

Basic Indian employment data are divided into ten sectors. The analysis in this section is carried out with a classification into five sectors: agriculture (including forestry and fishing), mining and quarrying, manufacturing, construction, and services. This level of aggregation is necessary to permit comparable employment and value added statistics to be presented. All five sectors are diverse, but the service sector—including utilities, defense, public administration, and various personal and business services—is perhaps the most diverse.

Basic data on the sectoral composition of Indian employment are available for the census years since 1901, except 1941, when sectoral data were not published because of wartime exigencies. The summary data presented in table 1-3 point up the extraordinary stability of the sectoral distribution of Indian employment during the seventy years covered. Agriculture, manufacturing, and services accounted for at least 94 percent of the total employment every year. In 1971 none of these sectors' shares of total employment differed from its corresponding 1901 share by more than 1.2 percentage points. The table reveals almost no evidence of long-term trends in sectoral shares. Agriculture's share rose erratically from 1901 to 1951 and has declined erratically since 1951. Manufacturing's share declined from 1901 to 1951 and has, for the most part, increased since then. Services' share declined from 1901 to 1921 and has increased in most census years since then.

India's sectoral distribution of employment is similar to that of other low-income countries. For 1980 the World Bank (1982, p. 146) reported the following labor force percentages in the three main sectors for India and for a weighted average of low-income countries other than India and China:

	India	*Average of low-income countries other than India and China*
Agriculture	69	73
Industry	13	11
Services	18	16

Although the share of the labor force in agriculture fell more in the average low-income country from 1960 to 1980 than it did in India, India's 1980 share was below that for the average low-income country.

Table 1-4 shows shares of gross domestic product (GDP) among the five sectors for which table 1-3 shows employment shares, for selected years from 1950 to 1977. Agriculture's GDP share decreased substantially and almost steadily while the shares of manufacturing and services rose substantially and almost steadily. Mining and quarrying and construction also recorded increases in their small GDP shares.

For 1980 the World Bank (1982, p. 114) reported the following GDP

Table 1-3. *Distribution of Employment by Sector, 1901–71*
(percent)

Sector	*1901*	*1911*	*1921*	*1931*	*1951*	*1961*	*1971*
Agriculture	69.4	70.4	72.1	70.6	75.1	69.5	69.7
Mining and quarrying	3.9	4.6	4.2	4.8	2.2	2.8	2.9
Manufacturing	10.1	9.6	8.8	8.5	7.5	10.6	9.6
Construction	1.0	1.2	1.0	1.1	0.8	1.1	1.2
Services	15.5	14.2	13.9	14.9	14.3	16.1	16.7
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Note: Columns may not add up to 100 because of rounding.
Sources: For 1901–31, Lal (1981, p. 174); for 1951–71, Government of India (1951, 1961, 1971).

Table 1-4. *Gross Domestic Product by Sector of Origin, 1950–77*
(percent)

Sector	*1950*	*1955*	*1960*	*1965*	*1970*	*1975*	*1977*
Agriculture	59.0	57.3	54.0	45.6	47.1	44.8	43.0
Mining and quarrying	0.7	0.8	0.9	1.1	1.0	1.2	1.2
Manufacturing	9.9	11.1	12.2	15.3	14.4	14.8	15.4
Construction	4.2	4.0	4.4	5.5	5.4	4.9	5.7
Services	26.1	26.8	28.5	32.6	32.1	34.3	34.9
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Note: Columns may not add up to 100 because of rounding.
Source: World Bank (1980, p. 102).

shares in the three main sectors for India and for an average of low-income countries other than India and China:

	India	*Average of low-income countries other than India and China*
Agriculture	37	45
Industry	26	17
Services	37	38

As with the sectoral labor force data presented in table 1-3, these data show that India's GDP shares are not out of line with those of other low-income countries. The important difference is that India has a larger GDP share in industry and a smaller share in agriculture than does the average of low-income countries.

The data in tables 1-3 and 1-4 imply certain trends in the relative labor productivity of the sectors. In 1950, manufacturing had almost equivalent shares of employment and GDP, which indicates that manufacturing value added per worker was about the same as that for the economy as a whole. This appears to be in contrast with the experience of most other low-income countries at the time. In India by 1970, manufacturing's GDP share exceeded its employment share by 60 percent, which indicates that labor productivity had risen much faster in manufacturing than it had in the economy as a whole during the two decades. Labor productivity in agriculture was below the average for the economy even in 1950, since agriculture's labor share exceeded its GDP share even then. But labor productivity rose less rapidly in agriculture than in the economy as a whole during the succeeding two decades, since by 1970 agriculture's GDP share had fallen more than its labor share. Labor productivity in services was about twice the average for the economy as a whole during the two decades the tables have in common. The same comparisons show that labor productivity rose more rapidly in construction and less rapidly in mining and quarrying than in the economy as a whole.

These findings are broadly consistent with those of Chenery and Syrquin (1975, p. 52), who found that in the lower-income countries labor productivity was about 1.5 times the average in industry and services and about 70 percent of the average in the primary sector.[2] But the substantial increase in relative labor productivity in Indian manufacturing appears to be contrary to the findings of Chenery and Syrquin for typical low-income countries.

In summary, during the period from 1950 to 1971, labor force migration out of agriculture was slower in India than it was in most other low-income countries. But GDP shares of manufacturing, construction, and services grew more rapidly in India than in most other low-income countries. For low-income countries other than India and China, the combined share of manufacturing, construction, and services in GDP rose from 51 to 55 percent (World Bank 1982, p. 114); for India, the rise was from 50 to 63 percent. The implication is that labor productivity has grown rather rapidly in these predominantly urban sectors. This suggests that there should have been a larger labor force shift from agriculture than has taken place. We return to this issue in chapter 3, where we analyze the implications of industrial shifts for urbanization.

The data in table 1-3 extend only to 1971. At the time of writing, only preliminary data have been published from the 1981 census, and the lack of detail makes them noncomparable to the data used in our analysis. Furthermore, in the 1981 data only main workers (those who have worked a major part of the year) are classified.[3] Comparable data for 1961 and 1971 are published in the 1981 census, so comparisons among the three years are made here.

Table 1-5 presents main-worker data for agriculture and all other sectors for 1961, 1971, and 1981. It is comparable with table 1-3. The major point to be made about table 1-5 is that the 1981 figure for agriculture's employment share is the smallest share recorded for that sector in India's history. The data in tables 1-3 and 1-5 suggest quite a large drop in agriculture's employment share during the thirty years for which postindependence data are available. It appears that the complete 1981 data will reveal shifts in the Indian sectoral labor force comparable with those in other low-income countries.

We conclude this section with data on capital formation. Table 1-6 shows net domestic capital formation as a percentage of net domestic product for selected years from 1960–61 to 1980–81. These data reflect a clear, upward trend in the share of net national product devoted to capital formation during the 1970s—growth that has been spurred by national government policy. A similar upward trend in capital formation rates can be found for the 1950s; the share of gross domestic capital formation in GDP rose from 10.0 percent in 1950–51 to 16.2 percent in 1960–61.

By no means all investment is in manufacturing: some is in agriculture, some is in services, and some is in other sectors. Even in the late 1970s, capital formation in manufacturing was only about 25 percent of total capital formation; the secondary sector contributed about 38 percent of total gross capital formation. However, the link between the increased investment rate (both for manufacturing and for all capital formation) and the increased GDP share in manufacturing shown in table 1-4 is somewhat tenuous. The increase in manufacturing's GDP share was not pronounced in the early 1960s and late 1970s, whereas the investment rate grew throughout the 1950s and 1970s. It is important to realize, though, that the monsoon and other short-term shocks cause considerable fluctua-

Table 1-5. *Distribution of Main Workers by Sector, 1961, 1971, 1981*
(percent)

Sector	*1961*	*1971*	*1981*
Agriculture	69.5	69.8	66.7
Other	30.5	30.2	33.3
Total	100.0	100.0	100.0

Source: Government of India (1981, paper 3).

tions in many variables and that the presence or absence of trends may be dependent on the sample of observations chosen. The year 1965, for example, was marked both by severe drought and by war. Consequently, the shares of nonagricultural sectors in GDP were unusually high, and agriculture's share was abnormally low. In contrast, 1977 was a relatively good year for agriculture but not for manufacturing.

Institutional Framework

Governments have important effects on resource allocation in all countries, developed or developing. In India, however, the effects of government on resource allocation are stronger, more direct, and more systematic than in most other developing countries. In this section we present a brief introduction to India's economic policymaking, concentrating on the mechanisms that are germane to the discussions in the chapters that follow. This section relies heavily on Bhagwati and Desai (1970), Frankel (1978), and Uppal (1975), which can be consulted for more detail.

India has been the world's largest functioning democracy since the country won its independence in 1947. The government is federal in character. The national government sets international and basic national policies, but state governments have important constitutional powers and economic functions. India's governmental structure is not unlike that of the United States.

India is ideologically socialist, but is in practice a mixed economy. Government owns most heavy industry, much of the transportation industry, many banks and insurance companies, and most utilities. In addition, about one-third of capital formation is by government. Finally, government regulates the private sector through a variety of mechanisms.

The primary instrument of government planning and policymaking is the Indian Planning Commission. Founded in 1950, it has promulgated five-year plans periodically since 1951. It is the foremost economic advisory body in the Indian government. Responsible for both domestic and international matters, it has been the focus of thought and controversy about both the speed and the direction of India's industrialization. The Planning Commission was responsible for decisions made during the 1950s to concentrate large investments in government-owned heavy industry and for decisions made in the early 1970s to increase government savings and to allocate most of the resulting capital formation to industry.

Five-year plans include targets for both capital formation and total spending by sector. The plans are implemented by government spending and by controls on private spending. Five-year plans set targets for produc-

Table 1-6. *Capital Formation as a Percentage of National Production*

Capital formation	*1950–51*	*1955–56*	*1960–61*	*1965–66*	*1968–69*	*1970–71*	*1973–74*	*1975–76*	*1976–77*	*1977–78*	*1978–79*	*198*
Net domestic capital formation as a percentage of net domestic product	n.a.	n.a.	12.7	13.8	n.a.	13.0	n.a.	15.3	n.a.	n.a.	n.a.	1
Gross domestic capital formation as a percentage of gross domestic product at market prices	10.0	14.3	16.9	18.2	15.8	17.8	20.0	n.a.	21.3	20.0	23.7	n
Gross domestic capital formation in manufacturing as a percentage of gross domestic product	1.2	2.7	4.7	5.1	3.6	4.9	5.5	n.a.	4.5	4.9	6.3	n

n.a. Not available.

Sources: Top row, Government of India (1982, p. 73); middle and bottom rows, World Bank (1980, pp. 221, 228).

tion sectors and for community development; targeted expenditures are divided between urban and rural activities and among sectors in great detail. Investments in government-owned heavy industries are specified precisely.

The private sector is regulated by price, wage, and rent controls and by the government's physical allocations of many important raw materials and produced commodities. Equally important is industrial licensing. Since 1951, private firms have been required to obtain government licenses to build, expand, or operate most manufacturing facilities. Licenses specify permitted activities in detail. Special licenses are required for capital investments, imports, and foreign collaboration. Licensing has been employed to implement five-year plans, to encourage small-scale and cottage industries, and to influence the choice of locations for licensed activities. The effects of licensing and other programs on India's city size distributions are discussed in chapter 7.

Although the government sector is large in India, most economic activity (production and employment) is in the private sector in agriculture, manufacturing, construction, and most services. Parts of the private sector are highly competitive. Yet in some sectors government policy has been to control entry by requiring licenses, by monitoring investment funds, by imposing import controls, and by placing limitations on foreign firms. Resource allocation is then moved in the direction of government targets with the help of price and wage controls, physical allocations of inputs and outputs, licensing requirements, and informal controls.

References

Bhagwati, Jagdish, and Padma Desai. 1970. *India: Planning for Industrialization.* London: Oxford University Press.

Chenery, Hollis, and Moises Syrquin. 1975. *Patterns of Development: 1950–1970.* London: Oxford University Press.

Frankel, Francine. 1978. *India's Political Economy: 1947–1977.* Princeton, N.J.: Princeton University Press.

Government of India. 1951. *Census of India, 1951.* New Delhi: Census Commissioner of India.

———. 1961. *Census of India, 1961.* New Delhi: Census Commissioner of India.

———. 1971. *Census of India, 1971.* New Delhi: Census Commissioner of India.

———. 1978. *Statistical Abstract: India, 1978.* New Delhi: Census Commissioner of India.

———. 1981. *Census of India, 1981.* Series 1. New Delhi: Census Commissioner of India.

———. 1982. *Economic Survey, 1981–82.* New Delhi: Census Commissioner of India.

Lal, Deepak. 1981. "Cultural Stability and Economic Stagnation: India, 1500 B.C.–1980 A.D." Washington, D.C.: World Bank, Development Economics Department. Processed. (To be published in 1986 by Oxford University Press as *The Hindu Equilibrium.*)

Mitra, Asok. 1978. *India's Population.* Vols. 1 and 2. Delhi: Abhinaw Publications.

Uppal, J. S., ed. 1975. *India's Economic Problems.* New Delhi: Tata McGraw-Hill Publishing Co.

World Bank. 1980. *World Tables: The Second Edition* [*1980*]. Baltimore, Md.: Johns Hopkins University Press.

———. 1982. *World Development Report 1982.* New York: Oxford University Press.

Notes

1. Here and subsequently, unless otherwise indicated, Indian monetary figures stated in U.S. dollars have been converted at the official exchange rate for the year in question.

2. Chenery and Syrquin (1975) define the three sectors in much the same way as the World Bank (1982). Their analysis is for 1965.

3. For 1971, main-worker figures are smaller than those for total workers, but for 1961 the two are the same. In 1971, 3.2 percent of total workers were secondary workers, whereas in 1981, 11.0 percent of workers were marginal workers. Criteria for these designations were somewhat different in the two censuses.

2

The Relationship between Urbanization and Economic Development

THIS CHAPTER PRESENTS an analysis of the relationship between urbanization and economic development, using data from most of the world's countries. Its purpose is to place Indian urbanization in a worldwide context. Chapter 3 is concerned with Indian urbanization in the context of India's historical development.

Urbanization has accompanied economic development in both developing and developed countries. Since World War II, according to the wealth of data available for the postwar period, no country has experienced sustained economic growth without also experiencing urbanization. Data from before World War II, most of which pertain to today's developed countries, suggest that the relationship has existed for several centuries.

That some amount of urbanization is a desirable and almost inevitable result of economic growth is hardly open to dispute. Given this premise, many interesting questions can be asked. First, why does urbanization accompany economic growth? Second, at different stages of the growth process, how much consistency of experience is there among countries? Do all countries attain about the same level of urbanization at a given stage of economic growth? Third, what is the form of the relationship, and how is economic growth to be measured in relating it to urbanization? Fourth, can such analysis be used to forecast urban growth in developing countries? Fifth, can any meaning be given to the notions of overurbanization and underurbanization? If so, can they be measured, and what, if anything, can governments do about the phenomena? Sixth, what can be said about the causes of urbanization and about the social implications of urbanization? How do the sizes and locations of urban areas, for example, affect the living standards and quality of life of the people?

Some of these important and complex questions are considered in this chapter. Some are addressed in other chapters, at least within the context of India's experience. Some are merely given definition here as subjects for future research.

Urbanization occurs because of massive shifts of labor and capital from predominantly rural to predominantly urban activities in the course of economic development.[1] In many respects the most satisfactory way to study such shifts is with a large model that analyzes and predicts changes in inputs and outputs as an economy develops. Such models are usually referred to as computable general equilibrium models. An urban modeling project for India is being undertaken by the authors and Jeffrey Williamson, and the findings will be reported in a separate document. In the present book our goal is more modest. We analyze simple statistical relationships that economic theory and common sense suggest may shed light on urbanization. The relationships studied have the advantage of being easy to understand and estimate. But they are approximate and intuitive, and they ignore important detail.

The outline of the chapter is as follows. To set the stage for subsequent sections, we first provide some factual background about the relationship between urbanization and economic development. Then we discuss the reasons urbanization regularly accompanies economic development. We go on to present an elaborate statistical analysis of worldwide relationships between urbanization and development, and in the last section we employ the statistical relationships to explain and predict Indian urbanization.

Factual Background

World Bank data make clear the strength and pervasiveness of the relationship between urbanization and economic development. Tables 2-1 and 2-2 present some cross-sectional data. Table 2-1 shows the weighted average GNP per capita for low-income, middle-income, and industrial market economies in 1980 and the weighted average percentage of the population that is urban (hereafter referred to as percent urban). These averages show that high-income countries are much more urbanized than low-income countries. The figures also suggest something about the form of the relationship between economic development and urbanization. The fivefold rise in income between the low- and middle-income countries is accompanied by a threefold increase in percent urban. The additional sevenfold rise in income between the middle-income and the industrial market economies

Table 2-1. *Average Percent Urban and* GNP *Per Capita, 1980*

Country group	GNP *per capita (U.S. dollars)*	*Percent urban*	*Number of countries*
Low-income	260	17	32
Middle-income	1,400	45	63
Industrial market	10,320	78	19

Note: Here and in tables 2-2 through 2-4, domestic GNP data have been converted to U.S. dollars at official exchange rates.
Source: World Bank (1982, tables 1 and 20).

is accompanied by only a 70 percent rise in percent urban. These data suggest that percent urban increases much more rapidly with economic growth at low income levels than at high ones. Given that percent urban is bounded from above by 100, whereas GNP per capita is not bounded, the relationship can hardly be linear. Neither general considerations nor the averages in table 2-1 can reveal much about the strength of the relationship.

Table 2-2 shows GNP per capita and percent urban for India and its neighbors in 1980. These countries have many elements of geography, climate, history, and culture in common, but of course display much less variability in income levels than would a worldwide sample. The table shows a strong, but by no means perfect, correlation between GNP per capita and percent urban. Definitions of urban areas vary somewhat among countries; more important, countries at given development levels have

Table 2-2. *Percent Urban and* GNP *Per Capita in India and Neighboring Countries, 1980*

Country	GNP *per capita (U.S. dollars)*	*Percent urban*
Bangladesh	130	11
Bhutan	80	4
Burma	170	27
China	290	13
India	240	22
Nepal	140	5
Pakistan	300	28
Sri Lanka	270	27

Source: Same as table 2-1.

Table 2-3. *Urbanization and Economic Growth, 1960–80*

Country group	*Percent urban* 1960	*Percent urban* 1980	*Average annual growth rate in GNP per capita, 1960–80*
Low-income	13	17	1.2
Middle-income	33	45	3.8
Industrial market	68	78	3.6

Source: Same as table 2-1.

somewhat different industrial structures, and it is reasonable to assume that this would lead to differences in percent urban.

A time series analysis tells much the same story: rapidly growing countries urbanize more quickly than slowly growing countries, and urbanization responds more rapidly to economic growth in low- than in high-income countries. Table 2-3 shows the weighted average percent urban for 1960 and 1980 and the growth rate of real GNP per capita during the two decades for countries classified as low-income, middle-income, and industrial market economies in 1980. The low-income countries increased their percent urban only slightly, because their economies grew slowly. The middle-income countries urbanized rapidly, and their GNP grew rapidly. The industrial market economies grew at almost the same rate as the middle-income countries, but they urbanized less rapidly because they were already at high levels of urbanization in 1960.

Table 2-4. *Urbanization and Economic Growth in India and Neighboring Countries, 1960–80*

Country	*Percent urban* 1960	*Percent urban* 1980	*Average annual growth rate in GNP per capita, 1960–80*
Bangladesh	5	11	0.0
Bhutan	3	4	−0.1
Burma	19	27	1.2
China	n.a.	13	n.a.
India	18	22	1.4
Nepal	3	5	0.2
Pakistan	22	28	2.8
Sri Lanka	18	27	2.4

n.a. Not available.
Source: Same as table 2-1.

Table 2-4 makes the same comparison for India and its neighbors. Bhutan and Nepal experienced little economic growth and little urbanization, whereas Bangladesh did undergo some urbanization despite the absence of economic growth. Burma, India, Pakistan, and Sri Lanka showed common responses of urbanization to economic growth.

The data presented in this section suggest the strength and pervasiveness of the relationship between urbanization and economic growth. They also show that the relationship is complex and varies somewhat among countries. But because the data are based on broad averages and represent only a narrow range of countries, they can be no more than suggestive. Much more careful statistical analysis is required for more conclusive results.

Why Urbanization Accompanies Economic Development

Urbanization is a natural and inevitable consequence of economic development. At the level of national urbanization, the reasons for urbanization are now well understood.[2] Urbanization accompanies economic development because economic development entails a massive shift of labor and other inputs from sectors that are predominantly rural to sectors that are predominantly urban. Although the migration of people from rural to urban residences and jobs is an important accompaniment of urbanization, the sectoral, not the locational, change is the crucial phenomenon. In extreme cases urbanization entails little or no physical movement.

Explanation is needed of two phenomena: why do the shifts to the sectors that are characteristic of urbanization take place and why is it the predominantly urban sectors, and not the rural or agricultural sector, whose relative growth accompanies development? Each phenomenon will be discussed in turn.

Food is the prime requirement for life, and in the poorest countries most production effort is devoted to agriculture. As economies develop, inputs and outputs shift from agriculture to industry, especially manufacturing, and to services.[3] In 1980, low-income countries had a weighted average of 71 percent of their labor forces in agriculture; in middle-income countries the percentage was 44, and in industrial market economies the percentage was 6. The labor force percentages in industry and services were: in the low-income countries, 15 and 15, respectively; in the middle-income countries, 22 and 34; in the industrial market economies, 38 and 56. Over the entire range of development, the percentage of the labor force in agriculture falls from about 70 in the poorest countries (those with a per capita GNP of between \$80 and \$200) to 2–8 in the richest countries (those with a per capita GNP of between \$11,000 and \$16,000).[4]

Why does the labor force share fall in agriculture and rise in industry and services during economic development? Two phenomena are involved: demand shifts and supply shifts resulting from capital accumulation and technological change.

On the demand side, it has been known for more than a century that a large share of income is spent on food at low income levels, but that the share falls as income increases. Income shares spent on industrial products and services rise with income. These phenomena are known loosely as Engel's law, and they reflect low-income elasticities of food demand and high-income elasticities of industrial goods and service demands. The effect of these demand shifts is to raise returns to labor and other inputs in industry and services relative to those in agriculture, and so to induce workers and other inputs to move from agriculture to the other sectors.

On the supply side, the costs and prices of industrial products may fall relative to those of agricultural products. Technical progress may be faster in industry than in agriculture, in part because industrial technology is often more directly transferable—from developed to developing countries—than agricultural technology. In addition, capital accumulation and scale economies may reduce industrial prices relative to agricultural prices. Finally, a more highly educated labor force may be of more benefit in industry and services than in agriculture.

The supply-side effects discussed in the previous paragraph depend on international considerations. Most industrial products are tradable, whereas most services are not. If the economy is open, cost reductions cause large increases in output and employment and small decreases in the prices of tradable commodities. Thus, cost reductions in industry are likely to cause large inflows of labor.

In conclusion, economic development causes a shift of labor from agriculture to industry because demand rises and costs fall in industry relative to agriculture. The shift to services is harder to be specific about. It certainly results from demand increases. It is doubtful that service costs fall relative to those in agriculture or industry.

The basic definition of an urban area is a set of contiguous places where population densities are much greater than elsewhere. Since workers tend to live near their place of work, most workers who reside in urban areas also work there and most workers who are employed in urban areas also live there. What needs to be explained, then, is why industrial and service production and consumption take place in dense and, in some cases large, urban areas.

A necessary condition for the high employment and population densities characteristic of urban areas is that production be possible without large amounts of land relative to other inputs. For the most part, commodities and services produced in urban areas simply require less land per unit of output than does agriculture—that is, urban production is less land-intensive. In addition, high urban land values induce some substitution of other inputs for land, thus permitting high urban densities. For example, services are frequently produced in tall office buildings in large cities, but in small buildings elsewhere.

Without the ability to economize on land inputs, typical urban densities of production and housing would be impossible.

Two other conditions are sufficient for the emergence of urban areas. The first is the existence of scale economies. Scale economies exist if the long-run unit cost declines as output increases, at fixed input prices. Scale economies exist in all economic activities, at least at small volumes of output. In some manufacturing activities, scale economies are exhausted only in a plant employing hundreds of workers (and, of course, appropriate amounts of other inputs). In service activities, scale economies may be exhausted by a few or a few dozen employees (and similarly modest amounts of other inputs). The size necessary to exhaust scale economies in service production depends on how specialized the services are. A law firm providing routine legal services may exhaust scale economies with a few employees, whereas a firm providing a wide range of specialized corporate legal services may require hundreds of employees to exhaust scale economies. Since demands per resident or per firm are small for specialized services, and since services must for the most part be consumed where they are produced, only the largest urban areas can provide a market large enough to support highly specialized service activities.

The second sufficient condition is the need for, and the cost of, transportation. Delivery of inputs and outputs requires movement of commodities and people. Material inputs must be shipped from the firms that supply them to the firms that use them, and workers must commute to and from work. Final products must be shipped to the customers. With services, the producer sometimes comes to the customer and the customer sometimes comes to the producer. A few services can be provided without face-to-face contact, but whenever face-to-face contact is required, people must be transported. If transportation were free, scale economies would not necessarily entail substantial urban concentrations: scale economies could be realized by producing a large volume of outputs in one place and having the inputs brought in from dispersed suppliers and the outputs shipped out to dispersed customers. But transportation, especially of people, is expensive. High transportation costs motivate producers of commodities and services and their business and household customers and input suppliers to locate in proximity to one another. Proximity permits production at economical scales without squandering the benefits on expensive transportation.

The above two conditions are sufficient to produce urban concentrations of considerable size. But most urban areas, and certainly the largest urban areas, owe their size in part to a regional advantage of some sort, most frequently the availability of certain natural resources. Farmland that is ideal for growing cotton may justify the location of textile and garment industries in a nearby city or port from which the processed materials can be shipped cheaply to other regions or countries.

Urban specialists identify other economies of urban areas, sometimes referred to as agglomeration economies. Most turn out, on examination, to be scale economies in one sector or another, commodity or service production, government or private. One kind of agglomeration economy is not an example of scale economies and is intriguing. Suppose two industries employ the same kinds of workers. Suppose their labor demands vary from time to time, from seasonal, cyclical, or random causes, in ways that are not perfectly correlated. An urban area that contains some of both industries can provide a better approximation to full employment than can two urban areas each containing one of the two industries. Thus a large urban area can generate higher real incomes than a small one. No one has studied the importance of this economy.

Urban areas vary enormously in size, organization, and functions. A large metropolis differs in many important ways from a crossroads town. Yet all result from the same basic set of economic circumstances. And the same economic considerations cause their sizes and numbers to grow as a country develops.

Statistical Analysis

In the previous section we concluded that urbanization accompanies economic development because sectors with incentives to locate at densities and proximities characteristic of urban areas grow more rapidly than other sectors. At the most abstract level, it is difficult to imagine that this explanation could be false. Thus, the main purpose of this section is not to test the explanation. Instead, our purpose is to choose particular indices of development and particular functional forms to estimate the relationship between urbanization and economic development as carefully as possible. Such statistical analysis will yield quantitative estimates of the importance of several variables, estimates of how closely countries follow the typical pattern, and forecasts of likely future urbanization. In particular, it will enable us to assess how closely India's urbanization experience parallels worldwide trends and to forecast the course of India's future urbanization.

In this section, percent urban is regressed on measures of economic development and related variables. Sample observations are for a large set of countries for 1980.

The most obvious measure of economic development to include in the regressions is income per capita. Three measures of income are available for a large set of countries: GNP, GDP, and Kravis-adjusted (for differences in purchasing power) GNP. All three measures were tried, and the results will be reported below. The second easily available and widely used measure of economic development is industrial structure—specifically, labor force shares in industry and agriculture. Economic base theory shows that since industrial output is, for the most part, sold outside the urban area in which it is produced, it influences but is not influenced by the population of the urban area. That makes industrial output a desirable independent variable in our regressions. There is also an argument for using the share of the labor force in agriculture. Both industrial employment and service employment influence the size of the urban area in which they are located. Especially in developing countries, an unknown, but probably significant, part of the industrial output (for example, construction) is consumed in the urban area, and some, again unknown, part of service production is consumed outside the urban area (for example, agriculture education, extension, and research services). Thus, there is no clear-cut distinction between industry and services as to which is exogenous to the size of the urban area.

Inclusion of the agricultural employment share is equivalent to inclusion of the sum of the industrial and service shares, since the three categories are exhaustive. It is by no means true that inclusion of agriculture's share makes the relationship with the urban shares of the population definitional. In both developing and developed countries, large amounts of nonagricultural employment are located in rural areas. Both industrial and agricultural employment shares are included in the regressions reported below. Both are correlated with income per capita (with opposite signs). Thus, separate effects may be hard to measure.

Three other variables are candidates for inclusion in the regressions: time, population density, and lagged urbanization. Many people believe that developing countries are urbanizing faster than is justified by the pace of their economic development and are therefore becoming overurbanized. If so, time should be correlated with urbanization, once the effects of development measures have been accounted for. Thus, inclusion of a time variable is a test of the overurbanization hypothesis.[5] The test requires data from two points in time.

Population density is another variable that can be included in the regressions. Some people claim that high rural population densities lower agricultural incomes and push people to urban areas. Some cite this phenomenon as another cause of overurbanization, although is it not clear why the effect should be thought of as excessive urbanization; if the relationship exists, population may be excessive, not urbanization.

An argument for including lagged urbanization in the regressions can also be made. If the equilibrium percent urban U^* depends on a set of variables according to a function $f(\ )$, then

$$U^* = f(\ ). \tag{2.1}$$

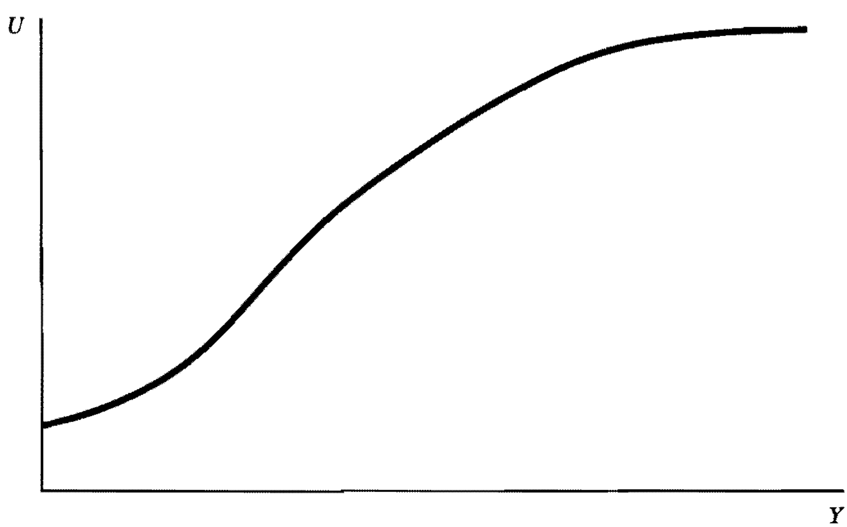
Adjustment of the population to changes in variables affecting equilibrium urbanization takes time, since major migration decisions may be involved. The most common assumption about such disequilibrium adjustments is that they follow a distributed lag

$$U = U_{-1} + \lambda\,(U^* - U_{-1}) \tag{2.2}$$

where U is percent urban and U_{-1} is its lagged value. Equation (2.2) says that U equals its lagged value plus a fraction λ of the difference between the equilibrium and lagged values. If $\lambda = 0$, $U = U_{-1}$ regardless of U^*, and if $\lambda = 1$, $U = U^*$ regardless of U_{-1}. Thus, one would expect λ to be in the interval $0 < \lambda < 1$. Substitution of equation (2.1) into (2.2) gives

$$U = \lambda f(\ ) + (1 - \lambda)U_{-1}. \tag{2.3}$$

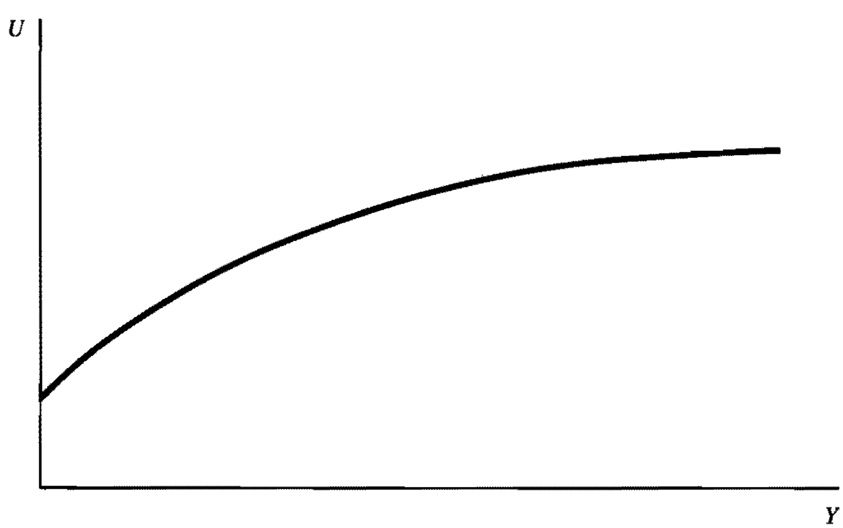
Figure 2-1. *Logistic Function*



Equation (2.3) contains only observable variables and justifies introduction of U_{-1} in the regressions.

The final issue is the choice of functional form for $f(\ )$ in equation (2.1). Many suggestions have been made in the literature that $f(\ )$ should be a logistic function, and the only other study we can find of cross-sectional determinants of urbanization uses the logistic (World Bank 1972).[6] The key characteristic of the logistic function shown in figure 2-1 is that urbanization responds slowly to economic development (indexed by income per capita Y in the figure) at low development levels, responds rapidly at higher levels, and again slowly at very high development levels. The preference for the logistic function seems to be based on confusion between urbanization as a function of time and urbanization as a function of economic development. It is reasonable to expect urbanization to be a logistic function of time, and many demographic models imply that it, or a closely related function, is best (Ledent 1980). Both economic development and urbanization are likely to start slowly as functions of time—as a country struggles to escape from one or more low-level traps—then accelerate, and finally decelerate after decades of development. But no extant model or known mechanism suggests that urbanization will respond more slowly to appropriate measures of economic development at low than at middle development levels. A more plausible functional form is that in figure 2-2, which shows continuously decelerating response of urban-

Figure 2-2. *Convex Urbanization Function*



ization to economic development. Figure 2-2 suggests that logarithmic or polynomial functions of development measures would be most appropriate. The choice of functional form is of practical importance. The graph in figure 2-2 indicates faster urbanization relative to income growth at low levels of income than does the graph in figure 2-1.

The logistic function can be written

$$U_{it}^{-1} = (a_0 + a_1 e^{g(\,)}) \tag{2.4}$$

where U_{it} is percent urban in country i and year t, $g(\,)$ is a function of development variables, a_0 and a_1 are constants to be estimated from data, and e is the base of the natural logarithm. Equation (2.4) can be estimated by taking natural logs of both sides and rearranging terms to get

$$\ln (U_{it}^{-1} - a_0) = \ln a_1 + g(\,). \tag{2.5}$$

If $g(\,)$ in equation (2.5) is linear in its parameters, then (2.5) is linear in all parameters except a_0. Equation (2.5) can be estimated by running linear regressions for a large number of a_0 values and choosing the a_0 value that maximizes R^2.

After some experimentation, it was discovered that the best results were obtained for the following functional forms and variables:

$$\begin{aligned} U_{it} = {} & 73.0286 - \underset{(15.19)}{0.7157A_{it}} + \underset{(1.67)}{13.3706Y_{it}} - \underset{(1.58)}{6.6886Y_{it}^2} \\ & + \underset{(1.30)}{1.0338t}. \qquad R^2 = 0.799 \end{aligned} \tag{2.6}$$

$$\begin{aligned} \ln (U_{it}^{-1} - 1.05) = {} & -1.6409 + \underset{(14.99)}{0.0448A_{it}} - \underset{(0.25)}{0.1282Y_{it}} \\ & + \underset{(0.60)}{0.1603Y_{it}^2} - \underset{(1.97)}{0.0996t}. \qquad R^2 = 0.769 \end{aligned} \tag{2.7}$$

$$\begin{aligned} U_{it} = {} & 13.98 - \underset{(0.70)}{0.0358A_{it}} - \underset{(2.73)}{18.2518D_{it}} + \underset{(2.24)}{17.0928Y_{it}} \\ & - \underset{(2.13)}{10.7269Y_{it}^2} + \underset{(15.87)}{0.9338U_{i,t-2}}. \qquad R^2 = 0.930 \end{aligned} \tag{2.8}$$

The notation is as follows:

U_{it} = percent urban in country i in year t

Y_{it} = real GNP per 10,000 population in country i in year t, measured in U.S. dollars

t = year ($t = 2$ is 1980, $t = 0$ is 1960)

D_{it} = population density in country i and year t, people per 1,000 square kilometers

A_{it} = percentage of the labor force in agriculture in country i in year t.

In all three regressions the sample is 105 countries in 1980 and 104 in 1960.[7] Absolute values of t-statistics appear in parentheses below the coefficient estimates.

Regression (2.6) is the type of regression represented in figure 2-2. Regression (2.7) is an estimate of the logistic (2.5), represented in figure 2-1. Regression (2.8) is an estimate of equation (2.3), using an $f(\ )$ function that includes population density. In all three regressions, inclusion of A, the agricultural share of the labor force, proved to be better than inclusion of the labor force share in industry. (See the appendix at the end of this chapter for alternative regressions.)

In all three regressions the coefficients have plausible signs and R^2's are large for cross-country studies. The R^2 for (2.7) is not directly comparable with those for (2.6) and (2.8) because (2.7) has a different dependent variable. High agricultural labor force shares always lead to low values of U. In each regression the derivative of U with respect to Y is positive and becomes small at high income levels. Regression (2.7) has the characteristic variation of U with Y and A that is suggested by figure 2-1. In no case did GDP or Kravis-adjusted GNP perform as well as GNP as a measure of Y (see the appendix regressions).

In regressions (2.6) and (2.7), urbanization increases as time passes for given values of other independent variables. But the significance levels of t coefficients are not high, and the coefficients are small. In (2.6) the urban share increases only one percentage point per decade as a result of the passage of time. In (2.7) the effect of time on urbanization disappears as high levels of urbanization are reached. The conclusion is that urbanization occurs independently of economic development to only a negligible extent.

It did not prove feasible to include both t and $U_{i,t-2}$ in the same regression. In regression (2.8) the coefficient of $U_{i,t-2}$ is large and highly significant. Inclusion of the lagged value of U in (2.8) captures not only the lagged adjustment of U to changes in development indices but also any national differences in urbanization that are not explained by the stage of development. Although inclusion of the lagged value of U raises R^2 by about 0.16, it is unclear which of the effects it captures is more important.

Density appears in regression (2.8) with a negative and significant coefficient. Our regression analysis provides no support for the notion that high population density leads to a large percent urban.

A and Y are correlated (about −0.35 in our sample), so only one tends to be significant in the same regression. Addition of A^2 added nothing to the regressions, whereas deletion of Y^2 made most regressions somewhat less satisfactory.

It is difficult to choose among the three regressions in terms of sign, magnitude, and significance of coefficients. The only clear difference is that (2.8) explains more of the sample variation of U than does (2.6). Choice among the three regressions should depend on their detailed characteristics in representing the urbanization process.

Regressions (2.6) and (2.8) have similar qualitative characteristics. In both, the increase in U per unit decrease in A is constant, regardless of development level, whereas U increases rapidly for given increases in Y at small values of Y and slowly at large values of Y. At very high Y values, U decreases as Y increases. In (2.6), U increases slowly, then rapidly, and then slowly again as A falls or Y rises. The R^2's tell us that (2.8) explains a larger fraction of the variance in U than (2.6), but for (2.7) we know only that it explains 76.9 percent of the variance in $\ln (U^{-1} - 1.05)$.

It is inconceivable that all three regressions provide equally good representations of the sample data. Table 2-5 shows actual values of U in 1980 for India and for the average of low-income countries, middle-income countries, and industrial market economies, together with values of U calculated from the three regressions.

All three regressions overpredict U for the average of high-income coun-

Table 2-5. *Calculated and Actual Values of U, 1980*

		Calculated U		
Country	*Actual U*	*From regression (2.6)*	*From regression (2.7)*	*From regression (2.8)*
India ($A = 71$, $Y = 190$, $D = 200.5$, $U = 18$, $t = 2$)	22	24.5	20.6	24.9
Average low-income ($A = 71$, $Y = 230$, $D = 66.9$, $U = 15$, $t = 2$)	17	24.6	20.7	24.6
Average middle-income ($A = 43$, $Y = 1{,}420$, $D = 25.4$, $U = 37$, $t = 2$)	50	46.1	47.1	48.7
Average industrial market ($A = 6$, $Y = 9{,}440$, $D = 22.1$, $U = 68$, $t = 2$)	77	77.5	79.2	83.4

Note: The units given here for A, U, and t are used in the regressions, but the regressions use $Y/10{,}000$ and $D/1{,}000$.

tries. The logistic (2.7) underpredicts, and the other two regressions over-predict, India's 1980 U value. Interestingly enough, despite the fact that (2.7) minimizes the sum of squared errors around a function of U, whereas (2.6) and (2.8) minimize the sum of squared errors around U, (2.7) has a much smaller sum of squared errors for the four observations in table 2-5 than do (2.6) and (2.8).[8]

Since our main field of inquiry is the urbanization of developing countries, the bottom line in table 2-5 is of less interest than the top three lines. Regression (2.8) gives the closest prediction of U for middle-income countries, whereas (2.7) gives the closest prediction of U for low-income countries and for India. We have already argued that the logistic regression (2.7) is less plausible than (2.6) and (2.8) as an explanation of urbanization at low levels of development. There is little in these regressions to suggest that low-income countries are urbanizing according to a pattern different from that of other countries.

Forecasts of India's Urbanization

In this section, regressions (2.6), (2.7), and (2.8) are used to forecast urbanization in India for 1990, just before the next Indian census will be taken. Logically, forecasts of U from the regressions must be based on forecasts of A, Y, and D. One might claim that it is easier to forecast U directly than to forecast A, Y, and D, but use of the regressions imposes consistency on the forecasts. A large economic model would, of course, provide forecasts of A and Y, and possibly of D.

Forecasting U for India is complicated by the fact that the Bank's estimate of U for 1980, 22, is almost certainly low. The 1981 Indian census shows $U = 23.7$. Urbanization increased about 0.3 percentage points a year between 1961 and 1981 (from 18.3 in 1961 to 23.7 in 1981), so U must have been at least 23 in 1980. Yet the Bank's estimate of 22 was used to estimate all the regressions. The underestimation of U for 1980 should be less serious for (2.8) than for (2.7) and (2.6) because (2.8) contains the value of U lagged one decade and the correct value can be inserted.

Our forecasts of A, Y, and D are based on historical trends, but they are no more than illustrative. Anyone can easily make alternative forecasts from the regressions by using other values of independent variables. In our forecasts we used $A = 66$, $Y = 215$, and $D = 244$. These projections were arrived at as follows: A was about 69.5 in India in 1961 and 69.8 in 1971, but it fell to 66.7 in 1981. Continued modest economic growth will almost certainly cause further decline, but 66 in 1990 is only a guess. Indian real GNP per capita grew about 1.25 percent per year during the

1960s and just less than 1.0 percent a year during the 1970s. Our guess is 1.24 percent a year during the 1980s. It is an optimistic guess and leads to a 1990 value of 215. The population of India has grown about 2.19 percent a year in the two decades since 1961. We assume that rate of growth will continue until 1990, which implies a 1990 population of 801.5 million (the population in 1981 was 658.1 million) and a density of 244 people per square kilometer (Government of India 1981).

In summary, the right-hand variables assumed in the forecasts for 1990 are $A = 66$, $Y = 215$, $D = 244$, $t = 3$, and $U_2 = 23.3$. The forecasts of U are on the top line of table 2-6. An alternative way to make the forecasts is to calculate 1980 and 1990 values of U from the regressions, take the difference between the two, and add it to the actual U for 1980. Using the notation $\hat{U}_3$ and $\hat{U}_2$ respectively for the 1990 and 1980 values of U calculated from the regressions, we have a forecast $\hat{\hat{U}}$ where

$$\hat{\hat{U}}_3 = (\hat{U}_3 - \hat{U}_2) + U_2.$$

These forecasts are in the bottom row of table 2-6. The forecasts are lower for (2.6) and (2.8) than those in the top row because, as shown in table 2-5, (2.6) and (2.8) overforecast $U_2 = 23.3$.

If the forecast from regression (2.7) is ignored, on the assumption that the logistic underforecasts the response of urbanization to economic development at low development levels, the forecasts in table 2-6 cluster between about 27 and about 29 percent urban. For illustrative purposes, take 28 as a reasonable forecast of India's 1990 percent urban. In calculating D for 1990, a total population forecast of 801.5 million people was used. Applying the 28 percent urban forecast to that total population forecast yields a 1990 Indian urban population forecast of 224.42 million people. The 1981 urban population was 156.19 million. The 1990 forecast implies an annual urban population growth rate of 4 percent, which is slightly higher than the urban population growth rate during the 1970s.

Urban population growth rates of 4 percent, or roughly double the national population growth rate, are not uncommon in developing coun-

Table 2-6. *Forecasts of Percent Urban in India, 1990*

	Regression		
Forecast	*(2.6)*	*(2.7)*	*(2.8)*
Calculated from regression	28.1	24.5	28.9
Calculated from regression increments	26.9	27.2	27.3

tries, and they are consistent with India's recent experience. It is the large numbers of people predicated by such a rate of growth that makes forecasts of Indian urban growth frightening. If India's 1990 urban population is 224.42 million, this implies an increase in the urban population of 68.23 million during the 1980s—an increase of almost 7 million people a year.

Appendix: Alternative Regression Estimates

Several statements appear in the text about the characteristics of alternatives to regressions (2.6), (2.7), and (2.8). This appendix presents a few alternatives to illustrate the characteristics.

First, consider the effect of replacing A with I, the percentage of the labor force in industry. With this change, (2.8) becomes

(2.9)
$$\begin{aligned} U_{it} &= \underset{}{10.83} + \underset{(0.33)}{0.0319 I_{it}} - 18.4264 D_{it} + \underset{(2.27)}{17.9817 Y_{it}} - \underset{(2.11)}{11.1152 Y_{it}^2} \\ &\quad + \underset{(19.93)}{0.9534 U_{i,t-1}}. \qquad R^2 = 0.93 \end{aligned}$$

This regression is similar to (2.8), but I in (2.9) is less significant than A in (2.8).

Second, consider the effect of using Kravis-adjusted GNP per capita instead of GNP per capita. With this change, (2.8) becomes

(2.10)
$$\begin{aligned} U_{it} &= 6.39 + \underset{(0.48)}{0.0331 A_{it}} - \underset{(2.19)}{19.5398 D_{it}} + \underset{(3.57)}{71.7795 Y_{itk}} - \underset{(3.23)}{65.1411 Y_{itk}^2} \\ &\quad + \underset{(11.47)}{0.8409 U_{i,t-1}}. \qquad R^2 = 0.93 \end{aligned}$$

Substitution of the Kravis-adjusted income measure makes the coefficients of Y and its square a little more significant than in (2.8), but A now has the wrong sign and is even less significant.

Third, consider dropping $U_{i,t-2}$ from (2.8). The resulting regression is

(2.11)
$$\begin{aligned} U_{it} &= 70.86 - \underset{(9.05)}{0.6378 A_{it}} - \underset{(2.07)}{25.0621 D_{it}} + \underset{(2.11)}{29.1122 Y_{it}} \\ &\quad - \underset{(1.67)}{15.2216 Y_{it}^2}. \qquad R^2 = 0.77 \end{aligned}$$

Regression (2.11) is the same as (2.6) except that (2.11) includes D and (2.6) includes t. Regression (2.11) has a lower R^2 than (2.8), but about the same R^2 as (2.6). Dropping $U_{i,t-2}$ from (2.8) makes A much more significant, but D, Y, and Y^2 less so.

Fourth, consider dropping Y^2 from (2.7). The result is

$$\text{(2.12)} \quad \ln (U_{it}^{-1} - 1.05) = -1.7230 + \underset{(18.19)}{0.0457A_{it}} + \underset{(0.70)}{0.1470Y_{it}} - \underset{(2.00)}{0.1008t}. \qquad R^2 = 0.769$$

There is no effect on R^2 and little on the coefficients of A and t. But the coefficient of Y now has the wrong sign.

Fifth, consider adding an A^2 term to (2.6). The result is

$$\text{(2.13)} \quad U_{it} = 73.8162 - \underset{(3.98)}{0.7446A_{it}} + \underset{(0.16)}{0.0002A_{it}^2} + \underset{(1.32)}{12.5441Y_{it}} - \underset{(1.45)}{6.4675Y_{it}^2} + \underset{(1.30)}{1.0402t}. \qquad R^2 = 0.799$$

R^2 is not affected and other coefficients are hardly affected, but the coefficient of A^2 is insignificant.

Finally, consider a regression estimated from data for just the developing countries. Regression (2.14) is the same as (2.11)—that is, regression (2.8) without the lagged U term—except that (2.14) is estimated from data for just the eighty-seven low- and middle-income countries.

$$\text{(2.14)} \quad U_{it} = 60.94 - \underset{(5.48)}{0.5523A_{it}} - \underset{(2.23)}{30.9477D_{it}} + \underset{(1.96)}{118.2569Y_{it}} - \underset{(1.36)}{170.9119Y_{it}^2}. \qquad R^2 = 0.72$$

R^2 is somewhat lower in (2.14) than in (2.11), and the coefficients are inevitably somewhat less significant because of the smaller sample size. The coefficients of A and D are similar in magnitude to those in (2.11), but those of Y and Y^2 differ greatly from the corresponding coefficients in (2.11). With the smaller sample variation of Y in (2.14) as compared with that in (2.11), collinearity of Y and Y^2 is more serious in (2.14) than in (2.11).

This sample of alternative regressions suggests that minor differences in specification make little difference to the three regressions in the text. But neither the regressions presented in this appendix nor a much larger sample of alternatives that was estimated but not reported shows any clear advantage over the three regressions reported and analyzed in the text.

References

Gaude, J. 1972. *Emploi agricole et migrations dans une économie dualiste.* Geneva: Librairie Droz.

Government of India. 1981. *Census of India, 1981.* Series 1, paper 2. New Delhi: Census Commissioner of India.

Kelley, Allen, and Jeffrey Williamson. 1984. *What Drives Third World City Growth?* Princeton, N.J.: Princeton University Press.

Kelley, Allen, Jeffrey Williamson, and Russell Cheetham. 1972. *Dualistic Economic Development.* Chicago: University of Chicago Press.

Ledent, Jacques. 1980. *Comparative Dynamics of Three Demographic Models of Urbanization.* Laxenburg, Austria: International Institute for Applied Systems Analysis.

Mera, Koichi. 1975. "A Multiple Layer Theory of National Urban Systems." In Harry Swain and Ross MacKinnon (eds.), *Issues in the Management of Urban Systems.* Laxenburg, Austria: International Institute for Applied Systems Analysis, pp. 134–54.

World Bank. 1972. *World Bank Operations: Sectoral Programs and Policies.* Baltimore, Md.: Johns Hopkins University Press.

———. 1981. *World Development Report 1981.* New York: Oxford University Press.

———. 1982. *World Development Report 1982.* New York: Oxford University Press.

Notes

1. Here and elsewhere, it must be understood that physical movement of people and assets is not necessarily implied. The text refers to shares of labor and capital, whose changes do not correlate with physical movements.

2. Kelley and Williamson (1984) and Kelley, Williamson, and Cheetham (1972) have made major contributions to such understanding in recent years.

3. Industry includes not only manufacturing, but also mining, construction, electricity, water, and gas.

4. The variation in output shares is less than the variation in labor force shares during development. In 1980, agriculture's GDP share averaged 36 percent in the low-income countries, 15 percent in the middle-income countries, and 4 percent in the industrial market economies. Industry's GDP share shows no trend with development level. The average GDP share of services rises from 29 to 45 to 62, going from low-income to middle-income to industrial market economies. These figures and those in the text indicate that labor productivity rises more in agriculture than in most other sectors as development proceeds. The statistics in the text and this footnote are from the World Bank (1982).

5. Inclusion of time is a weak test of the overurbanization hypothesis. If percent urban is positively correlated with time, it may be because important variables have been omitted. If the coefficient of time is negative or insignificantly different from zero, that is at least some evidence against the overurbanization hypothesis.

6. Mera (1975) has undertaken a similar study, but it uses Japanese prefectural data instead of national data. It is comparable to the Indian state study we report in chapter 6. Gaude (1972) makes extensive use of logistic functions in analyzing agricultural employment.

7. Excluded were Singapore and Hong Kong (city states), seven countries for which data were unavailable, and four high-income oil-exporting countries. All the data are from the World Bank (1981). U is from table 20, Y_{i2} is from table 1 and is for 1979, Y_{i0} was calculated from Y_{i2} using the growth rates in the table. A is from table 19, and A_{i2} pertains to 1979. D is calculated from data in table 1. The regressions reported in the appendix use Kravis-adjusted GNP instead of GNP for Y. These income data were obtained from World Bank data files. The units indicated were chosen simply to avoid large numbers of zeros in the coefficient estimates.

8. The sums of squared errors (SSE) are: for (2.6), 85.47; for (2.7), 28.9; for (2.8), 100.02. Of course, the overall sample SSE is greater in (2.6) than in (2.8) because R^2 is smaller in (2.6) than in (2.8).

3

Historical Analysis of Indian Urbanization

THE PURPOSE OF THIS CHAPTER is to present a historical description and analysis of Indian urban development. The preceding chapter showed that a country's urbanization is closely related to its income growth and to the labor force shift from agriculture to industry and services that occurs during economic development. This chapter explores the historical details of the relationship for India. Our concern here is with overall national and state urbanization; national and regional size distributions of urban areas are analyzed in subsequent chapters.

India has a richly documented history. Complete national censuses have been taken every decade since 1901. Private estimates of some data go back even further. This chapter focuses on the eighty years of the twentieth century for which official data, especially reliable censuses, are available. Interest is greatest about—and data are most plentiful and reliable for—the period since 1947, when independence was obtained. At the time of writing, only preliminary results are as yet available for the 1981 census. Because the data from the 1981 census are not directly comparable with those from earlier censuses, the 1981 data are discussed in a separate section.

Urban Growth

In this section twentieth-century trends in India's urban growth are reviewed. We have seen from our discussions in the preceding chapter that growth in the percentage of the population that is urban, or percent urban, always accompanies sustained economic growth. The fact that India's economic growth has been modest but sustained would lead one to expect a similar pattern of urbanization.

The basic measures of total and urban population for the census years since 1901 are presented in table 3-1. Although all nine censuses carefully distinguished between urban and rural areas, the Indian census definition of urban residence has changed somewhat over the years, especially since independence (see the appendix at the end of this chapter). The changes, however, do not seem to have been important enough to affect the discussion in this section.

India's population in 1981 was about three times its 1901 value. The population grew at a compound annual rate of 1.30 percent during the eighty-year period. The decrease in population during the 1911–21 decade resulted from the worldwide influenza epidemic that followed World War I. Although plagues, cholera, and famines killed many Indians throughout the early decades of this century, death rates from those causes were not higher in 1911–21 than they had been the previous decade, when the population increased. Influenza, however, devastated the country (Mitra 1978). After 1921 the annual population growth rate increased gradually in almost every decade, peaking at 2.20 percent in 1971–81. The acceleration of population growth resulted almost entirely from declining death rates. Before independence, such declines were most strongly associated with decreases in deaths from communicable diseases. Since independence,

Table 3-1. *Total and Urban Population of India, 1901–81*

	Total population		*Urban population*		
Year	*Millions of people*	*Average annual growth rate (percent)*	*Millions of people*	*Average annual growth rate (percent)*	*Percent urban*
1901	233.0	—	25.6	—	11.0
1911	246.0	0.59	25.6	0.00	10.4
1921	244.3	−0.07	27.7	0.67	11.3
1931	270.7	1.03	33.0	1.75	12.2
1941	309.0	1.32	43.6	2.79	14.1
1951	349.8	1.24	61.6	3.46	17.6
1961	424.8	1.94	77.6	2.31	18.3
1971	528.9	2.19	107.0	3.21	20.2
1981	658.1	2.20	156.2	3.78	23.7

— Not applicable.
Source: Government of India (1981, paper 2, p. 24).

reductions in the frequency and severity of famines have also been important in lowering the death rate.

The urban population grew at an average annual compound rate of 2.26 percent during the eighty-year period—1.74 times the 1.30 percent growth rate of the total population. Urban population growth also accelerated. The urban growth rate was zero or negligible until 1921. Then it accelerated rapidly; during the 1931–41 and 1941–51 decades the urban population grew more than twice as rapidly as the total population. Since 1951–61, the first decade after independence, the urban growth rate has been less than two times the total population growth rate. The unusually high 3.46 annual urban population growth rate during the 1941–51 decade probably resulted from the disruption that followed the partitioning of India in 1947. But that cannot be the explanation for the high rate of urban growth—2.79 percent in 1931–41.

The percentage of the population that is urban increased gradually and at a remarkably steady pace. Although India's percent urban dropped between 1901 and 1911, it has increased in every decade since then. In 1981 the percentage was more than twice its levels in the early censuses of the century. Percent urban, like the growth rate of the urban population, shows evidence of acceleration before independence. From 1901 to 1931, percent urban increased an average of only 0.04 percentage point a year, whereas from 1931 to 1981 it increased an average of 0.23 percentage point a year. From 1941 to 1981 it grew 0.24 percentage point a year. During the entire postindependence period 1951–81, it increased 0.20 percentage point a year.

It is clear that urbanization has been more rapid during recent decades than during the early decades of the century. Similarly, as was pointed out in chapter 1, real income per capita has grown more rapidly in recent decades than it did earlier in the century. There is thus a rough correlation between real income growth and urbanization, as expected.

Urbanization, however, accelerated in India well before 1941 (probably shortly after 1931), whereas income growth accelerated only later. Heston's estimates show a modest 0.30 percent annual growth rate of per capita real income between 1900 and 1945. But, as we saw in table 1-1 of chapter 1, all of that growth occurred before 1930 (Lal 1981, p. 171). Real income per capita fell between 1930 and 1940 and again between 1940 and 1945. Much of the urbanization that took place between 1941 and 1951 may have occurred after India was partitioned in 1947, but that does not explain the rapid pace of urbanization during the 1930s. Acceleration in urbanization seems to have preceded acceleration in real income growth.

Why urbanization should have accelerated during the 1930s, and perhaps during the 1940s before independence, is a bit of a puzzle. We pursue this issue in the next section.

Industrial Structure and Location

We saw in chapter 2 that urbanization accompanies economic development because employment and other inputs shift from the predominantly rural agricultural sector to the predominantly urban industrial and service sectors. The preceding section established a correlation between India's urbanization and its economic growth. The purpose of this section is to explore the relation between Indian urbanization and sectoral shifts.

When sectoral data for India are compared with those for other low-income countries, the puzzle noted in the preceding section grows. We saw in table 1-3 of chapter 1 that the percentage of the Indian labor force in agriculture did not decline between 1901 and 1971. We also saw that the percentage of the labor force in the predominantly urban manufacturing, construction, and service sectors hardly increased at all during the same seventy-year period. As was seen in table 3-1 of this chapter, however, the percentage of the population that was urban increased substantially during the period. These trends are somewhat at odds with the time series and cross-sectional data reported by the World Bank (1982), which show a closer correlation between urbanization and labor force shifts from agriculture to the manufacturing, construction, and service sectors. The Indian data also conflict with the finding of Chenery and Syrquin (1975), whose analysis of cross-sectional data disclosed development patterns similar to those reported in the World Bank study. The Indian data become less puzzling if one takes into account the decline in agriculture's labor force share between 1971 and 1981 (see table 1-5), but the puzzle remains.

To sort out the puzzle we must first analyze the urbanization of particular sectors. If the labor force has not shifted out of rural sectors, perhaps the sectors themselves have become urbanized. Table 3-2 shows the percentages of urban workers in each of the five sectors analyzed previously, for 1951, 1961, and 1971, as well as the percentages of urban workers in the total labor force for the same years and 1981. It is remarkable that the percentage of urban employment decreased in each sector between 1951 and 1961 and increased in each sector between 1961 and 1971. Yet, as was shown in table 3-1, the percentage of the population that was urban increased during both decades, and the increase in percent urban in 1961–71 was only slightly greater than it was in 1951–61. It is also remarkable that it was not until 1971 that majorities of manufacturing

Table 3-2. *Indian Urban Employment by Sector, 1951–81*

	Percentage of urban workers					
Year	*Agriculture*	*Mining and quarrying*	*Manufacturing*	*Construction*	*Services*	*Total labor force*
1951	2.8	13.5	44.2	48.0	55.2	14.0
1961	2.0	12.7	38.1	46.8	48.0	14.0
1971	2.8	16.5	52.2	50.5	58.2	17.7
1981	n.a.	n.a.	n.a.	n.a.	n.a.	19.8

n.a. Not available.
Source: Government of India (1951, 1961, 1971, 1981).

and construction workers were located in urban areas. The Indian census definition of manufacturing includes workers who repair and service manufactured products—persons who would be classified as service workers in the censuses of some other countries. This helps to explain the large percentage of manufacturing workers who are rural in Indian data.

Comparison of the 1951 and 1961 data in table 3-2 provides little insight into the causes of urbanization during that decade. It is not known why the census data show decreases in urban shares of employment during that period. But the 1951–71 and the 1961–71 comparisons show pervasive urbanization of all sectors except agriculture. Taking the two-decade period as a whole, we find that the increase in the percentage of urban employment shown in table 3-2 is roughly equal to the increase in percent urban shown in table 3-1. Comparing the percentage point increases of the five sectors, we also find that the pace of urbanization was greatest in the manufacturing sector, although manufacturing was still less urbanized than services in 1971.

The census data imply that Indian urbanization during the period 1951–71 resulted mostly from the urbanization of the manufacturing and service industries and, to a lesser extent, from the urbanization of mining and construction. In addition, as we saw in table 1-3, all four nonagricultural sectors made modest gains in their employment shares over the twenty-year period. Urbanization of a particular sector, say, manufacturing, does not imply that particular firms physically move. Urbanization may result from decreases in employment in rural manufacturing firms and increases in employment in urban manufacturing firms. The firms that shrink and the firms that grow may be engaged in altogether different activities. Similarly, the data on urban employment analyzed here do not necessarily mean that workers physically migrated, for changes in the sector of employment of a few workers may result in the reclassification of a community from rural to urban. Migration studies, however, show that Indian urbanization has been accompanied by large amounts of physical migration.

A recent study by the United Nations (1980) permits comparison of the Indian data in table 3-2 with data for other countries. Table 3-3 presents the relevant data from the U.N. study. The left column refers to the percentage of the total labor force in agriculture, which is the measure the U.N. study used as an overall index of development. The index goes from over 65, representing the lowest levels of development, to below 15, representing the highest levels of development. GNP per capita would have been a better index of development, but the U.N. study does not present it.

The entries in table 3-3 are based on fifty-nine observations for thirty-nine countries. The data are thus mostly cross-sectional, but they include two or more observations from a few countries, India among them. The countries sampled span the entire range of development levels. India was in the "65.0 or more" category in all census years from 1901 to 1971, as table 1-3 shows. The sectoral classification in table 3-3 is different from that in table 3-2, and the definitions used in the two tables are somewhat different, but the categories of data are similar enough to permit comparison.

The data in the top line of table 3-3 are broadly consistent with those in table 3-2. Both tables show only a negligible percentage of agricultural employment in urban areas, both show roughly half of industrial employment (dominated by manufacturing and construction in table 3-2) in urban areas, and both show nearly 60 percent of service employment in urban areas. Thus India's experience during the 1951–71 period is broadly typical

Table 3-3. *Distribution of Urban Employment in Thirty-nine Countries, by Sector and Stage of Development*

Country group (percentage of total labor force in agriculture)	*Percentage of urban workers*					
	Agriculture	*Industry*	*Professional and administrative services*	*Clerical and sales services*	*Traditional services*	*Unknown*
65 or more	3.8	50.0	51.2	57.0	59.0	33.7
50.0–64.9	5.7	53.7	59.7	64.9	60.9	51.3
35.0–49.9	13.2	60.1	74.0	75.3	70.8	53.4
15.0–34.9	14.5	67.2	77.3	78.1	74.8	59.1
14.9 or less	13.0	72.1	82.2	84.5	78.0	72.5

Source: United Nations (1980).

of that of other countries at a similar stage of development. A remarkably large share of manufacturing and construction is rural, and employment in services is more urban than that in industry. Furthermore, table 3-3 indicates that industrial employment becomes urbanized less rapidly than services at the next stage of development, where 50–64.9 percent of the labor force is in agriculture. If India continues to follow the worldwide pattern, industry will become only slightly more urbanized during the next decade or so.

Tables 3-4 and 3-5 show detailed data on urban and rural distributions of the labor force by sector for 1961 and 1971. One startling statistic is the decrease in the percentage of urban workers in nonhousehold, or factory, manufacturing between 1961 and 1971. In all seven other sectors, including the three service sectors, the urban share of employment increased. Not only are trade and commerce and transport, storage, and communication more urbanized than the larger other services sector, but also those two sectors grew as shares of total employment between 1961 and 1971, whereas other services shrank in total employment as well as in its share of employment.

The behavior of manufacturing employment in the 1961–71 decade illustrates the value of disaggregating its total into the household and factory components. Employment in household manufacturing decreased by almost 50 percent between 1961 and 1971, whereas employment in factory

Table 3-4. *Urban and Rural Employment by Sector, India, 1961*
(thousands of workers)

Sector	*Urban*		*Rural*			*Percentage of total*
	Number	*Percent*	*Number*	*Percent*	*Total*	*labor force*
Agriculture	2,651	2.0	128,491	98.0	131,143	69.5
Mining and quarrying	665	12.7	4,556	87.3	5,221	2.8
Household manufacturing	2,088	17.4	9,942	82.6	12,031	6.4
Nonhousehold manufacturing	5,540	69.5	2,435	30.5	7,975	4.2
Construction	964	46.8	1,096	53.2	2,060	1.1
Trade and commerce	4,308	56.8	3,345	43.2	7,654	4.1
Transport, storage and communication	2,125	70.4	894	29.6	3,019	1.6
Other services	8,088	41.3	11,485	58.7	19,572	10.4
Total	26,430	14.0	162,246	86.0	188,676	100.0

Note: Columns may not add up to totals because of rounding.
Source: Government of India (1961).

Table 3-5. *Urban and Rural Employment by Sector, India, 1971*
(thousands of workers)

Sector	Urban Number	Urban Percent	Rural Number	Rural Percent	Total	Percentage of total labor force
Agriculture	3,552	2.8	122,209	97.2	125,761	69.7
Mining and quarrying	859	16.5	4,361	83.3	5,220	2.9
Household manufacturing	1,589	25.0	4,763	75.0	6,352	3.5
Nonhousehold manufacturing	7,315	64.8	3,965	35.2	11,280	6.2
Construction	1,120	50.5	1,100	49.5	2,220	1.2
Trade and commerce	6,416	63.9	3,624	36.1	10,040	5.5
Transport, storage, and communication	3,192	72.5	1,210	27.5	4,403	2.4
Other services	7,963	50.5	7,809	49.5	15,772	8.7
Total	32,007	17.7	149,040	82.3	181,048	100.0

Note: Columns may not add up to totals because of rounding.
Source: Government of India (1971).

manufacturing increased by about 40 percent. Since factory employment is much more urban than household manufacturing employment, the increasing weight of factory employment implies an increasing urban share of total manufacturing employment, despite the falling urban share of factory employment.

There is no mystery, then, about the causes of Indian urbanization, at least between 1961 and 1971. In the five-sector employment classification all sectors—and especially the predominantly urban ones—became more urban. In the eight-sector classification only factory manufacturing employment became less urban. Since its share of total manufacturing and of total employment increased and it is among the most urban of the eight sectors, both total employment and total manufacturing employment became more urbanized.

The data in tables 3-4 and 3-5 also explain the steady and substantial expansion of manufacturing's GDP share. As we saw in table 1-4, manufacturing's GDP share rose from 12.2 to 14.4 percent between 1960 and 1970. That increase was undoubtedly a reflection of the labor shift, traced in tables 3-4 and 3-5, from low-productivity household to high-productivity factory manufacturing.

Increasing urbanization of predominantly urban sectors is to be expected during the course of economic development. The puzzle about the Indian data is why the urban share of factory manufacturing employment fell between 1961 and 1971, in opposition to the trend in all other sectors. The fact that factory employment has been the focus of government efforts to disperse employment from the cities raises the question of whether the trend toward reduced urban concentration is a result of the government programs.

Since 1956 it has been national government policy to disperse factory production from large cities and high-income states and districts to small towns, rural areas, and low-income states and districts. (Specific government programs to disperse industry are discussed in chapter 7.) Reduced urbanization of factory employment, such as that observed in the tables, is one of the goals of the government programs. To what extent are the trends caused by the government programs?

Sekhar (1983, p. 10) showed that there has been a decrease in concentration of manufacturing activity among states. In 1961, Maharashtra, West Bengal, Gujarat, and Tamil Nadu accounted for two-thirds of factory manufacturing value added and 58 percent of factory employment. In 1976 their share was down to 55 percent of value added and 52 percent of employment. In both 1961 and 1981 these states had about 29 percent of India's population. Presumably, government policies on industrial location played a role in these reductions in concentration, but it is not possible to say how much of the reduction resulted from government policy and how much from normal market forces. For example, improvements in the transportation system would permit manufacturing plants to locate in areas where wages are low, without loss of access to markets for outputs and inputs other than labor.

Sekhar (p. 83) also showed that between 1961 and 1971 there was no change in the location of factory employment by size of urban area. The concentration of factory employment in large urban areas was about the same in both years. This does not prove that government programs had no effect on factory employment by city size, because the concentration in large cities might have increased in the absence of policies to prevent it. In any case, government policies did not reduce the share of factory employment in large urban areas.

Sekhar (p. 84) went on to report that the share of backward districts in factory employment fell from the early 1960s until the early 1970s in the states for which such data are available.[1] Once again, it is not known what would have happened in the absence of government policies to locate factory employment in backward districts, but government policies did not increase the share of backward districts in factory employment.

It appears that government attempts to disperse industry probably have had little effect on industrial location by city size or by district but have had some effect on location by state. None of these conclusions refers

Table 3-6. *Percentage of Indian Main Workers Who Are Urban, 1961, 1971, 1981*

Sector	*1961*	*1971*	*1981*
Agriculture	2.1	2.9	3.5
Other	41.9	52.7	55.1
Total	14.2	17.9	20.7

Source: Government of India (1981, paper 3).

directly to the issue of urban versus rural location; however, city size and district are more closely related to the urban-rural effect than is location by state. If government policy has had little or no effect on industrial location by city size or by district, it is unlikely that government policy has had much effect on urban versus rural location.

Between 1961 and 1971, urban factory employment increased by about 1.8 million workers and rural factory employment increased by about 1.5 million workers. These numbers are small in the Indian context. They could be explained, at least in part, by government location policies or by changes in the criteria used to distinguish between urban and rural locations.

Preliminary data from the 1981 census permit tabulation of the percentages of main workers residing in urban and rural locations for the twofold sectoral classification employed in table 1-5. Once again the 1981 census presents comparable data for 1961 and 1971. Summary data are shown in table 3-6, which is comparable to table 1-5. Table 3-6 suggests a continuation through 1981 of the trend shown in table 3-2 for employment to become more and more urban in all industries. The most striking characteristic of the data in table 3-6 is the deceleration of employment urbanization, both in the "other" category and in total employment. Between 1971 and 1981 the percentage of main-worker employment that was urban increased by only 2.8 percentage points, compared with 3.7 percentage points the preceding decade. The percentage of the population that was urban, however, increased by 3.5 percentage points between 1971 and 1981, according to table 3-1, compared with 1.9 percentage points the preceding decade. The result was a slight drop in the urban work force participation rate during the 1971–81 decade.

Urbanization by State

The purpose of this section is to study the variation in, and determinants of, urbanization among Indian states. In the previous chapter, regressions from a large set of developing and developed countries were estimated to relate urbanization to development indices. In this section two of the regressions are estimated for Indian states. The two regressions are:

$$(3.1) \qquad U_{it} = a_0 + a_1 A_{it} + a_2 Y_{it} + a_3 Y_{it}^2 + a_t t$$

and

$$(3.2) \qquad \ln (U_{it}^{-1} - \bar{a}_0) = a_0 + a_1 A_{it} + a_2 Y_{it} + a_3 Y_{it}^2 + a_4 t$$

where

U_{it} = percent urban in state i at time t
A_{it} = percentage of total employment in agriculture in state i at time t
Y_{it} = real per capita income in state i in year t
t = the year in question.

The 1981 census gives U and A for all states for 1961, 1971, and 1981. For income, we used real per capita net domestic product (NDP) by state, base 1960–61. NDP data are available for only thirty-four observations: seventeen in 1961, eleven in 1971, and six in 1977–78. The latest year for which the income data are available is 1977–78, so we used those income data with the 1981 U, A, and t values.

The estimated combined cross-section, time series regressions are

$$(3.3) \qquad \begin{aligned} U_{it} = -10.758 &- \underset{(1.1022)}{0.1055A_{it}} + \underset{(3.6736)}{1{,}625.65Y_{it}} - \underset{(2.4197)}{13{,}826.64Y_{it}^2} \\ &- \underset{(0.4806)}{0.5953t} \qquad R^2 = 0.673 \end{aligned}$$

$$(3.4) \qquad \begin{aligned} \ln (U_{it}^{-1} - 1.9) = 4.0545 &+ \underset{(1.1819)}{0.0102A_{it}} - \underset{(3.9131)}{156.25Y_{it}} \\ &+ \underset{(2.6965)}{1{,}390.57Y_{it}^2} + \underset{(0.5882)}{0.6570t}. \qquad R^2 = 0.669 \end{aligned}$$

Absolute values of t-statistics appear in parentheses below the coefficient estimates. These regressions are analogous to regressions (2.6) and (2.7) in chapter 2. The R^2 values for (3.3) and (3.4) are high, but they are somewhat lower than those for the cross-country regressions. The coefficients of A, Y, and Y^2 have the same signs as in the cross-country regressions. The signs are those expected: U increases as A decreases and as Y increases. The signs of the coefficients of Y^2 mean that the effect of Y on U becomes small at high Y values and eventually becomes negative.

In the cross-country regressions, we were surprised to find that the A coefficient was more significant than the Y and Y^2 coefficients in some regressions. With the Indian state data in regressions (3.3) and (3.4), the opposite is true. (Elimination of A or substitution for it of percentage of employment in manufacturing reduced R^2 somewhat, but had little effect on other coefficients.) Comparison of Y and Y^2 in the regressions in this section with those in chapter 2 is difficult because the units in which Y is measured are different. The effects of Y on U, however, are not large in (3.3) and (3.4) because, for fixed A and t, U peaks at low values of Y. In (3.3), for example, U peaks at a Y value that is less than two times the average Y value and only about 10 percent greater than the largest Y value in the sample. That, of course, is not to be taken literally, since the turning point is outside the sample range, but it does mean that at income levels likely to be realized in coming years the calculated effects of income changes on U will not be large. Thus, despite A's lower significance level, changes in A are likely to have more effect than changes in Y in the regressions during coming years.

Appendix: Urban Concepts in Indian Censuses

Each country has its own definitions of urban concepts. In this appendix we comment briefly on the concepts employed in the Indian censuses. Two aspects of the census definitions need to be discussed: the criteria used to determine whether a given place is urban and the criteria used to ascertain the extent of an urban area. The first aspect concerns the number of urban people, the second aspect the number of urban areas. Each aspect will be discussed in turn.

Urban places

The Indian censuses since 1961 have been fairly consistent in their definition of an urban place.[2] A place is designated as urban if it meets either of two criteria: it has one of several urban forms of local government or it has certain urban characteristics. A set of characteristics sufficient to justify an urban designation is a population of at least 5,000, a minimum of 75 percent of the male labor force engaged in nonagricultural work, and a population density of at least 1,000 people a square mile. A few places not possessing all these characteristics but thought to be predominantly urban are also so designated.

No definition of an urban place is entirely satisfactory or precise. Ambiguity inevitably arises, especially with small places. Nevertheless, the Indian definition appears to be as precise and satisfactory as that employed in almost any other country. In addition, it should be noted that the problems of classification are more difficult in India than almost anywhere else. In 1981, India had more than 2,000 urban places with fewer than 20,000 people each. Under such circumstances many places must be near the margin between urban and rural, whatever reasonable definition is used.

In India, urban places with fewer than 100,000 people are typically referred to as towns. Places with 100,000 or more people are typically referred to as cities or as Class I cities. Beginning in 1961 the criteria for designation as a town were tightened. Earlier criteria, apparently applied rather consistently to census data from 1901 to 1951, had been less stringent than those listed at the beginning of this section. Between 1951 and 1961, 803 towns were declassified—that is, changed from urban to rural (Bose 1974, chapter 2). A town can be declassified, however, either because of a change in definition or, barring that, because its total population or number of urban workers declined. Apparently, changes in definition caused most of the declassifications in 1961, but no precise information seems to be available. In any case the 1961 census showed 3.9 million fewer people living in towns with fewer than 10,000 people each than the 1951 census—this during a decade in which the total urban population increased by 15.9 million.[3]

The new definition appears to have been applied consistently in the 1961, 1971, and 1981 censuses, but there is some lack of comparability with the 1951 census with regard to the number and population of small towns.

Between the 1961 and 1971 censuses the definition of "worker" was tightened. Starting in 1971 a certain number of hours in gainful employment was required for a person to be counted as a worker, whereas in earlier censuses one merely needed to have a gainful occupation. Presumably as a result of this change, the 1971 census reported 28.2 million fewer female workers than did the 1961 census, whereas the number of male workers increased by 20.0 million (Government of India 1971, table 1-4, pp. 6, 7). The change in definition may have brought changes in the number and population of small towns.

When describing urban places, Indian censuses have consistently employed six classes of population size, which will be referred to in subsequent chapters:

Class I	at least 100,000
Class II	50,000 to 99,999
Class III	20,000 to 49,999
Class IV	10,000 to 19,999
Class V	5,000 to 9,999
Class VI	less than 5,000.

The extent of an urban area

Most censuses identify urban areas that extend beyond the boundaries of local government jurisdictions. As urban places grow, the urban population may spill over local government boundaries and into what was formerly a rural area. Alternatively, two or more towns may grow together, so that they become one urban place. "Metropolitan area" is the name most commonly applied to such places.

The 1961 Indian census introduced the notion of a town group. The 1971 and 1981 censuses have employed the term "urban agglomeration." An urban agglomeration consists of one or more towns or cities and the adjoining urban outgrowths. In the 1981 census most large urban places are urban agglomerations, but some small urban places are also urban agglomerations.

References

Bose, Ashish. 1974. *Studies in India's Urbanization: 1901–1971.* New Delhi: Tata McGraw-Hill.

Chenery, Hollis, and Moises Syrquin. 1975. *Patterns of Development: 1950–1970.* London: Oxford University Press.

Government of India. 1951. *Census of India, 1951.* New Delhi: Census Commissioner of India.

———. 1961. *Census of India, 1961.* New Delhi: Census Commissioner of India.

———. 1971. *Census of India, 1971.* New Delhi: Census Commissioner of India.

———. 1981. *Census of India, 1981.* New Delhi: Census Commissioner of India.

Lal, Deepak. 1981. "Cultural Stability and Economic Stagnation: India, 1500 B.C.–1980 A.D." Washington, D.C.: World Bank, Development Economics Department. Processed. (To be published in 1986 by Oxford University Press as *The Hindu Equilibrium.*)

Mitra, Asok. 1978. *India's Population.* Vols. 1 and 2. Delhi: Abhinaw Publications.

Sekhar, Uday. 1983. *Industrial Location Policy: The Indian Experience.* World Bank Staff Working Paper no. 620. Washington, D.C.

United Nations. 1980. *Patterns of Urban and Rural Population Growth.* Population Studies no. 68. New York: United Nations.

World Bank. 1982. *World Development Report 1982.* New York: Oxford University Press.

Notes

1. Backward districts are defined by complex criteria by national and state governments. They are intended to be the least developed districts in the country. Backward districts include about 60 percent of the population. See Sekhar (1983, p. 62).

2. Definitions are presented at the front of the census volumes.

3. See Government of India (1961, statement 14, p. 38). If the number of new settlements grows slowly, gradual growth of existing settlements reduces the number in the smallest category.

4

City Sizes and City Growth in India

THIS CHAPTER IS CONCERNED with the relative sizes of India's urban areas and with the growth and sectoral structures of these areas. Both historical and regional comparisons are included. Although the term "city size distribution" is used throughout the chapter, the subjects of our study are generic urban areas: the set of contiguous urbanized places that makes up an urban complex. In small urban areas the generic urban area usually coincides with a local government jurisdiction, such as a city. In large urban areas the generic urban area often spills over the boundaries of legal cities.

The city size distribution is a matter of intense public and government concern in India. Of particular concern are the sizes and growth rates of the largest cities. Much popular and scholarly writing accepts as nearly axiomatic the premise that large cities are too large and grow too fast, at least in the absence of constraints imposed by governments. Many five-year plans and other government reports make the same claim. Since shortly after independence India's national policy has been to constrain the growth and sizes of large, and not so large, cities.

This chapter examines the positive, or factual, aspects of city sizes and growth. It attempts to build a foundation of facts and explanation from which normative or policy analysis can proceed fruitfully. Normative and government policy issues will be discussed in chapter 7.

Some Descriptive Measures

Indian census data on city sizes are complete and detailed for all the census years of the twentieth century through 1981. During the early census years of this century, the boundaries of generic urban areas coincided with those of local government jurisdictions, except in rare cases. In the period since independence, many urban areas have grown beyond the bounds of city jurisdictions. By 1981 most Class I cities—cities with a population of at least 100,000—were considerably smaller than their encompassing urban areas. The Indian census refers to such places as urban agglomerations, and these correspond roughly to what are sometimes referred to as metropolitan areas in other countries. Most of the calculations presented in this chapter are for urban agglomerations whose cores are Class I cities. In a few cases a city and its urban agglomeration are identical. To avoid circumlocution, urban agglomerations that contain a Class I city will be referred to as Class I cities, but the reference is to the urban agglomeration in all cases. An example of an urban area in which the city and the urban agglomeration are dramatically different is Calcutta. The population of the city of Calcutta was 3.3 million in 1981, whereas the population of the urban agglomeration was 9.2 million.

India has a rich diversity of city sizes. In 1981 its 156.2 million urban residents lived in 3,245 urban areas. Calcutta, the largest of these, had 9.2 million people, but 230 other urban areas had fewer than 5,000 people each. More than 2,000 urban places had fewer than 20,000 people each; twelve urban agglomerations had at least 1 million people each. Much writing on the allegedly excessive sizes of India's cities emphasizes the increasing percentage of the urban population living in Class I cities. The data, for selected census years from 1901 to 1981, are as follows:

	1901	*1921*	*1941*	*1961*	*1971*	*1981*
Percentage of urban population in Class I cities	25.7	29.4	37.9	50.8	56.2	60.4

As has been true of the statistics on other facets of Indian economic growth, the remarkable characteristic of these data is the smoothness and gradualness of the change they depict.[1]

The percentage of the urban population living in these relatively large cities has indeed grown. In 1981 the figure was about twice as large as it had been fifty years earlier and 2.4 times its value in 1901. The problem with using this measure as a guide in thinking about city size distributions is that it is difficult to imagine an economic growth process that would not lead to an increasing share of the urban population living in cities of at least modest size. As a country urbanizes, not only do its large cities grow—although typically less rapidly than the urban population—but also the number of cities with populations in excess of the threshold size increases.

By most easily available measures, India has among the most dispersed distribution of city sizes of any country in the world. India's largest urban area, Calcutta, contained 5.9 percent of the urban population in 1981. The only country in the world whose largest metropolitan area had a smaller share of the country's urban population was the Soviet Union, where 4 percent of the urban population resided in Moscow. In the low- and middle-income countries an average of 13 percent and 29 percent, respectively, of the urban population lived in the largest city in 1980, according to the World Bank (1982). In 1901, Calcutta contained 5.8 percent of India's urban population. The stability of this share over the eighty-year period is remarkable, and it indicates virtual equality between Calcutta's and India's urban population growth during the twentieth century.

We can broaden the scope of our comparisons with the following data on the 1981 populations and 1961–81 and 1971–81 annual growth rates of India's four largest cities:

	1981 population (millions)	*Annual growth rate (percent)*	
		1961–81	*1971–81*
Calcutta	9.2	2.3	2.7
Bombay	8.2	3.4	3.2
Delhi	5.7	4.4	4.5
Madras	4.3	3.9	3.0
India	658.1	2.2	2.2
India's urban population	156.2	3.5	3.8

In the two decades preceding 1981, the population of Delhi grew faster than India's total urban population and faster than the population of any other of India's four largest cities. The population of Madras also grew faster than the total urban population during the 1961–81 period, but from 1971 to 1981 the growth of the total urban population outpaced the population growth in all of the four largest cities except Delhi. Calcutta, the largest of the four, experienced the slowest growth during both periods. The combined share of the four largest cities in the urban population fell during the twenty-year period 1961–81.

None of the data presented above indicates that India has a growing or unusually large share of its urban population in its largest metropolitan areas. India's largest cities are of course very large, but not by the standards of large cities in other developing countries. Calcutta has about the same population as Seoul, and both Calcutta and Seoul are smaller than São Paulo and Mexico City. All four of these cities are smaller than Tokyo and New York. Nevertheless, India has two metropolitan areas that are among the largest dozen or so in the world. Calcutta is certainly the poorest of the world's very large cities.

Pareto Distributions of City Sizes

The statistics presented in the previous section are interesting, and they place large Indian cities in historical and international perspective, but they pertain to only a small part of the city size distribution. It is desirable to supplement them with comprehensive measures of the city size distribution—measures which are sensitive to changes in the relative sizes of urban areas throughout the distribution.

The best characterization of the city size distribution is a statistical frequency distribution that depends on city sizes throughout the distribution. The Pareto distribution has been estimated from data on city sizes in many countries. It is simple to calculate and to understand and has been found to fit city size data well, so it will be used here for our analysis of Indian data.

The Pareto distribution relates the rank of a city to its population in the city size distribution. Some technical comments on the distribution are presented in the appendix to this chapter. Here, we need only point out that the Pareto distribution contains two constants that must be estimated from data on city sizes. One of those constants, designated β in our notations, characterizes the dispersion of the city size distribution. β has been found in studies of city size in many countries and at many times to cluster around 1. The larger the value of β, the more nearly equal the populations of the cities at the various ranks. If β is very small, it means that most of the urban population is concentrated in a small set of large cities. Values of β in excess of 1.2 or less than 0.80 are rarely estimated. The natural dividing line for β is 1. If $\beta > 1$, the city size distribution is said to be dispersed.

β is a measure of relative city sizes. If the populations of all the cities in the distribution change by the same proportion, the value of β remains unchanged. Two countries in which cities of corresponding rank have populations that are the same proportion of the largest city inevitably have the same β, even though all the cities in one country may be much larger than those in the other.

The Pareto distribution is easy to estimate with data on the populations of cities of various ranks. We estimated it for a sample of urban populations reported by Mitra, Mukherji, and Bose (1980): the male populations of most of India's Class I cities in each census year from 1901 to 1971, except 1941. We supplemented these statistics with data from the 1981 census on the male populations of all the Class I cities. Mitra, Mukherji, and Bose used data on only the male population because of the relationship that is presumed to exist between the size of the male population and the industrial structure of cities. We have used the same set of statistics because the data were readily available. Restricting the data on city size to statistics on the male population only, however, is a serious issue, and we will come back to it in the next section.

The β estimates are shown in table 4-1. The R^2's are all in excess of 0.944 and cluster around 0.99. The t-values are all in excess of 28 and most are greater than 50. Thus, these statistics are not presented. Addition of terms in P^2 or P^3 to the right side of the estimating equation (4.9) in the appendix at the end of this chapter did not improve the goodness of fit. We conclude that the Pareto distribution provides an excellent fit to the data and that no other simple distribution would improve the fit. It should be noted that the constant term in table 4-1 is the natural log of the parameter A in equation (4.7). To double this term is to square A.

The remarkable characteristic of the data in table 4-1 is the smoothness and gradualness of changes in the estimates. The β estimates show a gradually rising trend, with a peak in 1961 and slight decreases during the succeeding twenty years. Thus, during the first sixty years of the century, the relatively small Class I cities grew slightly faster than the relatively large Class I cities. During the past twenty years the relatively large cities have grown very slightly faster than the somewhat smaller cities. The estimates thus imply that India's large cities became more nearly equal in size from 1901 to 1961 and slightly less so from 1961 to 1981.[2]

Table 4-7 in the appendix to this chapter gives a rough idea of the

Table 4-1. *Pareto Distributions of Indian Cities, 1901–81*

Year	Ln A	β	*Number of observations*
1901	13.96	1.02	49
1911	14.02	1.02	44
1921	14.36	1.04	53
1931	15.62	1.13	62
1951	15.73	1.07	75
1961	17.31	1.17	99
1971	17.17	1.13	99
1981	17.30	1.10	216

Sources: 1901–71 estimates based on Mitra, Mukherji, and Bose (1980, appendix tables 1–7, column 4); 1981 estimates based on Government of India (1981, provisional table 4, column 5).

magnitude of the changes reported in India's city size distribution in table 4-1. All the β estimates in table 4-1 are between 1.0 and 1.2. Thus, those two columns in table 4-7 bracket the P_1/P_R values implied by the estimated distributions.

Table 4-2 compares actual and calculated values of P_1/P_2 and P_1/P_{10} for 1961 and 1981 to provide insight into how much more primate the Indian city size distribution became during the twenty-year period. Since β fell from 1961 to 1981, the calculated P_1/P_2 and P_1/P_{10} increased moderately. In fact, P_1/P_{10} did increase slightly, but P_1/P_2 fell. Bombay and Calcutta held ranks one and two in 1961, but their ranks reversed in 1981 (for both the male and the total population). Poona was the tenth largest city in 1971, and Nagpur was the tenth largest in 1981. Bombay and Calcutta have switched ranks several times during this century, but throughout the eighty-year period their male populations have remained much closer in size than would be predicted by a Pareto distribution estimated from any significantly large set of Indian cities.

It can be seen from table 4-2 that the change in β from 1961 to 1981 results in only slight changes in the calculated values of P_1/P_2 and P_1/P_{10}. Both the actual and calculated values of the relative sizes of India's large cities have changed little during the twenty-year period.

The conclusion to be drawn from these estimations is that the trend during this century has been toward a more nearly even city size distribution in India, with a slight reversal since 1961. It is probable that not much significance should be attached to the modest increase in the relative sizes of the largest cities that is estimated to have occurred since 1961. The estimates in table 4-1 do not, however, indicate any success of national government programs to disperse population from large to medium-size cities. Nor do they suggest that the relatively large sizes of the largest Indian cities are a legacy of colonialism. The trend was toward equality in city sizes during the colonial period, whereas it has been erratically toward growth of the largest cities since independence.

Table 4-2. *Actual and Calculated Relative City Sizes*

	P_1/P_2		P_1/P_{10}	
Year	*Actual*	*Calculated*	*Actual*	*Calculated*
1961	1.22	1.81	7.31	7.16
1981	1.11	1.88	7.57	8.11

Sources: Actual: Government of India (1981) and Mitra, Mukherji, and Bose (1980); calculated: from table 4-7 and equation (4.8).

Regional City Size Distributions

India is a large country with many urban areas. That makes it possible to study regional changes and differences in the city size distribution. There is no theory about the appropriate geographical unit in which to fit a Pareto distribution to city size data. Intuition suggests that the more economically integrated a geographical area is, the better the Pareto distribution should fit the city size data. Even if the intuition is correct, there is no way to know how integrated a geographical area should be in order to make estimation of a Pareto distribution plausible. In fact, scholars have fitted Pareto distributions to regional city size data in the United States and elsewhere and have obtained results nearly as good as those obtained at national levels. Indian states are inappropriate units of analysis for city size distributions. Not only are there too few urban areas in many states, but also a few urban areas, notably Delhi and Chandigarh, are not in states.

Many Indian demographic, economic, and other studies divide India into five regions. The regions are defined somewhat differently from study to study. We have adopted the regional classifications used by Mitra, Mukherji, and Bose (1980) and other authors. They are:

North	*South*
Chandigarh	Andhra Pradesh
Delhi	Karnataka
Haryana	Kerala
Himachal Pradesh	Laccadive and Minicoy Islands
Jammu and Kashmir	Pondicherry
Punjab	Tamil Nadu
Rajasthan	
East	*West*
Assam	Gujarat
Arunachal Pradesh	Goa
Bihar	Maharashtra
Manipur	Nagar Haveli
Meghalaya	
Mizoram	*Central*
Nagaland	Madhya Pradesh
Orissa	Uttar Pradesh
Sikkim	
Tripura	
West Bengal	

Using the data on male populations from the Mitra, Mukherji, and Bose study, we are able to calculate Pareto distributions for each of the five regions for each of the census years for which national estimates were reported in table 4-1, with the exception of the eastern region for 1901, 1911, and 1921. At least two city populations are needed to estimate the two parameters of the Pareto distribution, and until 1931 there was only one city, Calcutta, in the eastern region that was large enough to be included in Mitra, Mukherji, and Bose's data. By 1931, five cities in the eastern region were included.

The estimates of the regional Pareto distributions are presented in table 4-3, and the estimates of the national distributions are repeated for comparison. Once again the R^2's and the t-statistics are large and thus are not presented in the table. The smallest R^2 is 0.792, and most are above 0.900. The smallest t-statistic is 4.5, and most are above 20.

Throughout the century four of the regions have been dominated by one of India's four largest cities: the East by Calcutta, the West by Bombay, the North by Delhi, and the South by Madras. Kanpur is the largest city in the central region, with a 1981 population of 1.7 million (0.9 million males). In 1981, Kanpur was India's eighth largest urban area.

The regional differences in the Pareto exponents are much larger than the temporal differences. In 1931, 1951, and 1961, the eastern region was the most primate of the five regions, which shows the extreme domination of that region by Calcutta until 1971. The domination was even greater before 1931. Until 1971 the western region was the second most primate region, which shows Bombay's domination of that region. The southern region is dominated by Madras, but for most of the census years of this century its Pareto exponent has been only somewhat smaller than the national exponent. Delhi's domination of the northern region has increased over the decades, which shows the growing importance of the national government. The central region is the least primate of the five regions, and its city size distribution is remarkably dispersed.

Although the regional disparities in Pareto exponents are large, they have become smaller as time has passed. In most years until 1961 the largest regional Pareto exponent was about twice the size of the smallest. In 1931 the largest was more than three times the smallest. By 1981 the largest was only 1.3 times the smallest. There are no analytical studies that indicate what forces might produce reductions in regional disparities among Pareto exponents. It is tempting to conjecture that such reductions reflect the increasing integration of the Indian economy. Suppose, for example, that particular economic activities are most productive in cities of particular sizes. The more mobile the inputs among cities, the more similar the city size distribution among regions. Some such phenomenon must be behind the increasingly similar regional Pareto distributions.

It will be recalled that the Pareto distributions reported have been esti-

Table 4-3. *Pareto Distributions by Indian Regions, 1901–81*

Year	*Entire country*	*North*	*South*	*East*	*West*	*Central*
1901						
Ln *A*	13.96	11.50	11.87	n.a.	8.06	15.24
β	1.02	0.950	0.962	n.a.	0.607	1.22
Number of observations	49	8	14	1	8	18
1911						
Ln *A*	14.02	10.92	10.94	n.a.	7.77	15.95
β	1.02	0.897	0.851	n.a.	0.589	1.284
Number of observations	44	8	12	1	6	17
1921						
Ln *A*	14.36	10.35	12.42	n.a.	8.66	15.97
β	1.04	0.835	0.975	n.a.	0.643	1.283
Number of observations	53	8	15	1	9	20
1931						
Ln *A*	15.62	9.99	13.95	6.61	8.15	18.48
β	1.13	0.784	1.085	0.491	0.607	1.483
Number of observations	62	8	22	5	7	20
1951						
Ln *A*	15.73	11.14	13.74	7.36	9.36	18.19
β	1.07	0.831	1.004	0.542	0.657	1.382
Number of observations	75	11	24	8	1	22
1961						
Ln *A*	17.31	11.51	15.50	11.05	11.67	17.81
β	1.17	0.837	1.127	0.792	0.813	1.319
Number of observations	99	12	28	15	19	25
1971						
Ln *A*	17.17	11.19	14.68	11.636	12.14	17.71
β	1.13	0.785	1.030	0.827	0.827	1.282
Number of observations	99	12	28	15	19	25
1981						
Ln *A*	17.30	13.09	15.88	13.82	12.98	16.12
β	1.10	0.907	1.082	0.955	0.862	1.122
Number of observations	216	30	66	37	39	44

n.a. Not available.

Source: Calculated from data in Government of India (1981) and Mitra, Mukherji, and Bose (1980).

mated from the male populations of selected Class I cities. If, as has been concluded, the city sizes fit the Pareto distributions closely, estimation from samples of cities is adequate provided the sample sizes are large enough. Nevertheless, it would be interesting to know whether enlarging the sample to include smaller urban areas would affect the estimates.

Much more important is the exclusion of females from the estimates.

If the sex ratio were the same in all urban areas, the exclusion of females would affect the estimate of A but not that of the exponent β in the Pareto distribution. There were, however, only 935 females per 1,000 males in all of India in 1981. Almost no Class I city had as many females as males, and a few had fewer than 750 females per 1,000 males. It is possible that the sex ratio varies systematically by city size or by region.

To test these possibilities, we reestimated the Pareto exponents for 1981, since the data were readily available. Table 4-4 shows three estimates of the Pareto distribution for the country as a whole and for each of the five regions: the first set of estimates is repeated from the 1981 row in table 4-3 for comparison; the second is based on the total population of all Class I cities; and the third is based on the total population of all Class I and Class II cities. The inclusion of Class II cities—those with a population of at least 50,000 but less than 100,000—more than doubles the national and most of the regional samples. Again it is unnecessary to report the R^2's and t-values. The smallest R^2 in table 4-4 is 0.9129, and most are above 0.95. The smallest t-value is 18.9, and most are above 25.

The Pareto exponents for the total population of all Class I cities are extremely close to those for the male population. Certainly, the differences are not significant. Interestingly, all the regional exponents estimated for the total population exceed the corresponding estimates for the male population. This indicates that the female-to-male ratio is slightly higher in small Class I cities than in large Class I cities. Presumably, the reason is that males are somewhat more likely to migrate by themselves to large than to small cities.

The inclusion of Class II cities in the sample makes the exponent larger nationally and in all regions except the central. Thus, Pareto distributions estimated from cities in Classes I and II give less evidence of population concentration in the largest cities than do those estimated from Class I cities. The differences in the estimated Pareto exponents brought about by the inclusion of females and Class II cities are not large. For no geographical designation is the range of the three estimates in table 4-4 as great as 0.15, and only for the eastern region does it exceed 0.064. The estimates in table 4-4 suggest that the preceding analysis, which was based on estimates from data on the male populations of the Class I cities, is not far off the mark.

Determinants of City Size Distributions

We have seen that there are large differences in Pareto exponents among regions and moderate differences through time and that the trend has

Table 4-4. *Alternative Estimates of Pareto Distributions, 1981*

Region	Ln *A*	β	*Number of observations*
North			
Male population of Class I cities	13.09	0.907	30
Total population of Class I cities	13.74	0.915	30
Total population of Class I and Class II cities	14.44	0.968	56
South			
Male population of Class I cities	15.88	1.082	66
Total population of Class I cities	16.65	1.086	66
Total population of Class I and Class II cities	17.00	1.113	151
East			
Male population of Class I cities	13.82	0.955	37
Total population of Class I cities	14.56	0.969	37
Total population of Class I and Class II cities	16.28	1.102	85
West			
Male population of Class I cities	12.98	0.862	39
Total population of Class I cities	13.64	0.871	39
Total population of Class I and Class II cities	14.35	0.926	84
Central			
Male population of Class I cities	16.12	1.122	44
Total population of Class I cities	16.91	1.130	44
Total population of Class I and Class II cities	16.73	1.116	110
Entire country			
Male population of Class I cities	17.30	1.10	216
Total population of Class I cities	18.12	1.109	216
Total population of Class I and Class II cities	18.27	1.121	486

Source: Calculated from Government of India (1981).

been toward regional convergence in recent decades. What explains these regional differences and temporal changes?

Economic analysis is of little help on this matter. Several empirical studies, notably those of Henderson (1982) and Rosen and Resnick (1980), have shed some light on the subject. Both studies regressed measures of country city size distributions on country characteristics, employing data from developed and developing countries for these cross-sectional regressions.

Both studies found that countries with large populations, or large urban populations, have relatively even city sizes. Whereas one large city may serve the specialized production needs of a small country, several large cities are needed for this purpose in a large country. India is a large country, and it is not surprising that it contains several large cities. Similarly, it is not by accident that India's four largest cities are near the four geographical extremes of the country. Rosen and Resnick also found that countries with a large per capita GNP have relatively even city size distributions. High income presumably means that demands are great enough to support the production of a relatively large number of large and intermediate-size cities. Henderson found that the city size distributions of countries with federal systems of government are much more nearly even than those of countries with unitary governments. His explanation is that states or provinces have constitutional authority in federal countries and can bid for funds for infrastructure and government services for many cities, whereas such funds are likely to go to the national capital in a unitary system. India is of course a federal country, and that variable does not differ among regions. The studies by Henderson and by Rosen and Resnick differ from each other as to specification of the dependent variable and as to choice of independent variables. Several independent variables either are not available or do not vary among regions or through time within India.

It would not be fruitful to undertake a statistical analysis of the eight national estimates of the Pareto exponents in table 4-1. The sample of estimates is too small, and the relevant variables change too slowly. What is needed is a combined cross-sectional and time series analysis that uses all the estimates in table 4-3 to enlarge the sample and to introduce variability.

Income is the key variable that other scholars have found to be important but that cannot be supplied from Indian census data. No income measure is available for Indian cities. Income measures are available by state, but only for the postindependence period. Both of the studies referred to above found high income to be a significant correlate of a relatively even distribution of city sizes.

Total population is the second variable found in earlier studies to be consistently correlated with the city size distribution. The city size distribution becomes more nearly equal as population increases. Not surprisingly, the effect has been found to be nonlinear. As population becomes large, the effect of additional population on the city size distribution diminishes. Regional population statistics can be compiled from Indian census data for the entire eighty-year period covered by this study. After some experimentation it was found that the square root of population provided the best results.

Henderson found that a high ratio of manufacturing employment to service employment was correlated with a relatively even distribution of city sizes. His interpretation of this finding is that the existence of a resource-based sector that is large in relation to the mobile, or footloose, service sector permits several large cities to develop. The regressions reported below include the percentage of the labor force in manufacturing. Less adequate results were obtained using Henderson's variable.

The discussion in the previous section suggests that some of the regional variation in Pareto exponents may not be explained by available regional variables. If so, the inclusion of regional dummies is appropriate.

The resulting regressions, with and without the regional dummies, based on the regional β estimates in table 4-3 are

$$\begin{aligned}(4.1)\qquad \beta_{it} = 1.055 &+ \underset{(2.3835)}{0.6984P_{it}^{1/2}} + \underset{(1.591)}{0.0115I_{it}} \\ &- \underset{(9.385)}{0.4468N} - \underset{(6.227)}{0.2996S} - \underset{(12.324)}{0.6762E} - \underset{(10.064)}{0.7439W} \qquad R^2 = 0.91\end{aligned}$$

and

$$(4.2)\qquad \beta_{it} = 1.200 + \underset{(1.156)}{0.9138P_{it}^{1/2}} - \underset{(2.440)}{0.0260I_{it}} \qquad R^2 = 0.18$$

where

β_{it} = Pareto exponent for the i^{th} region and t^{th} year (treated as positive numbers)

$i = N, S, E, W, C$

$t = 0$ in 1901, 1 in 1911, and so on

P_{it} = population of the i^{th} region and t^{th} year

I_{it} = percentage of workers in manufacturing in the i^{th} region and t^{th} year

N = northern region

S = southern region

E = eastern region
W = western region
C = central region.

N, S, E, and W take the value 1 when i refers to that region and 0 otherwise. Only four dummies are needed; when i is the central region, its locational impact is included in the constant term. The four dummy variables simply permit the constant term, but not the slope coefficients, to vary by region. Absolute values of t-statistics appear in parentheses below the coefficient estimates.

The R^2 in regression (4.1) is remarkably large, whereas that in (4.2) is similar to the values obtained by Rosen and Resnick in their cross-country regression. Henderson, using a different dependent variable, obtained an R^2 of 0.69. The fact that the R^2 in (4.1) is much larger than it is in (4.2) means that much of the regional variation in β is not correlated with the regional variation in P and I. The regional dummies permit unbiased estimates of the coefficients of P and I to be obtained, and yield a large R^2, but they leave the regional variation in β unexplained.

Both regressions imply, and other studies have found, that larger populations lead to relatively even city size distributions. The positive coefficient of I in regression (4.1) implies, as Henderson found, that a large manufacturing employment share leads to an even city size distribution. The negative coefficient of I in regression (4.2) casts doubt on the specification in (4.2).

The regional dummy coefficients confirm what was observed in the previous sections. All four regional dummies have negative coefficients, which implies that the central region has the most nearly even distribution of city sizes even apart from the effects of P and I. Then, regional effects cause increasingly unequal city sizes as one moves from S to N to E to W.

The gradual convergence of regional city size distributions noted in the previous section raises the question of the extent to which convergence of the β's is attributable to the convergence in P and I values among regions. That question can be answered by reestimating regression (4.1) with regional dummies that are functions of time. In that specification two variables, N and Nt, represent the northern region. As before, N is 1 when β refers to the northern region and 0 otherwise, and t refers to the census year. Similar changes are made in the dummies for other regions. For example, if the coefficient of N were negative and that for Nt were positive, it would mean that the northern regional effect was negative in 1901 but became positive as time passed.

The new regression is

$$
(4.3) \quad \begin{aligned}
\beta_{it} = 1.287 &- \underset{(0.245)}{0.281 P_{it}^{1/2}} + \underset{(0.207)}{0.0020 I_{it}} \\
&- \underset{(4.060)}{0.404N} - \underset{(5.273)}{0.364S} - \underset{(7.354)}{1.110E} - \underset{(6.122)}{0.748W} \\
&- \underset{(0.336)}{0.0099Nt} + \underset{(0.998)}{0.0288St} + \underset{(2.875)}{0.0954Et} + \underset{(1.298)}{0.040Wt} \\
&+ \underset{(0.628)}{0.0135Ct}. \qquad R^2 = 0.950
\end{aligned}
$$

In this formulation the coefficients of the population and manufacturing employment share terms are both insignificant and the coefficient of the population term has the wrong sign. Once again the four regional dummies all have negative coefficients and are highly significant. Except for Et, the coefficients of the products of the regional dummies and time are insignificant. All but Nt have positive signs.

Regression (4.3) confirms the findings from table 4-4. In the early years of the century, when t was small, all four other regions had smaller β's than the central region. The β's for all regions except the northern have tended to increase through time. The time trend in the eastern regional dummy is much larger than it is in the others. The R^2 in regression (4.3) is somewhat larger than it is in (4.1).

The conclusion from these regressions must be that the increasingly similar regional distributions of city sizes have been the result of specifically regional influences and not the result of regional shifts in population and manufacturing employment.

Growth of Urban Areas

The preceding analysis of the distribution of city sizes in India at the national and regional levels has been based on the assumption that city sizes in a nation or region form a system whose component cities move in relation to one another. The stability of the city size distribution in India and in other countries, over long periods of time and despite economic, population, and urban growth, makes compelling the notion that city sizes can be understood only in relation to one another.

Yet relative city sizes do change. Cities change positions in the distribution, and the distribution itself changes. The city size distribution is a statistical summary of a complex set of phenomena that affect the growth of one city, several cities, or all the cities in a country or region. It seems worthwhile to step down from the level of national or regional aggregation and to provide some analysis at the level of individual cities. The basic purpose of the analysis in this section is to shed light on the determinants of growth of individual cities. In the following chapter, using a more detailed data set, we will go on to analyze city growth in a single state: Madhya Pradesh.

The basic idea of the analysis in this section is that cities grow to the extent that their ability to provide employment grows. This notion is founded on economic base theory, the assumption that most of the goods manufactured in a city are produced for export outside the city. If so, manufacturing employment in a city depends on demands elsewhere for products best manufactured in the city. In this view other production in the city, such as services, is undertaken to satisfy the demands of the city's residents. Economic base theory leads to the notion that manufacturing employment is an important determinant of the total population of a city but that causation in the opposite direction is unimportant.[3]

Economic base theory is at best an approximation. An unknown, but probably substantial, part of the manufacturing production in India's cities is consumed in the city of manufacture. In all likelihood, the proportion of manufactured goods consumed in the city in which they are produced increases with city size. In addition, some services produced in cities are exported outside the city in which they are produced. Bombay provides financial services for much of India; Delhi provides national government services for all of India; and many urban areas provide educational, agricultural, and other services for the residents of neighboring rural areas. Thus, the strength of the correlation between manufacturing employment and city size is to some extent a test of economic base theory. Relevant Indian data are available for every census year since 1901.

There are, however, several other variables that are likely to influence the growth of individual cities. First, our finding earlier in this chapter that regional dummy variables are important to an understanding of regional variations in city size distributions suggests that regional dummies might also be helpful in explaining the growth of individual cities. Second, many people in India and elsewhere believe that large cities grow fast merely because they are large. If so, the initial population of a city will be correlated with the city's growth rate. Third, the income in the state in which a city is located may influence the growth of the city: the higher the income, the faster the growth. Fourth, the farther a city is from other large cities, the faster it might grow.[4]

In principle, the dependent variable might be either a city's population or its growth rate. Both because of dimensionality considerations and to avoid introducing heteroscedastic error terms, which may cause poor coefficient estimates, we used the growth rate of the urban area as the dependent variable. The sample is the percentage growth of the male population from one census year to the next of the Class I cities employed in Mitra, Mukherji, and Bose's 1980 study—the same data we used earlier to estimate the Pareto distributions. A city was included for any decade for which the 1980 study presented population data at both ends of the decade. All growth rates available for any part of the period from 1901 to 1971 were included. The 1971–81 growth rates could not be included because 1981 manufacturing employment data were not yet available. More data are available for more recent decades.

State income data (real net domestic product per capita) were available for a significant number of states for the 1961–71 census decade only, and the distance variable was calculated for that decade alone. Thus, two regressions are presented. For the 1901–71 period a large sample of city growth rates can be regressed on the city's initial population, its growth rate of manufacturing employment, and the regional dummies. For the 1961–71 decade a smaller sample of city growth rates can be regressed on the above variables, income in the state in which the city is located, and the distance variable.

The regressions are

$$\begin{aligned} G_{it} = 4.052 &+ \underset{(7.500)}{0.105G_{Mit}} + \underset{(7.669)}{0.905G_t} + \underset{(0.542)}{0.560P_{i,t-1}} \\ &- \underset{(0.679)}{0.036P^2_{i,t-1}} + \underset{(0.747)}{2.955N} + \underset{(0.903)}{2.454S} \\ &- \underset{(0.534)}{2.244E} + \underset{(2.677)}{8.972W} \qquad R^2 = 0.355 \end{aligned} \tag{4.4}$$

and

$$\begin{aligned} G_{i\,71} = 20.251 &+ \underset{(6.450)}{0.287G_{Mi\,71}} + \underset{(0.896)}{1.329P_{i\,61}} - \underset{(1.125)}{0.0745P^2_{i\,61}} \\ &- \underset{(1.529)}{10.294N} - \underset{(0.055)}{0.283S} + \underset{(0.182)}{0.947E} + \underset{(0.684)}{4.733W} \\ &+ \underset{(0.639)}{0.0217\mathrm{NDP}_{i\,61}} - \underset{(1.904)}{0.000813P_{ic\,61}/d_{ic}} \qquad R^2 = 0.416 \end{aligned} \tag{4.5}$$

where

G_{it} = percentage growth of city i's male population between census years $t-1$ and t

G_{Mit} = percentage growth of city i's manufacturing (both factory and household) employment during the same decade

G_t = percentage growth of national population between census years $t - 1$ and t

$P_{i,t-1}$ = city i's population in census year $t - 1$

$\text{NDP}_{i\,61}$ = 1961 real per capita net domestic product in state in which city i is located

$P_{ic\,61}$ = 1961 population of the nearest Class I city to city i

d_{ic} = airline distance from city i to the nearest Class I city

and N, S, E, and W are the regional dummies (the central region is included in the constant term).

G_M and G have highly significant coefficients of the expected signs in regression (4.4), as does G_M in regression (4.5). G could not be included in (4.5) because it does not vary among the observations from which that regression is estimated. Rapid growth of a city's manufacturing employment and of the national population induces fast growth in that city's population. The coefficient of G in (4.4) is nearly 1, which means that, other things being equal, cities grow almost as fast as the national population. Some writers contend that rapid growth of the national population causes an even faster growth in the urban population. The opposite is true of the large sample of Class I cities on which regression (4.4) is based, since (4.4) says that national population growth leads to a somewhat smaller proportionate growth of cities. (Substitution of regional or national urban population for G in (4.4) led to slightly less satisfactory estimates.) The coefficient of G_M is highly significant but small in both regressions. If a city's population were proportionate to its manufacturing employment, as economic base theory predicts, the coefficient of G_M would be 1. Contrary to the indication of economic base theory, the estimated coefficients of 0.105 and 0.287 in regressions (4.4) and (4.5) imply elasticities of a city's population with respect to its manufacturing employment of about 0.1 and 0.3 respectively.

In both regressions it was found that the inclusion of terms in P and P^2 provided more satisfactory results than did the inclusion of alternative functional forms. In both regressions the coefficient of P is positive and that of P^2 is negative, which indicates that a large initial population stimulates growth at low initial population levels, but that the effect disappears at high initial population levels. Both equations indicate that the effect of large initial population on population growth turns negative at modest initial population levels: about 780,000 in (4.4) and 890,000 in (4.5).[5] This implies that a large initial population discourages further city growth starting at initial populations somewhat below 1 million. This evidence is in opposition to the view that large cities grow rapidly because they are large. Unfortunately, the significance levels are not great for the P and P^2 coefficients in (4.4) and (4.5).

In both regressions the West is the region most favorable to city growth. But none of the coefficients of the regional dummies except that of W in (4.4) is highly significant.

The coefficient of the state income variable in (4.5) has a positive sign, as should be expected. It indicates that cities grow faster in high-income than in low-income states. The explanation for this is undoubtedly the same as that given in an earlier chapter for the correlation between urbanization and income by country. High income promotes urbanization, and whatever promotes urbanization causes rapid city growth. It should also be mentioned that regressions not reported show that cities grow rapidly if they are in states with high agricultural wage rates. City and agricultural incomes no doubt interact in causing rapid city growth, but regressions containing agricultural wages were not quite as satisfactory as (4.5).

The coefficient of the distance variable in (4.5) also has the expected sign. Other things being equal, the farther cities are from the nearest Class I city, the faster they grow. Regression (4.5) confirms the indication of gravity models that the effect of distance from the nearest Class I city on a city's growth depends on the population of the nearest city. Other things being equal, a large nearby city deters growth more than a small nearby city. The effect is quite significant. The proximity of cities other than the nearest Class I city may also affect the growth of a city, but we have not been able to study such effects.

Regressions (4.4) and (4.5) provide other insights into the determinants of individual city growth. As was true of such studies in other countries, a considerable amount of the variance in city growth rates is left unexplained by available economic or demographic variables. In addition, in regressions (4.4) and (4.5) a considerable amount of the explanatory power comes from the regional dummies, despite their uneven significance levels. The regressions leave unspecified the content of the regional influences: resource availability, state or national government policies, or other influences.

Regressions (4.4) and (4.5) suggest that as industrialization proceeds and spreads to a large set of cities, many cities will grow to substantial size. The regressions provide no support for the belief that large cities grow faster than small ones or that growth (at least beyond modest size) is fostered by size. The regressions do provide evidence to support the proposition that manufacturing employment stimulates city population growth, but the correlation turns out to be much less than that indicated by economic base theory.

Industrial Characteristics

We conclude this chapter with an analysis of industrial distribution by city size. Many Indian scholars and government officials are concerned about what they fear is an excessive concentration of manufacturing employment in large cities. The policy issue will be discussed in chapter 7. Here we provide some factual background.

Table 4-5 shows the percentage of all manufacturing employment located in the cities with the most manufacturing employment in census years 1901 through 1971. The cities with the largest manufacturing employment are not the same as the cities with the largest populations. For example, Bombay had the largest manufacturing employment, but not the largest population, in all the census years listed in the table.

All three columns show an almost uniformly upward trend until 1951 and a strongly downward trend during the succeeding two decades. The trends in the three columns are remarkably uniform, with the sole exception of the 1931 drop in column 1. The 1951 shares are all 3.76 to 3.92 times the corresponding 1901 shares and 1.72 to 1.91 times the 1971 shares. Thus, the trend of increasing and then decreasing shares has been more or less uniform among the ten cities with the most manufacturing employment, although the cities occupying particular ranks have changed.

Beyond a doubt, part of the phenomenon represented by the figures

Table 4-5. *Percentage of Total Indian Manufacturing Employment in Cities with Large Manufacturing Employment*

Year	*Bombay (largest)*	*Five largest cities*	*Ten largest cities*
1901	0.92	2.29	2.98
1911	1.25	2.33	3.08
1921	1.62	2.89	3.75
1931	1.38	3.05	4.19
1951	3.61	8.87	11.20
1961	3.31	7.73	9.71
1971	2.10	4.65	5.93

Note: "Largest" refers not to city size but to amount of manufacturing employment.
Sources: Mitra, Mukherji, and Bose (1980) and Lal (1981).

in table 4-5 is the urbanization of manufacturing that was discussed in the preceding chapter. But that does not explain the falling shares after 1951, because manufacturing urbanized rapidly after 1951. Another remarkable fact about the data in table 4-5 is the smallness of the shares. For example, table 4-5 shows that in 1971 Bombay had only 2.10 percent of India's total manufacturing employment, whereas it had 5.58 percent of the total urban population and 1.13 percent of the total population.

The most important question about the trends depicted in table 4-5 is the extent to which they reflect the success or failure of government programs to disperse manufacturing from large cities. Undoubtedly, market forces were at work to some extent. Manufacturing would naturally have dispersed from its traditional centers as other cities grew in size and improved their infrastructure. In addition, improvements in intercity transportation would have permitted manufacturing to disperse from consuming centers. Sekhar (1983) concluded that government dispersal programs have had little effect on manufacturing location by city size. Nevertheless, the strength of the dispersal of manufacturing from its traditional centers and the coincidence of the dispersal with government programs designed to foster it suggest that governmental policy may have influenced some of the postindependence dispersal recorded in table 4-5.

Additional light can be shed on this issue by one additional regression. Equation (4.6) shows the regression of the percentage of the male population that is employed in manufacturing (PPM) on total male population (P) and on time (t). The data are all the cities studied by Mitra, Mukherji, and Bose (1980) for census years 1901–71. Again the population data are given in hundreds of thousands.

$$\text{(4.6)} \qquad \text{PPM}_{it} = \underset{}{13.844} + \underset{(4.5093)}{4.1390P_{it}} - \underset{(0.9899)}{0.1176t}. \qquad R^2 = 0.041$$

As expected, the coefficient of population is positive and highly significant. Perhaps surprisingly, the coefficient of time is negative. Most surprisingly, R^2 is only 0.041. Thus, when the data are averaged over the long time and the many Class I cities covered by the Mitra, Mukherji, and Bose study, only 4 percent of the variance in the percentage of the male population employed in manufacturing is explained by total population and time. Hence, the size of a city's total population has only a minor influence on the percentage of males in manufacturing.

The final data to be presented in this chapter are in table 4-6, which shows the correlation coefficients between the percentage of the male population in manufacturing and the total male population for the cities in Mitra, Mukherji, and Bose's study. Despite the declining share of all manu-

Table 4-6. *Correlation between Percentage of Male Population in Manufacturing and Total Male Population, by Year*

Year	*Correlation*
1901	0.1617
1911	0.1861
1921	0.2736
1931	0.0624
1951	0.2683
1961	0.2180
1971	0.2693
All years	0.1975

Source: Calculated from data in Mitra, Mukherji, and Bose (1980).

facturing employment in the ten largest (in terms of manufacturing employment) cities since 1951, as shown in table 4-5, the correlation between city size and manufacturing share shows no tendency to become smaller in the most recent years in table 4-6.

The conclusion of this section is that there is hardly any relation between city size and manufacturing's share of total employment. In no census year in the twentieth century has male population explained more than 7.5 percent of the variance in manufacturing's share in male employment in India's Class I cities. Furthermore, there is no evidence of a time trend in this relationship. Manufacturing location must be explained by variables other than city population—presumably, variables concerned with the efficiency of locations.

Conclusions

India has a rich and dispersed distribution of city sizes. The distribution has persisted with only gradual changes throughout the twentieth century. The city size distribution varies greatly among India's five large regions. The regional distributions have become more similar to one another and to the national distribution during this century. Only a small part of the convergence of the regional distributions is explained by regional population growth and by industrialization.

The growth rates of individual cities are affected by industrialization and by income levels in the states in which the cities are located. The regional effect on individual city growth is slight.

There has been rapid dispersal of manufacturing employment from the largest manufacturing centers during the period since independence, and yet the correlation between city population and the share of its labor force in manufacturing has not decreased since independence.

Appendix: The Pareto Distribution

This appendix presents some technical material concerning the Pareto distribution that may be helpful to readers.

The Pareto distribution can be written

$$R = AP_R^{-\beta} \tag{4.7}$$

where R is the number of cities with at least P_R population, and A and β are constants to be estimated from data on the populations of cities of various ranks. Both A and β are positive. It can be shown that cities of various ranks are of nearly equal population if β is large. If β is small, the urban population is concentrated in a few cities that are large relative to the cities of lower rank. $\beta = 1$ is a natural dividing line. In that special case the Pareto distribution is sometimes referred to as the rank-size rule. In that special case A equals P_1, and P_R is P_1/R. The rank-size rule has no normative significance.

Equation (4.7) implies the following relation between the largest and the R^{th} largest urban areas:

$$\frac{P_1}{P_R} = R^{1/\beta}. \tag{4.8}$$

Table 4-7 shows P_1/P_R for $R = 2$, 5, 10, 20, and 40, and for selected values of β between 0.7 and 1.3. The table shows why $\beta = 1$ is a natural dividing line for β values. When $\beta = 1$, P_1/P_R equals R. When $\beta > 1$,

Table 4-7. *Selected Values of P_1/P_R for the Pareto Distribution*

	β						
P_1/P_R	*0.7*	*0.8*	*0.9*	*1.0*	*1.1*	*1.2*	*1.3*
P_1/P_2	2.69	2.38	2.16	2.00	1.88	1.78	1.71
P_1/P_5	9.99	7.48	5.97	5.00	4.33	3.80	3.45
P_1/P_{10}	26.92	17.78	12.88	10.00	8.13	6.76	5.89
P_1/P_{20}	72.52	42.29	27.81	20.00	15.27	12.02	10.04
P_1/P_{40}	195.41	100.50	60.02	40.00	28.70	21.37	17.12

Note: This range of β values includes most of those reported in the text.

P_1/P_R is less than R. When $\beta < 1$, P_1/P_R is greater than R. Of course, the further β and R are from 1, the further P_1/P_R is from R.

In one of the most elaborate studies of city sizes, Rosen and Resnick (1980) fitted the Pareto distribution to 1970 data for the fifty largest cities in each of forty-four countries. In their sample, β ranged from a low of 0.81 for Morocco to a high of 1.96 for Australia, with an average of 1.14. Most of Rosen and Resnick's analysis is based on Pareto distributions estimated from data from legal cities, not metropolitan areas. This biases estimates of β upward, since large metropolitan areas are larger relative to their legal cities than are small metropolitan areas or urban areas below metropolitan size.

Regression (4.7) is easy to estimate if logarithms are taken of both sides:

$$\ln R = \ln A - \beta \ln P_R. \tag{4.9}$$

Tables in the text report least squares estimates of $\ln A$ and of β from (4.9).

References

Government of India. 1981. *Census of India, 1981.* Series 1, paper 2. New Delhi: Census Commissioner of India.

Henderson, J. Vernon. 1982. "The Impact of Government Policies on Urban Concentration." *Journal of Urban Economics* 12, no. 3 (November): 280–303.

Lal, Deepak. 1981. "Cultural Stability and Economic Stagnation: India, 1500 B.C.–1980 A.D." Washington, D.C.: World Bank, Development Economics Department. Processed. (To be published in 1986 by Oxford University Press as *The Hindu Equilibrium.*)

Mitra, Asok, Sekhar Mukherji, and Radendran Bose. 1980. *Indian Cities.* New Delhi: Ashimar Publications.

Rosen, Kenneth, and Mitchell Resnick. 1980. "The Size Distribution of Cities: An Examination of the Pareto Law and Primacy." *Journal of Urban Economics* 8:165–86.

Sekhar, Uday. 1983. *Industrial Location Policy: The Indian Experience.* World Bank Staff Working Paper no. 620. Washington, D.C.

World Bank. 1982. *World Development Report 1982.* New York: Oxford University Press.

Notes

1. The statistics in this paragraph are from Government of India (1981).

2. The 1981 sample is much larger than those for earlier years. We reestimated the 1981 Pareto distribution using the ninety-eight of the ninety-nine cities in the 1971 census that are also included in the 1981 census. The estimates are $A = 16.52$ and $\beta = 1.04$. Comparison of these estimates with those in table 4-1 shows that the cities that were already large in 1971 became more primate (β fell) by 1981, whereas the newcomers to the set of 1981 Class I cities made the overall distribution nearly as even as it had been in 1971.

3. Recent studies employ more elaborate concepts of basic employment, but these concepts cannot be applied with Indian data.

4. Airline distance, measured from maps, was used as the variable. Road or rail measures might be better, but they are not easily available.

5. P is in units of 100,000 in regressions (4.4) and (4.5).

5

Urbanization and City Characteristics in Madhya Pradesh

THIS CHAPTER PRESENTS a detailed analysis of the determinants of growth of individual cities in the state of Madhya Pradesh. We find that a simple equilibrium model of city growth is remarkably successful in explaining intercensal differences in city growth rates and that the impression that growth rates are related directly to city size has almost no empirical basis.

In the preceding chapter our analysis of the determinants of city growth was based on a national sample of large Indian cities. Here, we narrow the focus to a more homogeneous region than India in its entirety and limit our study to the cities of a particular state. We chose Madhya Pradesh because it has a large number of cities and sizable towns but is not dominated by a particular urban agglomeration. In addition, Madhya Pradesh is part of the central Indian hinterland, a region that has been the subject of relatively little empirical economic research.

The 1971 census offers extensive information on the labor force and the industrial and infrastructural characteristics of the cities of Madhya Pradesh. Data from recently published parts of the 1981 census make it possible for us to calculate city growth rates for the 1971–81 decade as well as for 1961–71. Using both data sets we can relate city growth to a wider variety of urban characteristics than has been possible in previous studies of city growth in developing countries.

Note: We thank Kathryn Anderson, John R. Hansen, Koichi Mera, Rakesh Mohan, Felipe Morris, and Koichi Nakajima for their valuable comments, and Lynne Sherburne for her suggestions and diligent research assistance.

In the first section of this chapter we describe the basic characteristics of the cities of Madhya Pradesh and present a preliminary analysis.[1] In the next section we explain the theory of urban migration used in our subsequent analysis of city growth in Madhya Pradesh. The main empirical analysis of the chapter is presented in the final section.

Characteristics of the Cities of Madhya Pradesh

Madhya Pradesh does not contain any of India's largest cities, but it does contain several large cities. In 1971, Madhya Pradesh had eleven Class I and thirteen Class II cities. While nearly all of these cities grew at rates in excess of the national population growth rate, most experienced small declines in their growth rate in the 1970s compared with the 1960s. The 1971–81 average growth rates of the Class I and Class II cities of Madhya Pradesh slightly exceeded the unweighted average growth rates experienced by India's twenty largest urban areas.[2] These twenty urban areas averaged a decadal growth rate of 43.2 percent, a figure just slightly greater than the average recorded for the Class II cities of Madhya Pradesh.

The data in table 5-1 point up a few characteristics of the urban structure of Madhya Pradesh. Class I cities tended to grow at slightly higher rates than Class II cities during the 1971–81 decade. Moreover, the largest Class I cities grew faster still: those with a 1971 population in excess of 200,000 experienced an unweighted average 1971–81 growth rate of 51.1 percent, while the figure for the remaining Class I cities was only 44.4 percent. These statistics seem to confirm the hypothesis that forces related to city size are powerful determinants of city growth. The regressions reported below, however, indicate that this is a superficial interpretation. The data in table 5-1 also show that city growth rates in Madhya Pradesh varied considerably and that five of thirteen Class II cities grew more rapidly than did the average Class I city. Obviously, initial size is not the sole determinant of a city's population growth rate. This impression is bolstered by the data in table 5-2: the growth rates of the forty-two Class III cities for which data are available. Class III cities—cities with a population of at least 20,000 but less than 50,000—on average grew more rapidly than Class II or Class I cities.

We begin by considering two popular hypotheses: that great initial city size causes rapid city growth and that "growth pole" effects are important determinants of a city's growth. The presence of growth pole effects is frequently taken to imply—and our findings in chapter 4 suggest—that proximity to a large city will stunt the growth of small cities. If so, a city's growth rate will be positively related to its distance from the nearest

Table 5-1. *Madhya Pradesh Class I and Class II Cities, Arranged in Order of 1971–81 Population Growth Rate*

City	Population, 1971 (thousands)	Population growth rate (percent) 1971–81	1961–71
Class I cities			
Bhilai Nagar	174	83.2	102.5
Bhopal	385	74.7	72.6
Raipur	206	64.6	47.4
Indore	561	47.4	42.0
Jabalpur	442	46.8	49.8
Gwalior	394	37.8	35.1
Sagar	127	37.1	49.1
Bilaspur	131	36.7	50.8
Ujjain	209	35.2	44.7
Burhanpur	105	34.0	28.3
Ratlam	119	31.2	36.3
Average	259	48.1[a]	50.8[a]
Average of top five	354	63.3[a]	62.9[a]
Average of bottom six	181	35.3[a]	40.7[a]
Class II cities			
Durg	71	67.6	50.2
Dewas	52	60.7	50.0
Satna	62	55.1	63.4
Rajnandgaon	56	54.7	25.0
Shivpuri	51	48.8	77.3
Murwara	63	47.5	35.7
Rewa	69	45.3	60.6
Chhindwara	54	40.4	43.7
Mandsaur	57	36.3	35.9
Khandwa	85	34.0	34.5
Damoh	60	27.8	28.6
Jabalpur Cantonment	50	21.1	22.4
Mhow Cantonment	64	18.8	32.7
Average	61	42.9[a]	43.1[a]
Average of top six	59	55.7[a]	50.3[a]
Average of bottom seven	63	32.0[a]	36.9[a]

a. Unweighted average.

Source: Government of India (1981). The 1971 populations are calculated from the 1981 census figures for 1981 population and 1971–81 intercensal growth rates. These figures differ slightly from the population statistics reported in the 1971 census.

Table 5-2. *Madhya Pradesh Class III Cities, Arranged in Order of 1971–81 Population Growth Rate*

City	Population, 1971 (thousands)	Population growth rate (percent) 1971–81	1961–71
Class III cities			
Korba	33	115.5	162.8
Rajhara Jharandalli	49	111.8	11.8
Kurasia (urban area)	30	75.0	358.7
Nagda	33	73.8	100.1
Jagdalpur	37	72.2	80.9
Bhind	46	62.6	62.3
Ambikapur	24	61.4	55.8
Chhatarpur	32	59.8	45.7
Balaghat	33	59.5	75.6
Dabra	21	55.9	69.3
Morena	45	55.6	58.4
Hoshangabad	29	54.9	52.6
Shahdol	32	53.9	45.2
Guna	42	52.7	36.4
Tikamgarh	28	51.8	36.3
Vidisha	43	51.7	55.9
Betul	31	50.0	55.4
Itarsi	47	48.3	39.4
Dhanpuri	20	46.3	92.2
Raigarh	48	45.4	30.1
Sehore	36	44.4	26.8
Bhatapara	22	41.0	29.1
Seoni	38	40.6	26.8
Neemuch	50	38.2	37.1
Panna	24	38.2	45.6
Mandla	27	36.7	41.5
Dhar	36	35.1	27.7
Shajapur	25	34.8	45.5
Narsimhapur	26	34.0	42.4
Harda	29	32.2	27.9
Datia	37	31.9	27.2
Dhamtari	43	28.7	37.4
Sironj	22	28.1	29.6
Khargone	41	27.7	34.8
Jaora	37	26.8	20.4
Pandhurna	23	26.1	26.6
Barwani	22	25.6	26.7
Bina Etawa	33	25.2	21.8
Sagar Cantonment	27	19.6	42.5
Badnagar	20	19.3	16.8
Khamaria	42	11.9	37.7
Average	32	46.4[a]	53.6[a]
Average of top twenty-one	35	62.0[a]	73.9[a]
Average of bottom twenty	31	30.1[a]	32.3[a]

a. Unweighted average.
Source: Same as for table 5-1.

large city. In contrast, extreme proximity—location in the suburbs of a larger city—can be expected to have a positive impact on the growth rate of a city. This is because as rents rise in the central city decentralization of the urban area accelerates. In addition, if the entire urban area is growing rapidly as a result of its already large size, the suburbs will grow rapidly as well. Thus, proximity to a large city may spur the growth of extremely close towns and hinder the growth of moderately distant small cities.

These distance effects, along with the alleged positive effect of city size on its growth rate, imply

$$G_{it} = f(P_{i,t-10}, d_i, d_i^2) \qquad f_1, f_3 > 0 > f_2 \tag{5.1}$$

where

G_{it} = intercensal percentage growth rate of city i between census years $t - 10$ and t

$P_{i,t-10}$ = city i's base year population, in thousands

d_i = road distance in kilometers to the nearest Class I city.

A variety of functional forms was tried in considering these hypotheses. Both 1961–71 and 1971–81 growth rates were used as dependent variables. The coefficient estimates had the predicted signs but did not have statistically significant values. The coefficient estimates for the population variables rarely differed from zero, even at low confidence levels. The best of these poor fits was for the 1961–71 growth of Class I, II, and III cities:

(5.2)

$$\begin{aligned} G_{i\,71} = 42.57 &+ \underset{(1.56)}{0.195P_{i\,61}} - \underset{(-1.39)}{0.000473P_{i\,61}^2} - \underset{(-1.72)}{0.229d_i} + \underset{(2.08)}{0.00102d_i^2} + v_i \end{aligned}$$

$R^2 = 0.16$; $F = 1.84$

number of observations = 44

where

$G_{i\,71}$ = percentage intercensal growth rate, 1961–71

$P_{i\,61}$ = 1961 city population, in thousands

d_i = road distance in kilometers to the nearest Class I city

and v_i is a random error term.[3] The figures in parentheses are t-statistics.

Equation (5.2) hardly demonstrates that initial population level is a major explanatory variable in determining the population growth rate, although all P coefficients are statistically significant at the 80 percent confidence level.[4] In (5.2), the effect of P on G peaks at a level of 206 (206,000), which indicates that any positive impact of city size on its growth rate disappears at modest sizes. Using mean values for d and d^2, equation (5.2) indicates that a city at the bottom of Class III (with

$P = 20$) can be expected to grow by 38 percent in the next decade, whereas a city with a population of 206,000 will grow by 54 percent. The absence of robustness in these estimates, however, suggests the need for caution in accepting these predictions. Equation (5.2) provides the signs anticipated in (5.1) for the "distance to nearest large city" variables and suggests the presence of some positive effects from proximity to a large city. In this regression the effect of distance on growth becomes positive beyond 112 kilometers, a length only just exceeding the average distance to the nearest Class I city, 105 kilometers.

Restricting the sample to Class I and II cities does not improve the fit, nor does it alter the coefficient signs. Similar patterns are observed when we perform regressions on Class III cities alone. Our conclusion from these regressions is that city size is not among the major forces driving the growth of particular cities; nor are the underlying forces that determine growth strongly correlated with city size. If they were, we would observe statistically significant coefficients on the P variables, since they would serve as instrumental variables for the omitted underlying forces in these regressions.

If the existence of so-called runaway cities cannot be demonstrated empirically, why has the theme been so popular in the Indian literature? A likely reason is given by Mohan and Pant (1982). Analysis of Indian urban growth has typically been based on cross-tabulations of the growth of cities by size class. With a gradual upward trend in city size, the number of cities in the largest size class (Class I) rises relative to the numbers in other classes.[5] Consequently, the growth rate of the population living in Class I cities tends to be high since it is augmented by new entrant cities and suffers no exits. Population growth differentials are smaller when one calculates growth according to the city population class in the base year rather than the terminal year.

While the growth of the cities of Madhya Pradesh may be linked only tenuously to their geographical isolation and depends little on their initial sizes, the cities' 1971–81 growth rates do follow the pattern of their earlier growth rates. The 1961–71 city growth rates explain nearly half the variance in 1971–81 growth rates. The $G_{i\,71}$ coefficient estimates are affected little by the inclusion of P and d terms, whereas the coefficients of those variables are even less statistically significant than before.[6] In all Class I, II, and III cities for which data are available, we find

$$G_{i\,81} = 16.8 + \underset{(7.45)}{0.643} G_{i\,71} + v_i. \tag{5.3}$$

$R^2 = 0.48$; $F = 55.4$

number of observations $= 60$

Perhaps reassuringly, the coefficient on the 1961–71 city growth rate is below unity, which indicates that any impact of current growth on future growth is not explosive. That is, the equation $G_{it} = 16.8 + 0.643G_{i,\,t-10}$ has a stable long-run growth rate solution: $G = 47.059$ percent, which implies a doubling of city population every eighteen years. More important, these findings, coupled with the lack of significance of the population level coefficients, suggest that economic and geographic traits are responsible for the observed growth rate differentials and that such traits vary among cities but are stable over time.

Note that splitting the city sample does not reduce the impact of previous city growth rates on current rates. When the sample is restricted to Class I and II cities, we find[7]

(5.4)
$$G_{i\,81} = 17.0 + \underset{(4.54)}{0.628}G_{i\,71} + v_i$$
$R^2 = 0.52$; $F = 20.58$
number of observations = 21

and

(5.5)
$$G_{i\,81} = 34.7 + \underset{(4.49)}{0.593}G_{i\,71} - \underset{(-1.88)}{0.348}d_i + \underset{(1.56)}{0.00136}d_i^2 + v_i.$$
$R^2 = 0.63$; $F = 9.55$
number of observations = 21

The figures in parentheses here and below are t-statistics. Restricting the sample to Class III cities yields

(5.6)
$$G_{i\,81} = 16.9 + \underset{(5.87)}{0.647}G_{i\,71} + v_i$$
$R^2 = 0.48$; $F = 34.42$
number of observations = 39

and, as the only form where d was significantly different from zero,

(5.7)
$$\ln G_{i\,81} = 3.123 + \underset{(4.04)}{0.00938}G_{i\,71} + \underset{(1.69)}{0.00141}d_i + v_i.$$
$R^2 = 0.43$; $F = 13.43$
number of observations = 39; $\rho(G_{i\,71}, d_i) = 0.33$

The coefficients of $G_{i\,71}$ are strikingly similar in regressions (5.4), (5.5), and (5.6) and are extremely close to the coefficient in (5.3) for the full sample. This similarity suggests that similar economic and geographic variables are responsible for growth rate differentials within groups of large cities and groups of small cities alike. Note finally that the signs of the coefficients of d and d^2 in regression (5.5) are as anticipated in (5.1).

The patterns we find indicate that specific economic and geographic characteristics of cities should be highly relevant in explaining the determinants of city growth rates. We consider now the set of variables predicted to be relevant by a standard equilibrium model of city growth.

Migration Theory and City Growth

The growth of particular Indian cities is linked to household migration decisions, since India enjoys substantial labor mobility. Although the portion of total urban growth attributable to net rural-urban migration is not as large in India as in some other developing countries, it is substantial. Only about half of Indian city growth can be attributed to the natural increase in the number of inhabitants; the figure is much less in rapidly growing cities such as Bhilai Nagar, Bhopal, and Raipur. In this section a general theory of city growth is presented to suggest hypotheses. In the next section we estimate special cases with available data.

Neoclassical migration theory begins with the assumption that locational choices are based on constrained welfare-maximizing activities by households. If utility or well-being is higher in city i than elsewhere after adjusting for moving costs, mobile households locate in city i. If utility is a decreasing function of the city population (a condition that is implied if wages move inversely with the labor supply for a fixed capital stock) and if the number of mobile households is adequate to ensure that marginal conditions hold as equalities rather than inequalities, labor market equilibrium conditions are simple. They may be expressed as the standard requirement from the urban economics literature that

$$U_i = \bar{U}_A \text{ for all } i \in \{I\} \tag{5.8}$$

where

$\{I\}$ = the set of cities in the relevant region
U_i = utility for a household in city i
$\bar{U}_A$ = utility in rural areas, which is assumed to be constant.

The first issue that arises is whether one can anchor all urban functions at an exogenous rural level. In terms of numbers, the assumption seems plausible. In 1981 only 20.3 percent of Madhya Pradesh's 52 million inhabitants were urban; a decade earlier 16.3 percent of the population were urban. The assumption requires only that no major category of migrants be large relative to the group with similar characteristics that remains, so that the living standard at the place of origin stays approximately constant. In view of the rural preponderance in Madhya Pradesh, it seems appropriate to treat urban household welfare levels as exogenous and equal to a fixed rural level. The assumption of constant welfare is familiar both in development economics (dating from the formulation of Lewis's surplus labor model) and in U.S. intracity location models. Of course, $\bar{U}_A$ increases as the rural sector develops.

Our model concentrates on the determinants of a city's labor force growth. For this purpose it is useful to divide the total labor force L_i into four components:

$$L_i \equiv L_i^M + L_i^{AP} + L_i^{SV} + L_i^{SWS} \tag{5.9}$$

where

L^M = number of workers in manufacturing
L^{AP} = labor force engaged in agricultural processing
L^{SV} = the set of primary workers (household heads) engaged in services
L^{SWS} = the set of secondary workers (nonheads or part-time workers) engaged in services.

The practical distinction among these categories is detailed further in the discussion below.

Sector M is distinguished in the following ways: It produces a tradable commodity (or set of commodities) with a price determined in a nonlocal market. A simple but unnecessary assumption is that producers in city i set their product's price P^M parametrically.[8] A simplifying assumption is that M imports all nonprimary factors other than local infrastructure, at a fixed price. We assume the value added component of production depends on capital inputs K_i^M and labor inputs L_i^M, and that the productivity of these inputs depends on the extent of the city's social infrastructure S_i. Value added then depends on these three variables and the product price P_i^M.

Efficiency conditions can be used to derive an M-sector labor demand function that depends on variable input prices, product price, and stocks of fixed factors.[9] If it is assumed that Indian capital markets are imperfect and that exogenous local savings determine the level of capital accumulation, K_i^M and S_i are regarded as fixed factors. Labor demand is then equated to the implicit labor supply function, which is given by

$$U_i^M = U_i^M(w_i^M, Z_i) = \bar{U}_A. \tag{5.10}$$

That is, a worker's utility depends on his wage w_i^M and his consumption of publicly provided goods Z_i. If we assume that workers receive the value of their marginal product, we can substitute for w_i^M in equation (5.10). In that case, total differentiation of (5.10), given the usual restrictions on utility and production functions, enables us to express growth of employment in the M sector as an increasing function of increments to Z_i, S_i, K_i^M, and P_i^M.

Consider now the second source of employment in city i, agricultural processing (AP). As with M, AP produces largely for export, so that there is an exogenous component to the demand function facing it. In contrast, however, much of the AP sector's intermediate input A_i is provided by the surrounding hinterland; the price and aggregate output of A_i may be viewed as being determined by the fixed welfare assumption and by rural profit-maximizing behavior by landowners and peasants.

Thus, the AP-sector production function f^{AP} will depend on the sector's capital and labor inputs, on the extent of urban infrastructure, and on agricultural inputs, $A_i : f^{AP} = f^{AP}(L_i^{AP}, K_i^{AP}, S_i, A_i)$.[10] In turn, we assume that A_i depends on the available agricultural labor L_i^A and a set of effects characterized in appendix A of this chapter by a shift variable c_i. The shift variable accounts for such elements as the adoption of Green Revolution techniques and an increase in the amount of land devoted to cash crops, such as grain or rice. A city's geographical setting will also affect its equilibrium use of agricultural produce. Since A_i is agricultural output processed in city i, it is plausible that the larger city i's effective hinterland and the larger the share of the countryside devoted to Green Revolution cash crops, the greater the impact on A_i of changes in rural production conditions.

Determining the demand for AP-sector labor in a manner similar to that used in determining the demand for manufacturing labor, substituting into the fixed welfare constraint, and totally differentiating yield an expression for the change in equilibrium use of agricultural processing labor. This change is an increasing function of the infrastructure, capital stock, and own-price variables; it also rises with shifts that increase A_i.

The third sector consists of services—SV. As with the M sector, it is assumed that each city's production of services can be written as a function of the sector's labor and capital stocks and the city's infrastructural network. Regarding the capital stock as exogenous enables us to write a set of equations for the SV sector in each city analogous to those for the M sector. A critical distinction to be made, however, is that services are nontradable. Hence, the equilibrium price P_i^{SV} must be endogenous. We assume that the service demand increases with expanded real labor income in the M and AP sectors as well as with the level of infrastructural services (S_i and Z_i) provided by the urban authorities. Thus, the service demand function may be written:

$$
\begin{aligned}
(5.11) \qquad P_i^{SV} &= P_i^{SV}(U_i^M \cdot L_i^M + U_i^{AP} \cdot L_L^{AP}, Z_i, S_i, Q_i^{SV}) \\
&= P_i^{SV}[\bar{U}_A (L_i^M + L_i^{AP}), Z_i, S_i, Q_i^{SV}] \\
&\qquad P_1^{SV}, P_2^{SV}, P_3^{SV} \geq 0 > P_4^{SV}
\end{aligned}
$$

where Q^{SV} = quantity of services, and the P^{SV} numbered subscripts denote partial derivatives.

As is shown in appendix A, total differentiation of the SV-sector labor market equilibrium condition indicates that L_i^{SV} increases in capital stocks, prices of exportable goods AP and M, infrastructure provision, and the

AP-sector shift variable. The only exception to this conclusion is that dL_i^{SV}/dS_i and dL_i^{SV}/dK_i^{SV} can be negative if $\partial P_i^{SV}/\partial Q_i^{SV}$ is sufficiently large in absolute value. In that unlikely case an increment to K_i^{SV} or S_i causes such a fall in the price of services that the marginal revenue product of labor in that sector declines.

Labor supply in the final sector—secondary worker services (SWS)—may have both equilibrium and disequilibrium characteristics. Incorporation of secondary workers into a model of city growth also raises complexities if one views households of fixed composition as the appropriate decisionmaking unit. In such a case the equilibrium conditions would involve household rather than worker welfare equalization across urban and rural areas, thus adding to the system's simultaneity.

Alternatively, one may assume that welfare levels are equated for individuals. This characterization seems preferable for India, particularly in view of the characteristics of migrants. A large portion of Indian rural-urban migrants are prime-aged males who are single or migrate without families. Families tend to follow after some interval, presumably after the household head has obtained stable employment. Rural ties are maintained, though, and the marginal secondary worker is likely to have to choose between staying with his or her rural family component and moving to the urban family.

We therefore write

$$U(w_i^{SWS}, Z_i) = \bar{U}_A^{SWS}. \tag{5.12}$$

Since secondary worker services are also nontradable, equation (5.12) gives rise to an equilibrium labor force growth equation similar to that for SV-sector workers.

A third possibility is that the secondary worker labor market may not clear or there may be lags in the secondary worker adjustment process. Rapidly growing urban areas have low sex ratios (females per 1,000 males), although sex ratios rise over time. It is apparent from table 5-3 that slowly growing cities tend to have sex ratios that are higher than and that rise more slowly than those of rapidly growing cities.

The reason for this pattern is not obvious. Housing in mining boomtowns such as Korba may initially consist largely of barracks for male miners, and an influx of secondary workers and dependents may be delayed until family dwellings are erected. Or secondary workers may be hesitant to migrate until primary workers have established themselves. Such hesitation is particularly likely with secondary household members who are rarely in the labor force.

For simplicity, we assume that any lagged adjustment is confined to dependents outside the labor force. In that case the SWS market reaches

Table 5-3. *Changes in the Sex Ratio and Population Growth of Rapidly Growing and Slowly Growing Cities in Madhya Pradesh*

City	Population growth rate		Sex ratio		Change in sex ratio
	1961–71	1971–81	1971	1981	
High growth			0.840[a]	0.874[a]	0.34[a]
Bhilai Nagar	102.5	83.2	0.79	0.85	
Bhind	62.3	62.6	0.80	0.82	
Bhopal	72.6	74.7	0.83	0.87	
Dewas	50.0	60.7	0.90	0.89	
Durg	50.2	67.6	0.90	0.92	
Jagdalpur	80.9	72.2	0.80	0.90	
Korba	162.8	115.5	0.82	0.85	
Nagda	100.1	73.8	0.82	0.86	
Raipur	47.4	64.6	0.90	0.91	
Low growth			0.884[a]	0.898[a]	0.14[a]
Badnagar	16.8	19.3	0.90	0.91	
Barwani	26.7	25.6	0.89	0.90	
Bina Etawa	21.8	25.2	0.86	0.89	
Burhanpur	28.3	34.0	0.93	0.94	
Damoh	28.6	27.8	0.90	0.89	
Mhow Cantonment	32.7	18.8	0.80	0.84	
Ratlam	36.3	31.2	0.91	0.91	

a. Unweighted average.
Source: Same as for table 5-1.

equilibrium, and its labor growth equation takes a form identical to that of L^{SV} (implying that demand for *SWS* output is a function of $L_i^{AP} + L_i^M$).[11]

Besides accounting for labor force growth, total population growth incorporates additions to the stock of adults ("secondary adults") not in the labor force (N^{SA}) and children (N^C). Total population growth is, definitionally:

$$(5.13) \qquad dN_i \equiv dL_i + dN_i^{SA} + dN_i^C.$$

Assuming that there is a disequilibrium sex ratio adjustment coefficient Θ and that the pools of secondary adults and secondary service workers are predominantly female while *AP, M,* and *SV* workers are males, the ratio $(N_i^{SA} + L_i^{SWS})/(L_i^{AP} + L_i^M + L_i^{SV})$ should approach an equilibrium value σ. Part of the N_i^{SA} population increase, then, is an adjustment to a disequilibrium sex ratio. The second component of N_i^{SA} growth is its natural increment, which is equal to σ times the growth in employment in the *AP, M,* and *SV* sectors:

(5.14)

$$dN_i^{SA} = \sigma(dL_i^{AP} + dL_i^{M} + dL_i^{SV}) + \Theta(\Sigma L_i^j)\left(\sigma - \frac{(L_i^{SWS} + N_i^{SA})}{\Sigma_j L_i^j}\right) \qquad j = AP, M, SV$$

where the disequilibrium adjustment coefficient equals a coefficient multiplied by the male population.

A demographic model would help explain the increase in the final population category—nonadults—but data limitations prevent us from proceeding in this direction. Instead, we assume the nonadult population is proportionate to the adult female population. Moreover, while recognizing that sectors *AP, M,* and *SV* do not employ male workers exclusively and that $L^{SWS} + N^{SA}$ includes many males in any city, we further assume that

$$dN_i^C = b(L_i^{SWS} + N_i^{SA}). \qquad b > 0 \tag{5.15}$$

Combining the terms that describe the growth of different labor force and dependent groups enables us to obtain an expression for total population growth dN_i. It is an increasing function of growth in the prices of tradable commodities, of growth in the provision of urban infrastructure, of growth in the adjacent rural economy, and of capital stock growth in sectors *AP* and *M* and probably *SV* and *SWS*. The value of dN_i also rises with the number of secondary service workers, a stock that is itself dependent on past capital and infrastructure investments. Finally, as we can see from equations (5.14) and (5.15), dN_i also may rise or fall with the secondary adult population. This indeterminacy stems from the assumption that a low ratio of females to males generates a rapid influx of female migrants; a small female population, however, implies fewer births. The equation for the total population growth rate is given by (5.27) in appendix A.

City Growth Estimates

Although the model generating equation (5.27) includes many simplifying assumptions, the equation is too complex to be estimated from Indian data. Even for variables that are available, the data are rarely complete. We thus have a choice between omitting variables and excluding observations. We proceed by adding variables incrementally.

Data on increments to sectoral capital stocks and on changes in prices of export goods for each city cannot be obtained. Nevertheless, the importance of these terms suggests that a search for proxy variables would be useful. We begin by determining which types of manufacturing output have been favored in India's industrialization drive. In recent years India has experienced heavy investment in intermediate and capital goods industries, some basic raw material industries, and some high technology. But growth in investment in the traditional industries—small-scale manufacturing, textiles, and other light consumer goods—has lagged.[12] Price movements have been less consistent but have supported investment trends. The model presented in the previous section suggests that cities with a proportionately large intermediate and capital goods manufacturing sector will enjoy a relatively high growth rate.

A second force that spurs city growth is rapid output expansion in the adjacent agricultural sector. If regions were closed communities, rapid innovation and productivity growth in rural areas might hamper a city's growth. The model presented above assumes that the system is open and that both rural and urban areas realize rapid population growth.[13] Areas with growing marketed agricultural output are likely to specialize in grains and rice, crops which have benefited in recent decades from the technological advances of the Green Revolution. As production grows, the need for a larger processing industry increases, as does demand for local goods and services associated with increased consumption of intermediate inputs in agriculture and increased final consumption by rural households. A growing agricultural supply may also lead to falling input prices in the *AP* sector or economies of scale in agricultural processing; these trends generate additional labor demands. In sum, the model implies that cities with a dynamic rural hinterland experience relatively rapid growth.

Such growth is expected to be particularly great if the city's effective hinterland is large. If the city is viewed as a monopolistic competitor, the farther it is from other large cities and the more densely populated the nearby countryside, the larger its hinterland is likely to be. The *AP* shift variable vector thus is expected to contain an element c_{ik} that reflects city i's distance from other cities, and c_{ik} is hypothesized to vary positively with distance. This prediction must be hedged, however, since, as noted above, interurban trade from close cities may provide growth stimulus. Whether proximity stunts or promotes growth is a complex issue; at this point we simply recognize the possibility of ambiguity in the results.

Another factor that influences population growth rates in equation (5.27) is the sex ratio (females to males), which serves as a proxy for $(N_i^{SA} + L_i^{SWS})/N_i$. Unfortunately, the predicted sign is indeterminate.

The final factors that determine city growth rates for which information is available are per capita increments to infrastructure for industrial, commercial, and household uses. Because these public services are influenced by government policy, their effect on population growth should be of particular interest to policymakers. Unfortunately, estimation problems become most severe when public services are considered.

The first problem is that, although data on "stocks" of public services to both firms and households are available for 1971, increments in public services cannot be calculated. This problem is temporary; stock data for 1981 will become available when the census is completely tabulated. For now, we are able to calculate only intercensal increases in literacy rates for the fourteen Class I cities listed in the 1981 census and for the total urban population in each of Madhya Pradesh's forty-five districts. Changes in literacy rates do not, however, constitute an ideal proxy for increments in public service provision.

A second problem that emerges with regard to public services is that of simultaneity. To some extent, public services respond to demand—that is, to other independent variables such as dK_i^{AP} or dK_i^M. Were the problem this simple, we would simply suppress the dZ_i and dS_i terms and concentrate on estimating a reduced form. It appears likely, however, that both levels of and growth in per capita public services depend on population growth (in complex form) as well, so that the structure is nonrecursive.

Absence of data prevents us from simultaneously estimating equations via two-stage least squares in which population growth and public service growth are dependent variables. Moreover, the underlying simultaneity is likely to involve a lagged adjustment process. Cities with large influxes of population find it difficult to maintain services at an equilibrium level. Even though longtime residents in these cities may enjoy extensive public services, recent migrants may be compelled to wait to obtain a similar level of services. It is, however, largely the long-term level of services a city provides that concerns migrants and firms: presumably they are aware that explosively growing cities experience temporary constraints in providing services.

In view of these considerations, we have adopted the following estimation procedure. First, we estimate a public service supply function, using data from the 1971 census. This function depends on the demographic composition of the city (sex ratio, percentages of the population belonging to scheduled tribes and castes) and on the 1961–71 city growth rate. Rather than use the fitted values of public services in the 1971–81 population growth rate equation, though, we use the residuals. These residuals are taken to indicate the extent of an abnormal effort in providing public services—that is, we assume that the difference between actual and predicted levels of public service provision offers the best indication of the city's future performance in providing public services.

This procedure does not eliminate the problem of using stocks instead of first differences. Cities that have made an abnormally large effort to provide public services may not continue to do so. Implicitly, people are assumed to anticipate myopically that past trends will continue, at least with regard to the provision of public services.

A final point regarding public service provision concerns the applicability to India of the Tiebout model of decentralized local government service provision (Tiebout 1956). Public services are less attractive to migrants if they have to pay for them. Although Indian taxes are not carefully matched to the marginal cost of the public services consumed, it is likely that property values to some extent incorporate the value of public services provided to sites. Such capitalization is incomplete in the presence of imperfect land markets (including some squatter households that pay no rent) and services, such as schooling, that are only partially site-specific. Consequently, the specification embodied in equation (5.20) seems appropriate, although S_i should be understood to include only the noncapitalized portion of public services.

Table 5-4. *Major Industries and Exports of Class I and II Cities of Madhya Pradesh*

City	*Population growth rate, 1971–81 (percent)*	*Leading commodities manufactured*	*Leading exports*
Growth rate above 45 percent			
Class I			
Bhilai Nagar	83.2	*Pig iron, steel shapes*	Pig iron, steel shapes
Bhopal	74.7	Cardboard,[a] shoes	**Wheat,** pulses
Raipur	64.6	Poha, biri	**Rice,** *poha*
Indore	47.4	Cloth, *chemicals, medicines*	Cloth, biri
Jabalpur	46.8	Biri, saris	Gram, masoor
Class II			
Durg	67.6	Poha, oil	**Paddy,** brass metals, utensils
Dewas	60.7	*Water pumps,* cloth	Water pumps, cloth
Satna	55.1	*Cement, ironstick*	Cement, stone
Rajnandgaon	54.7	Pulses, poha, *cement pipe and curtains*	Cloth, size-wood
Shivpuri	48.8	*Iron products,* handloom products	Catechu, stone
Murwara	47.5	*Limestone, white bricks*	**Rice,** oilseed
Rewa	45.3	Wooden toys, biri	Oilseed, betel leaves

Table 5-4 *(continued)*

City	*Population growth rate, 1971–81 (percent)*	*Leading commodities manufactured*	*Leading exports*
Growth rate below 45 percent			
Class I			
Gwalior	37.8	Pottery, cloth	Pottery, cloth
Sagar	37.1	Biri, oil	Timber,[b] oil
Bilaspur	36.7	Biri, soap	**Rice,** Tendu leaves
Ujjain	35.2	Cloth	Cloth, Kirana
Burhanpur	34.0	Cloth, biri	Cloth, biri
Ratlam	31.2	Iron scale, brass pots	Cotton, alcohol
Class II			
Chhindwara	40.4	Oil	Vegetables, teakwood
Mandsaur	36.3	Slate pencils, starch	Glutin, slate pencils
Khandwa	34.0	Groundnut oil, pulses	Cotton bales, groundnut oil
Damoh	27.8	Biri, earthenware	Biri, earthenware
Jabalpur Cantonment	21.1	Bakery products, bamboo furniture	Vegetables, bakery products
Mhow Cantonment	18.8	(None listed)	Mawa, rawhide

Notes: Commodities in italics were given a binary variable value of 1, indicating the city has a major intermediate goods, capital goods, or high technology industry; those in boldface received a binary variable value of 1, indicating the city is a major grain processing center.

a. Govindpura, in the Bhopal urban agglomeration, is also a major producer of electrical goods and machines.

b. Timber is, debatably, an intermediate good and was treated as such in alternate regressions (not presented); the effect on regression results is minor.

Source: Government of India (1971). Unfortunately, the census volume does not discuss the criteria used to determine the "most important commodities" exported and manufactured (although the volume does restrict the number given to three or fewer).

Having discussed the variables that are to be used in estimating the population growth equation, we turn to the estimates. The first set of regressions omits a public service term because several observations are lost when the term is included. The independent variables include the city's 1971 ratio of females to males (SEX) and the distance to the nearest Class I city (d). A third term is a binary variable that equals 1 if either of the city's two main industries is an intermediate, capital, or high technology industry, or if the city is a mining town.[14] The fourth term is a binary variable that equals 1 if either of the city's two main exports (to the rest of India) is grain.

Table 5-5. *Madhya Pradesh City Growth Regressions 1–7*

		Coefficient								
Equation	*Class*	*Constant*	DUMMY 1[a]	DUMMY 2[b]	d[c]	d^2	SEX[d]	R^2	F	*Number of observation*
1	I-II-III	154.6	18.40	5.451	0.04499		−136.7	0.36	7.13	55
	t-ratio		3.06	1.07	1.34		−2.08			
2	I-II-III	39.81	27.90	2.0382	−0.09056	0.0005284		0.34	7.48	62
	t-ratio		4.89	0.38	−0.80	1.39				
3	I-II	202.0	10.53	17.12	−0.08163		−175.1	0.70	9.45	21
	t-ratio		2.25	3.26	−1.93		−2.75			
4	I-II	38.59	18.43	18.35	−0.02782			0.57	8.27	23
	t-ratio		3.53	2.78	−0.63					
5	I-II	50.56	18.13	14.53	−0.3311	0.00144		0.62	7.43	23
	t-ratio		3.63	2.16	−1.75	1.64				
6	III	153.9	19.25	2.391	0.08473		−141.6	0.39	4.59	34
	t-ratio		1.76	0.32	1.84		−1.43			
7	III	36.67	33.92	−3.010		0.0002407		0.37	6.83	39
	t-ratio		3.66	−0.41		1.72				

Notes: In all cases, the dependent variable is G_{81}, the 1971–81 city growth rate.

a. DUMMY 1 takes the value 1 if the city has a major intermediate goods, capital goods, or high technology industry, and 0 otherwise.

b. DUMMY 2 takes the value 1 if the city is a major grain processor, and 0 otherwise.

c. d = distance in kilometers to the nearest Class I city.

d. SEX = ratio of females to males in the city.

Source: Authors' calculations.

The correlation between city growth and production structure is striking. Table 5-4 shows the major exports and industries of Class I and II cities in Madhya Pradesh. None of the slower growing cities has a large modern industry, and only one of them is a major grain processing center.

This pattern extends to the Class III category of cities, which includes all the mining towns in the sample. The three fastest growing Class III cities are mining towns, and again the slowest growing cities lack major manufacturing industries. Several of the slowest growing Class III cities, though, are major grain exporters.

The regression results shown in table 5-5 are striking. For the full sample the dummy variables and city distance terms account for roughly 35 percent of the variance in growth rates. The industrial structure binary variable (DUMMY 1) alone explains nearly 29 percent of the variance and is highly significant. By contrast, DUMMY 2 is not highly significant, and d is only at the 90 percent level. The negative sign on SEX indicates that the tendency of females to migrate to cities where SEX is small dominates the child-bearing effect. Unfortunately, however, the sex ratio is correlated with the other independent variables, and this hinders interpretation of the coefficients in equation 1 in the table.[15]

When the sample is limited to Class I and II cities, the model is even more successful. Both dummy variables are highly significant; alone they account for more than 55 percent of the variance in the growth rates of those cities. It is difficult to escape the conclusion that economic structure is a major determinant of a city's growth. Large cities are not subject to mysterious, runaway growth; even with highly imperfect proxy variables it is possible to predict growth with considerable accuracy.

Class III cities appear to be influenced more by their distance from large cities and less by their role as agricultural processors. The model is nevertheless able to account for 37 percent of their growth variance with only two significant variables, DUMMY 1 and d^2. These results suggest that distance from other large urban areas is more important to small cities than large ones. In equation 7 in table 5-5 an increase in d^2 by one standard deviation implies an increase in city growth of 5.9 percent.

Even without tests for verification, the results in table 5-5 indicate that the determinants of city growth differ among categories of cities. The industrial structure dummy remains consistently important, but the other variables are influential only for one of the two subsamples. The coefficient signs, however, consistently accord with the predictions of the model discussed in the previous section, while the R^2's are surprisingly large.

Efforts to estimate service provision equations enjoyed limited success. From equations 8 through 11 in table 5-6 it can be seen that the independent

Table 5-6. *Madhya Pradesh City Service Regressions 8–11*

			Coefficient						
Equation	*Class*	*Dependent variable*	*Constant*	TRIBE	SEX	CASTE	R^2	F	*Number of observations*
8	I-II-III	EXP	−53.153	−0.06561	90.439		0.24	8.10	55
	t-ratio			−1.88	2.96				
9	I-II	EXP	−70.399	−0.204	114.21		0.36	4.78	21
	t-ratio			−1.49	2.02				
10	III	EXP	−42.751	−0.0407	76.03		0.20	3.64	34
	t-ratio			−1.09	2.12				
11	III	ELEC	85.08	−0.1332	−69.07	−0.08319	0.41	6.95	34
	t-ratio			−4.45	−2.42	−2.72			

Notes: TRIBE = percentage of population belonging to a scheduled tribe.
CASTE = percentage of population belonging to a scheduled caste.
SEX = ratio of females to males in the city.
EXP = municipal expenditures per capita.
ELEC = commercial and industrial electric hookups, per capita.

Source: Data for TRIBE, SEX, CASTE, EXP, and ELEC are from Government of India (1971).

variables explain 20 to 36 percent of the variance in per capita expenditures (EXP) and about 40 percent of the variance in electricity hookup provision (ELEC) for Class III cities (a similar equation for Class I and II cities had little explanatory power). Surprisingly, the industrial structure variables had virtually no impact on EXP or ELEC, and the percentage of the population belonging to a scheduled caste (CASTE) failed to influence EXP. Previous growth experience (G_{71}) was highly negatively correlated with SEX, but in regressions not shown it was less successful in explaining variation in per capita services. In the equations presented, the coefficients of the percentages of the population belonging to scheduled tribes and castes have the expected negative signs: the larger the proportion of the population in disadvantaged groups, the lower the per capita service provision. The ratio of females to males, however, gives a predominantly significant, positive sign in the EXP regressions. This sign may reflect the higher sex ratios in older, more slowly growing urban areas which experience less "catching-up" in service provision. In addition, demand for public services such as education is likely to rise with the dependency ratio, while demand for site amenities such as water taps and electricity can be expected to rise with the importance of multiperson households. Both of these variables will be highly correlated with the sex ratio as well.

Residuals from the equations for public service provision are themselves used as explanatory variables in the city growth equations in table 5-7. These residuals provide an indicator of effort in the provision of services after correcting for interurban variation in demographic composition. In the growth equations the residuals from the EXP regressions nearly double the explanatory power of the Class I-II-III and Class III regressions. Their contribution to the Class I-II regressions is more modest, but positive. However, it is difficult to accept that anticipated service provision is the major determinant of variation in the growth rate of Class III cities. This skepticism is reinforced when we consider the impact of the residuals from ELEC on Class III growth: the growth equation exhibits little additional explanatory power, and the coefficient of RESID has an unexpected negative sign.

If we are reluctant to accept per capita expenditures as the major determinant of city growth, we must be prepared to explain its significance as being attributable to its correlation with an important omitted variable. One likely correlation is between EXP and the presence of government administrative bodies. Cursory examination of the data indicates that rapidly growing cities tend to have a large government presence, although any attempt to establish such a link would require a detailed examination that is beyond the scope of this study.[16] We believe that such a connection is quite likely, however, and that the presence of a tie between government

Table 5-7. *Madhya Pradesh City Growth Regressions 12–16*

		Coefficient									*Number of*
Equation	*Class*	*Constant*	DUMMY 1	DUMMY 2	*d*	d^2	SEX	RESID[a]	R^2	*F*	*observations*
12	I-II-III	44.49	14.49	4.20	−0.1640	0.0007313		0.0965	0.74	27.88	55
	t-ratio		3.98	1.28	−2.17	2.93		9.6			
13	I-II	58.42	11.49	12.43	−0.4661	0.002175		0.0442	0.88	21.09	21
	t-ratio		3.85	3.09	−3.79	3.88		4.4			
14	III	31.97	12.68	0.9834	0.06045			0.1156	0.78	25.36	34
	t-ratio		2.13	0.23	2.16			5.8			
15	III	160.4		0.06005			−143.6	0.1177	0.81	42.83	34
	t-ratio			2.39			−3.29	11.8			
16	III	28.85	27.30	0.08948				−0.9516	0.40	6.54	34
	t-ratio		3.07	1.99				−1.60			

Notes: See notes to table 5-5.

a. The variable RESID refers to the residuals (actual less estimated dependent variable values) in table 5-6. The RESID term in equation 12 is generated from equation 8; the RESID terms in equations 13, 14, and 15 are generated from equation 10; and the RESID term in equation 16 is generated from equation 11.

Source: Authors' calculations.

presence and rapid city growth is plausible. According to World Bank data, employment growth has proceeded at a far greater pace in the public sector than in the private organized sector.

An alternative explanation for the highly significant coefficients on the EXP residuals involves reverse causality. Cities anticipating rapid future growth of industry and population attempt to provide an infrastructural base for them. While there may be a lag in the equilibrium level of *services* per capita, *expenditures* per capita may be abnormally high during a period of rapid accumulation of infrastructural capital.

The need for further modeling is apparent from the results we have obtained. Expenditures, population growth, and the demographic structure are all likely to be simultaneously determined. Single-equation estimates suffer either from the bias of simultaneous equations or from omitted variables. Fortunately, Indian data may not render the task of simultaneous estimation impossible. In the meantime, we may be confident that the major economic and geographic variables—DUMMY 1, DUMMY 2, and d—are important determinants of city growth experience.

Appendix A: The Mathematics of a Model of City Growth

Let us start with the utility maximization framework (5.8) and the division of the labor force and production into four sectors (5.9). Our assumptions on input use into value added in manufacturing (VA_i^M) permit us to write

$$(5.16) \qquad VA_i^M = VA_i^M(P_i^M, L_i^M, K_i^M, S_i) = P_i^M \cdot f^M(L_i^M, K_i^M, S_i).$$

If we assume competitive labor and capital factor markets, the first-order efficiency conditions are

$$(5.17) \qquad \frac{\partial VA_i^M}{\partial L_i^M} = w$$

and

$$(5.18) \qquad \frac{\partial VA_i^M}{\partial K_i^M} = r$$

where

w = labor's wage
r = capital's rental rate.

It is debatable whether we should view India as having efficient capital markets. If we do, r is constant in equation (5.18), and equations (5.17) and (5.18) together with a zero-profit restriction may be solved to give, in view of equation (5.16), labor demand as a function of P_i^M, S_i, r, and equilibrium wage w_i^{M*}. We believe it is more realistic to view r as endogenous to this system and to treat K_i^M as a function of exogenous forces (such as the investment decisions of India's public sector and the accumulation rate of city i). This yields

$$L_i^{M*} = L_i^{M*}(P_i^M, S_i, K_i^M, w_i^{M*}). \tag{5.17'}$$

If we proceed with the assumption of a less than perfect capital market, the equilibrium wage may be written as a function of P_i^M, S_i, K_i^M, and L_i^M—that is, labor demand (5.17') involves only two endogenous variables and does not require further substitution from the capital market equilibrium condition.

To find the equilibrium labor supply L_i^M, we further assume that a worker's quality of life may be written as a function of his or her consumption both of private goods and of publicly provided goods Z_i. Substituting in the wage constraint, the expected quality of life for a worker in city i's M sector is

$$U_i^M = U(w_i^M, Z_i). \tag{5.10}$$

Labor market and wage equilibria then imply

$$U_i^M = \bar{U}_A \tag{5.19}$$

or

$$U[P_i^M \cdot f_1^M(L_i^M, K_i^M, S_i), Z_i] = \bar{U}_A. \tag{5.20}$$

Equation (5.20) then yields equilibrium employment in the manufacturing sector of city i. More important, it also implies

(5.21)
$$dL_i^M = \left(-\frac{f_{13}^M}{f_{11}^M}\right) dS_i + \left(-\frac{f_{12}^M}{f_{11}^M}\right) dK_i^M + \left(-\frac{f_1^M}{f_{11}^M}\right)\frac{dP_i^M}{P_i^M} + \left(\frac{-U_2}{U_1 P_i^M f_{11}^M}\right) dZ_i.$$

If we make the usual assumptions on the production and utility functions ($f_{11}^M < 0 < f_{12}^M$, f_{13}^M and U_1, $U_2 > 0$), the increment in the equilibrium manufacturing labor force dL_i^M is an increasing function of dS_i, dK_i^M, dP_i^M/P_i^M, and dZ_i.

A similar procedure can be used to derive employment growth in the agricultural processing sector AP_i. The only effective difference is that we recognize the impact of the agricultural surplus supply curve $A[L^A(c_i),$

c_i], which contains a regional size and high-yield variety adoption shift variable. Output is thus

(5.22)
$$f_i^{AP} = f^{AP}(L_i^{AP}, K_i^{AP}, S_i, A_i) = f^{AP}\{L_i^{AP}, K_i^{AP}, S_i, A[L^A(c_i), c_i]\}.$$

The corresponding equilibrium employment growth equation is:

(5.23)
$$\begin{aligned} dL_i^{AP} = &\left(-\frac{f_{13}^{AP}}{f_{11}^{AP}}\right) dS_i + \left(-\frac{f_{12}^{AP}}{f_{11}^{AP}}\right) dK_i^{AP} + \left(\frac{-f_1^{AP}}{f_{11}^{AP}}\right) \frac{dP_i^{AP}}{P_i^{AP}} \\ &+ \left(\frac{-U_2}{U_1 P_i^{AP} f_{11}^{AP}}\right) dZ_i + \left(-\frac{f_{14}^{AP}}{f_{11}^{AP}}\right) \left(\frac{\partial A_i}{\partial L_i^A} \cdot \frac{\partial L_i^A}{\partial c_i} + \frac{\partial A_i}{\partial c_i}\right) dc_i \end{aligned}$$

where once again all terms in parentheses are positive under usual convexity and concavity conditions.[17]

Turning now to the service sector, recall equation (5.11):

$$P_i^{SV} = P_i^{SV}\left[\bar{U}_A(L_i^M + L_i^{AP}), Z_i, S_i, Q_i^{SV}\right].$$

The assumptions underlying equation (5.11) are not overly restrictive.[18] The nontradability assumption is standard. The *SV* sector produces local social infrastructure and commercial and personal services. The prices of these services are competitive among cities only if the *M* and *AP* sectors are sufficiently mobile to move in response to price differentials. Such mobility seems unlikely, particularly for the *AP* sector, which is tied to its primary product inputs.

The equations could also be expanded to handle other general equilibrium effects concerning the *SV* sector that implicitly have been ignored. If, in addition to that portion of *SV* devoted to infrastructure provision, *SV* output is an input into *M* and *AP* products, then labor force growth conditions (5.21) and (5.23) will vary inversely with dP_i^{SV}. This addition would make the sectoral employment sign somewhat indeterminate, but since P_i^{SV} will itself vary inversely with L_i^M and L_i^{AP}, the main effect is to reduce the absolute value of the sign in these equations. Similarly, we have abstracted from the effects of price changes on utility levels. Strictly speaking, the income term in equation (5.8) should be deflated by a consumer price index; its absence is equivalent to positing the presence of a degree of money illusion. We expect that such general equilibrium effects will have only minor implications.

Total differentiation of the *SV*-sector labor market equilibrium condition yields

(5.24)

$$dL_i^{SV} = \left[\frac{f_{13}^{SV} + \left(\dfrac{f_1^{SV} f_3^{SV}}{P_i^{SV}} \cdot \dfrac{\partial P_i^{SV}}{\partial Q_i^{SV}}\right)}{F_{11}^{SV}}\right] dS_i + \left[\frac{f_{12}^{SV} + \left(\dfrac{f_1^{SV} f_2^{SV}}{P_i^{SV}} \cdot \dfrac{\partial P_i^{SV}}{\partial Q_i^{SV}}\right)}{-F_{11}^{SV}}\right] dK_i^{SV}$$

$$+ \left(\frac{-U_2}{U_1 P_i^{SV} F_{11}^{SV}}\right) dZ_i + \left(\frac{-f_1^{SV}}{F_{11}^{SV} P_i^{SV}}\right)$$

$$\cdot \left[\bar{U}_A \cdot \frac{\partial P_i^{SV}}{\partial L} \cdot dL_i^M + \bar{U}_A \cdot \frac{\partial P_i^{SV}}{\partial L} \cdot dL_i^{AP} + \frac{\partial P_i^{SV}}{\partial Z_i} \cdot dZ_i + \frac{\partial P_i^{SV}}{\partial S_i}\right] dS_i$$

where $F_{11}^{SV} = f_{11}^{SV} + \dfrac{f_1^2}{P_i^{SV}} \cdot \dfrac{\partial P_i^{SV}}{\partial Q_i^{SV}} < 0$. This reduces to

(5.25)

$$dL_i^{SV} = \left\{\left[\frac{f_{13}^{SV} + \left(\dfrac{f_1^{SV} f_3^{SV}}{P_i^{SV}} \cdot \dfrac{\partial P_i^{SV}}{\partial Q_i^{SV}}\right)}{F_{11}^{SV}}\right] - \frac{f_i^{SV}}{F_{11}^{SV} P_i^{SV}} \cdot \frac{\partial P_i^{SV}}{\partial S_i}\right.$$

$$\left. - \left[\left(\frac{f_1^{SV} \cdot \bar{U}_A}{F_{11}^{SV} \cdot P_i^{SV}}\right)\left(\frac{-\partial P_i^{SV}}{\partial L} \cdot \frac{f_{13}^M}{f_{11}^M} - \frac{\partial P_i^{SV}}{\partial L} \cdot \frac{f_{13}^{AP}}{f_{11}^{AP}}\right)\right]\right\} dS_i$$

$$+ \left(\frac{f_1^{SV} \bar{U}_A}{F_{11}^{SV} P_i^{SV}} \cdot \frac{\partial P_i^{SV}}{\partial L} \cdot \frac{f_{12}^{AP}}{f_{11}^{AP}}\right) dK_i^{AP} + \left(\frac{f_1^{SV} \bar{U}_A}{F_{11}^{SV} P_i^{SV}} \cdot \frac{\partial P_i^{SV}}{\partial L} \cdot \frac{f_1^M}{f_{11}^M P_i^M}\right) dP_i^M$$

$$+ \left(\frac{f_1^{SV} \bar{U}_A}{F_{11}^{SV} P_i^{SV}} \cdot \frac{\partial P_i^{SV}}{\partial L} \cdot \frac{f_1^{AP}}{f_{11}^{AP} P_i^{AP}}\right) dP_i^{AP}$$

$$+ \left\{\left[\frac{f^{SV} \bar{U}_A}{F_{11}^{SV} P_i^{SV}} \cdot \frac{\partial P_i^{SV}}{\partial L} \cdot \frac{f_{14}^{AP}}{f_{11}^{AP}}\right] \cdot \left[\left(\frac{\partial A_i}{\partial L_i^A} \cdot \frac{\partial L_i^A}{\partial c_i}\right) + \frac{\partial A_i}{\partial c_i}\right]\right\} dc_i$$

$$+ \left(\frac{-U_2}{U_1 P_i^{SV} F_{11}^{SV}}\right) \cdot \left(1 - \frac{f_1^{SV}}{P_1^M f_{11}^M} - \frac{f_1^{SV}}{P_i^{AP} f_{11}^{AP}}\right) dZ_i.$$

The assumptions on L_i^{SWS} enable dL_i^{SWS} to be expressed in a manner analogous to equation (5.25).

Our incorporation of secondary adults and children is given by equation (5.14). Total population growth in a given city is simply the sum of employment growth in the four sectors plus the increases in nonworkers. The final step is to move from growth to growth rates. This last transformation is made to ensure homoscedastic error terms in the equations estimated, but it requires a slight reinterpretation of the equations. We circumvent the potential problems by expressing independent variables as relative to the population stock (for example, $\mathrm{d}K_i^M/N_i$ instead of $\mathrm{d}K_i^M$).

Noting that equation (5.14) can be rewritten as

(5.26)

$$\mathrm{d}N_i^{SA} = \sigma\,(\mathrm{d}L_i^{AP} + \mathrm{d}L_i^M + \mathrm{d}L_i^{SV}) + \sigma\Theta N_i - (1+\sigma)\Theta(L_i^{SWS} + N_i^{SA})$$

we then have

(5.27)

$$\frac{\mathrm{d}N_i}{N_i} = \sigma\Theta + \sum_{j=AP,M,SV,SWS} e_j \left(\frac{\mathrm{d}K_i^j}{N_i}\right) + e_1\left(\frac{\mathrm{d}S_i}{N_i}\right) + e_2\left(\frac{\mathrm{d}Z_i}{N_i}\right)$$
$$+ \sum_{j=AP,M} g_j \left(\frac{\mathrm{d}P_i^j}{P_i^j N_i}\right) + [b - (1+\sigma)\Theta]\left(\frac{N_i^{SA} + L_i^{SWS}}{N_i}\right) + e_3\left(\frac{\mathrm{d}c_i}{N_i}\right)$$

where the e_j, e_1, e_2, e_3, and g_j terms denote constants.

The linearity of equation (5.27) obscures the fact that the coefficients e_1, e_2, and e_3 involve complex nonlinear terms that are likely to vary among cities. In the absence of scale economies, however, we have no reason to expect these terms to vary systematically, and we therefore assume that the coefficients can be treated as constants, at least to a first approximation. As was discussed in the final section of this chapter, scale economies are most likely to occur in the terms with coefficients e_1 and e_2 which relate to infrastructural provision.

Appendix B: Correlation Matrices

Correlation Matrix for Table 5-5

	DUMMY 2	d	d^2	SEX
Equation 1				
DUMMY 1	−0.14	0.06	—	−0.47
DUMMY 2		−0.07	—	0.26
d			—	−0.22
Equation 2				
DUMMY 1	−0.11	0.08	0.07	—
DUMMY 2		−0.01	0.02	—
d			0.95	—
Equation 3				
DUMMY 1	0.12	−0.04	—	−0.39
DUMMY 2		−0.13	—	0.03
d			—	−0.37
Equations 4 and 5				
DUMMY 1	0.15	−0.05	−0.02	—
DUMMY 2		0.13	0.04	—
d			0.97	—

Correlation Matrix for Table 5-5 (*continued*)

	DUMMY 2	d	d^2	SEX
Equation 6				
DUMMY 1	−0.23	0.23	—	−0.57
DUMMY 2		−0.15	—	−0.39
d			—	−0.16
Equation 7				
DUMMY 1	−0.21	0.23	0.18	—
DUMMY 2		−0.03	−0.02	—
d			0.96	—

Correlation Matrix for Table 5-6

	TRIBE	G_{71}	CASTE
Equation 8			
SEX	−0.26	−0.64	—
TRIBE		0.33	—
Equation 9			
SEX	−0.35	−0.69	—
TRIBE		0.21	—
Equations 10 and 11			
SEX	−0.25	−0.64	−0.31
TRIBE		0.41	−0.30

Correlation Matrix for Table 5-7

	DUMMY 2	SEX	d	d^2	RESID	
Equation 12						
DUMMY 1	−0.14	−0.47	0.06	0.06	0.33	
DUMMY 2		0.26	−0.07	−0.02	−0.14	
SEX			−0.22	−0.23	−0.31	
d				0.96	0.07	
d^2					0.09	
Equation 13						
DUMMY 1	0.12	−0.39	−0.03	0.00	0.28	
DUMMY 2		0.03	−0.13	−0.02	−0.09	
SEX			−0.37	−0.42	−0.49	
d				0.97	−0.26	
d^2					−0.19	
					RESID (eqs. 14, 15)	RESID (eq. 16)
Equations 14, 15, and 16						
DUMMY 1	−0.23	−0.57	0.23	0.19	0.38	−0.02
DUMMY 2		0.39	−0.15	−0.11	−0.15	−0.03
SEX			−0.16	−0.20	−0.23	−0.01
d				0.97	0.20	0.04
d^2					0.17	−0.02

References

Crook, Nigel, and Tim Dyson. 1982. "Urbanization in India: Results of the 1981 Census." *Population and Development Review* 8, no. 1 (March): 145–55.

Government of India. 1971. *Census of India, 1971.* Series 10, Madhya Pradesh. Part VI-A: Town Directory. New Delhi: Census Commissioner of India.

———. 1981. *Census of India, 1981.* Series 1, papers 2 and 3, Provisional Population Tables. New Delhi: Census Commissioner of India.

Mohan, Rakesh, and Chandrashekhar Pant. 1982. "The Morphology of Urbanization in India: Some Results from the 1981 Census." *Economic and Political Weekly* 16 (August).

Prakesh, Shiri, and P. Rejan. 1977. "Economic Development and Structural Changes in the Economy of Madhya Pradesh." *Indian Journal of Economics* 58, no. 229 (October): 203–37.

Rao, V. K. R. V. 1983. *India's National Income, 1950–1980.* New Delhi: Sage Publications.

Tiebout, Charles. 1956. "A Pure Theory of Local Public Expenditure." *Journal of Political Economy* 64:416–24.

Notes

1. For a more detailed analysis of the Madhya Pradesh economy, see Prakesh and Rejan (1977).

2. The ranking is by 1981 population. The twenty largest urban areas include the Madhya Pradesh cities of Indore (seventeenth) and Jabalpur (twentieth). See Crook and Dyson (1982). Because Crook and Dyson's figures are based on entire "urban agglomerations," whereas we define specific cities as "municipal corporations plus outgrowths," the sets of figures differ slightly.

3. The correlation coefficient between $P_{i\,61}$ and d_i is negligible:

$$\rho(P_{i\,61}, d_i) = -0.03.$$

4. Hereafter, "statistical significance" will be taken to refer to a 90 percent significance level.

5. The relative numbers of Class II cities have also enjoyed a secular rise. See Government of India (1971). This census report is also an example of a publication that reports population growth rates by terminal- rather than base-year class.

6. For the Class I, II, and III cities together, $\rho(G_{i\,71}, P_{i\,71}) = 0.04$—a surprisingly low correlation—and $\rho(G_{i\,71}, d_i) = 0.26$.

7. Other functional forms were tried, but they did not improve on the linear fit. The $G_{i\,71}$ term remained highly significant; the d and d^2 variables, significantly different from zero at the 90 and 80 percent levels, respectively, in regression (5.5), were not as robust. For this sample, $\rho(G_{i\,71}, d) = 0.05$ and $\rho(G_{i\,71}, d_i^2) = 0.09$.

8. A less restrictive assumption is that P^M varies with the levels of external demand-related variables, as well as with local purchasing power. M-sector producers may also be permitted to behave as monopolists and to choose an optimal price. None of these extensions seriously affects the mathematics of the model (at least as long as M goods are not highly inferior in the eyes of the local population), and all of them are excluded from the analysis for the sake of clarity.

9. The mathematics of the model appears in appendix A of this chapter. The labor demand function here can be viewed (by Sheppard's lemma) as the negative of the partial derivative of a city's manufacturing-sector profit function with respect to the wage rate. Investment flows respond to regional differences in rates of return to an unknown extent, but the capital stock depends on investments during a long historical period and responds only very slowly to regional differences.

10. Note that f_i^{AP} is strictly a value added production function that is net of nonagricultural inputs but that includes A_i.

11. The total differential dL_i^{SWS} takes the same form as dL_i^{SV} in equation (5.25) in appendix A.

12. Within the factory sector, capital growth (1960–77) has been most rapid in electric and nonelectric machinery, motor vehicles, chemicals and chemical products, nonferrous basic metals, shipbuilding, and, especially, electric light and power. Sectors with slow accumulation include tobacco products, spinning and weaving, miscellaneous food products, petroleum refineries, railroad equipment, repairs to motor vehicles, and cement. Thus, the investment pattern noted is strong but is not without exception. Unregistered manufacturing's share of gross domestic capital formation also rose markedly between the early 1960s and 1970s. See Rao (1983, pp. 157, 160).

13. In equation (5.27) the impact of a dynamic rural sector is captured by high values for the shift variable c_i and by the AP sector's capital stock growth dK_i^{AP}.

14. Data on cities' industrial structure are taken from Government of India (1971). Unfortunately, the data are insufficiently detailed to permit us to expand beyond using binary variables.

15. Correlation coefficients for the regressions presented in tables 5-5, 5-6, and 5-7 appear in appendix B of this chapter.

16. Of the ten fastest growing Class III cities, five were district headquarters and two were project towns (and one other was a coal mining center). See Government of India (1971).

17. The last term in equation (5.23) is sensitive to the assumptions regarding the underlying product and labor market structures. As long as U_A is fixed and $\partial A_i/\partial c_i$ is positive, a rise in c_i must have a positive impact on the equilibrium value of L_i^{AP}. If the labor market is viewed as "regional" and U_A no longer fixed, while $L_i^{AP} = L_i - L_i^M - L_i^{SV} - L_i^A$, the above determinacy would disappear. If rural wage formation is nonneoclassical, the sign of $\partial A_i/\partial c_i$ may also be questionable.

18. Expressing P_i^{SV} as a function of SV labor stock as well has only the minor effect of changing the F_{11}^{SV} terms in equations (5.24) and (5.25) to

$$F_{11}^{SV} + \bar{U}_A \cdot \frac{f_1^{SV}}{P_i^{SV}} \cdot \frac{\partial P_i^{SV}}{\partial L}.$$

Only if the urban economy is highly unstable will the latter term in this expression be large enough to change the comparative static results.

6

Urbanization and Productivity in Indian States

Satyendra Verma

THE STATES AND UNION TERRITORIES of India differ not only in their social customs, language, and literacy rates but also in their economic background—infrastructure, natural resources, public services, and utilities. From 1950–51 to 1979–80, they experienced different growth rates of such measures of economic development as labor force size, value added in the manufacturing sector, and per capita domestic product. The states and union territories also vary greatly in the extent to which they have become urbanized. Those in the West and North are the most urban. According to the 1971 census 30 percent of the total population of those regions is urban; the figure is 23 percent in the South and Central regions, and only 14 percent in the East. Data from the 1981 census indicate that among the states the percentage of the population that is urban varies from 7.7 percent in Himachal Pradesh to 35 percent in Maharashtra; among the union territories percent urban varies from 6.3 percent in Arunachal Pradesh to 93.6 percent in Chandigarh.

This chapter examines the significance of urbanization with regard to the temporal and spatial differences in factor and total productivity of the manufacturing sector of Indian states for 1961, 1971, and 1975. First an overview of the literature on the relation between productivity and city size is presented, followed by a description of the scope and hypothesis of the present study. The chapter then addresses the methodology of calculating translog multilateral productivity indices. Subsequent sections de-

Note: I am indebted to Edwin Mills, Koichi Mera, Mieko Nishimizu, and Rakesh Mohan for the valuable guidance and comments they provided in the preparation of this paper.

scribe the sources of data and the results of productivity indices, relate productivity to migration, and present the productivity index for the agricultural sector. The conclusions of the study are summarized in the final section.

India is a vast country, divided into twenty-two states and nine union territories. The composition of some of the states and union territories has changed since 1961, however. For example, Himachal Pradesh, Manipur, and Tripura, which were classified as union territories in the 1961 census, later became states and were classified as such in the 1971 census. Other states, such as Punjab and Haryana, were reorganized. Madras was reorganized as Tamil Nadu, and Mysore as Karnataka.

The present study includes seventeen states (Andhra Pradesh, Assam, Bihar, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, and West Bengal) and one union territory (Delhi). The other states and union territories are either negligibly small or incomparable because a consistent series of data is not available for them as a result of reorganization since 1961.

The basic definition of an urban area is a set of contiguous places where population densities are much greater than elsewhere. With economic development, incomes rise and the demand for manufactured articles rises. The resulting industrial expansion attracts people from rural to urban areas, and these migrants further increase the demand for manufactured goods. In such a growing urban area, therefore, not only does horizontal and vertical integration of industries proceed simultaneously, but city services such as banking and other financial institutions, managerial firms, educational institutions, and transport facilities develop rapidly. These things provide scale and transport economies—also called agglomeration economies—to the manufacturing sector and thereby affect its productivity. Since the modern manufacturing and service sectors are becoming more complex and capital-intensive, only large cities can provide a market that is large enough to support such activities. In addition, high transportation costs motivate consumers and producers of commodities and services to locate near each other.

Alonso (1968), Mera (1973), Aberg (1973), Sveikauskas (1975), Segal (1976), and, more recently, Wheaton and Shishido (1981) have shown that productivity is higher where urban agglomeration economies are more prominent. For example, Mera has shown that productivity is higher in the more densely populated areas. Similarly, Aberg found that in Sweden productivity of the manufacturing sector is higher in the more populous regions. Segal estimated the net benefits of agglomeration in production in large metropolitan areas, and Sveikauskas calculated the productivity of labor in cities and found that a doubling of city size is typically associated with a 6 percent increase in labor productivity. Wheaton and Shishido attempted to explain cross-national variations in urban concentration that result in per capita income differentials. Confirming these findings, the large cities of India, such as Bombay, Calcutta, Madras, and Delhi, have shown high growth rates of population and industrial output in the past three decades.

How urban agglomeration economies have influenced the productivity of the manufacturing sector in Indian states is the subject of this chapter. Since data on capital were not available by city, I have examined the effect of urban agglomeration on the multilateral productivity index of states. States with large geographical areas but with small urban populations—Andhra Pradesh, Rajasthan, and Uttar Pradesh, for example—have low productivity. Small states with large urban populations have high productivity. High productivity also appears to be positively related to gross migration into urban areas.

The present study includes in the manufacturing sector all factories registered under the Factory Act of 1948, excluding defense installations, factories engaged in the storage and distribution of oil, and technical training institutions. An urban area is defined as a place with municipal corporations of cantonments or a notified town area,[1] a population of at least 5,000, at least 75 percent of its male workers in nonagricultural work, and a density of at least 400 people a square kilometer. An urban agglomeration is defined as an area or a city with more than 100,000 inhabitants; the core of an urban agglomeration is usually a Class I city.

Measurement of Productivity Differentials

Since Solow (1957), productivity analysts have become increasingly aware of the close relationship between productivity accounting by index numbers and the form of the cost functions. Jorgenson and Griliches (1967) used the Tornqvist approximation to the Divisia index of productivity growth for a particular spatial entity. Diewert (1976) also made an important contribution to the literature on productivity measurement: he introduced superlative index numbers which use quadratic functional forms, and he showed that for intertemporal comparisons the Tornqvist approximation to the Divisia index of inputs is exact for the linear homogeneous translog production function. Until recently the intertemporal growth rate, not the productivity level, was the primary concern. Jorgenson and Nishimizu (1978) introduced a framework for investigating relative productivity levels. Both the intertemporal productivity growth rate index introduced by Jorgenson and Griliches (1967) and Jorgenson and Nishimizu's productivity level comparison index are contained in Denny, Fuss, and May (1981). Using a quadratic approximation to the function

$$\log Q_{it} = F(\log X_{it}, \log T_{it}, D) \tag{6.1}$$

where

i = region
t = time
Q = output vector
X = input vector
T = intertemporal total factor productivity growth index
D = regional dummy,

Denny, Fuss, and May defined the intertemporal total factor productivity growth index as

(6.2)
$$T_{t,t+1} = (\log Q_{t+1} - \log Q_t) - \tfrac{1}{2} \sum_k (S_{k,t+1} + S_{kt})(\log X_{k,t+1} - \log X_{kt})$$

and interspatial productivity differences δ as

$$\delta_{i0} = (\log Q_i - \log Q_0) - \tfrac{1}{2} \sum_k (S_{k,i} + S_{k,0})(\log X_{k,i} - \log X_{k,0}) \tag{6.3}$$

where

k = k^{th} input
S_k = share of K^{th} input in production cost
i and 0 = two regions to be compared.

All the indices defined above were derived from an unrestricted constant returns-to-scale translog transformation function. Thus, these indices entail neither separability of inputs and outputs, nor neutrality of differences in productivity. The indices for interspatial comparisons are bilateral and are good for only binary (input or output) comparisons. Introduction of a third country, state, or region would not satisfy the circularity test, namely $\delta_{io} \neq \delta_{ij}/\delta_{oj}$, where i, j, and o are states.

To avoid this problem, Caves, Christensen, and Diewert (1982) defined an interspatial multilateral translog productivity index. That index is also used in the present study, and the technique can be described, in short, as follows.

In Caves, Christensen, and Diewert's notation, for an economic entry S, using a vector of inputs X^s to produce output Y^s, the structure of production is

$$F(\ln Y^s, \ln X^s, S) = 1. \tag{6.4}$$

Here, region s is included as an argument of F to indicate that the structure of production, and hence productivity, is allowed to differ in a nonneutral manner among observations. Observations u and v are assumed to fall on the translog transformation (6.4). A comparison of output (or input) for states u and v could use state u or v as a base. If u is used as a base, the interspatial ratio of outputs of u to v is defined as the maximum proportional increase in all elements of y^u such that the resulting output vector can be produced with the input and productivity levels of v. If v is used as the basis of comparison, the output of u relative to v is defined as the maximum proportional increase in all elements of Y^v such that the resulting vector of outputs can be produced with the input and productivity levels of u. The geometric mean of these two ratios thus defines the interspatial translog productivity index as

(6.5) $$\ln \delta_{uv} = \tfrac{1}{2} \sum_i (R_i^u + R_i^v) \ln \left(\frac{Y_i^u}{Y_i^v}\right)$$

where R_i's are output shares in total revenue. This index, as was pointed out earlier, is bilateral. To modify the definition of output comparisons so that transitive results are obtained in the multilateral setting, Caves, Christensen, and Diewert defined the output of region u relative to the output of all s regions as the geometric mean of the bilateral output comparisons between u and each of the regions. Thus:

(6.6) $$\overline{\ln \delta_u} = \frac{1}{S} \sum_s^S \ln \delta_{us}.$$

Substituting in the translog bilateral output index:

(6.7) $$\overline{\ln \delta_u} = \tfrac{1}{2} \sum_i \left[(R_i^u + \bar{R}_i)(\ln Y_i^u) - (\overline{\ln Y_i}) + R_i \ln Y_i^u - \overline{R_i \ln Y_i} \right].$$

An alternative way of defining such a translog multilateral output index is by considering a representative state h with output vector $\overline{\ln Y}$, input vector $\overline{\ln X}$, and revenue share $\bar{R}_i$ (the bars indicate averages). Then the translog productivity index can be written

(6.8) $$\ln \delta_{uh} = \tfrac{1}{2} \sum_i (R_i^u + R_i)(\ln Y_i^u - \ln Y_i).$$

In the same way, a translog factor productivity index ρ_{uh} for state u as compared with the hypothetical state h can be given by

(6.9) $$\ln \rho_{uh} = \tfrac{1}{2} \sum_j (W_j^u - \bar{W}_j)(\ln X_j^u - \overline{\ln X_j})$$

Table 6-1. *Interspatial Productivity of the Manufacturing Sector,* ASI *Group, 1961*

State	*Millions of man-hours worked*	*Total productive capital (millions of rupees)*	*Share of labor in total cost (percent)*	*Value added (millions of rupees)*	*Interspatial productivity index*: *Factor productivity* (ρ)	*Interspatial productivity index*: *Total productivity* (λ)
Andhra Pradesh	213	858	13.5	276	0.3933	−0.4206
Assam	1,249	658	12.3	273	0.4118	−0.4500
Bihar	340	2,719	14.5	568	1.4444	−0.7500
Gujarat	604	1,884	20.1	1,007	1.0359	0.2311
Haryana	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Himachal Pradesh	3	15	11.1	14	−3.6860	0.6774
Jammu and Kashmir	16	36	22.5	10	−2.6521	−0.6930
Karnataka	221	861	18.3	341	0.4101	−0.2265
Kerala	272	467	16.7	239	−0.0713	−0.0999
Madhya Pradesh	205	712	15.9	285	0.2340	−0.2292
Maharashtra	1,376	537	17.9	2,664	0.3249	1.9150
Orissa	50	581	11.9	121	0.0728	−0.9247
Punjab	155	740	11.3	242	0.2189	−0.3777
Rajasthan	87	254	19.4	105	−0.6753	−0.3185
Tamil Nadu	479	1,455	16.6	806	0.9278	0.1166
Uttar Pradesh	516	1,676	13.5	929	1.0944	0.0920
West Bengal	1,379	5,194	16.8	2,043	2.2100	−0.0555
Delhi (union territory)	87	257	16.0	161	−0.7143	0.1814

n.a. Not available.

Source: Central Statistical Organization (1978b).

where W_j's are the shares of inputs in the total cost, and the bars indicate averages over s states.

If economic entities minimize cost conditional on output levels and input prices, and Diewert's (1976) quadratic index is applied, the total productivity index $\ln \lambda_{uh}$ becomes

$$\text{(6.10)} \qquad \ln \lambda_{uh} = \tfrac{1}{2} \sum_i (R_i + \bar{R}_i)(\ln Y_i^u - \overline{\ln Y_i}) - \tfrac{1}{2} \sum_j (W_j^u + \bar{W}_j)(\ln X_j^u - \overline{\ln X_j}).$$

The index $\ln \lambda_{uh}$ is known as the interspatial translog total productivity index. Since states may differ in their level of technology and elasticities of substitution, the advantage of using such a translog multilateral productivity index is that the elasticity of substitution is no longer assumed to be constant across the states.

Data and Productivity Estimates

In this section both the interspatial factor productivity index ρ and total productivity index λ are compared for seventeen states and one union territory. As stated before, the other states and union territories are either negligibly small or incomparable. For example, Manipur, Nagaland, Sikkim, Tripura, and Meghalaya were new states added in the 1981 census. Of these, Nagaland, Sikkim, and Tripura are very small states. The productivity index is calculated for the entire manufacturing sector by the Annual Survey of Industries (ASI). The ASI—which covers the whole of the Indian Union, including Jammu and Kashmir—defines a factory as an installation that uses electric power and has twenty or more workers. Therefore, the coverage is nearly exhaustive and is representative of the modern portion of the manufacturing sector. The data on man-hours worked, productive capital, share of labor in total cost, and value added are shown in tables 6-1, 6-2, and 6-3 for 1961, 1971, and 1975. Also listed are ρ and λ for the manufacturing sector. The stock of capital could not be deflated because proper price deflators and data on age or vintage of capital are lacking. Hence, all the figures listed in these tables are at current prices. Some states have a positive index, which indicates above-average productivity; a negative sign indicates below-average productivity. Because of biases introduced by inflation, intertemporal productivity growth only relative to the average can be analyzed.

In table 6-4 the states are ranked according to the interspatial productivity index and percent urban. Since the ranks by percent urban do not

Table 6-2. *Interspatial Productivity of the Manufacturing Sector,* ASI *Group, 1971*

State	*Millions of man-hours worked*	*Total productive capital (millions of rupees)*	*Share of labor in total cost (percent)*	*Value added (millions of rupees)*	*Interspatial productivity index*: *Factor productivity* (ρ)	*Interspatial productivity index*: *Total productivity* (λ)
Andhra Pradesh	411	5,603	8.3	1,076	0.3633	−0.0888
Assam	148	2,433	9.4	391	−0.4868	−0.2510
Bihar	442	8,753	6.6	1,668	0.7809	−0.0680
Gujarat	867	6,762	5.9	2,399	0.5905	0.4858
Haryana	169	2,681	6.8	602	−0.3855	0.0792
Himachal Pradesh	26	253	10.4	123	−2.7018	0.8075
Jammu and Kashmir	22	443	17.6	16	−2.2216	−1.7123
Karnataka	349	5,508	8.6	1,671	0.3339	0.3808
Kerala	267	2,565	9.1	800	−0.3871	0.3650
Madhya Pradesh	319	6,580	6.8	1,002	0.4921	−0.2889
Maharashtra	1,604	19,911	5.9	7,716	1.6372	0.6073
Orissa	162	4,846	7.8	567	0.1544	−0.5206
Punjab	142	2,906	5.9	503	−0.3236	−0.1634
Rajasthan	169	3,204	9.1	539	−0.2241	−0.1927
Tamil Nadu	861	10,782	8.1	2,696	1.0248	0.1682
Uttar Pradesh	636	11,282	7.7	1,879	1.0417	−0.2097
West Bengal	1,621	15,046	9.6	3,957	1.3862	0.1904
Delhi (union territory)	143	1,277	7.5	421	−1.0819	0.4180

Source: Same as table 6-1.

Table 6-3. *Interspatial Productivity of the Manufacturing Sector,* ASI *Group, 1975*

State	*Millions of man-hours worked*	*Total productive capital (millions of rupees)*	*Share of labor in total cost (percent)*	*Value added (millions of rupees)*	*Interspatial productivity index*		*Intertemporal productivity index (1971–75)*	
					Factor productivity (ρ)	*Total productivity* (λ)	*Factor productivity*	*Total productivity*
Andhra Pradesh	512	9,674	11.7	2,899	0.3136	−0.2041	0.5135	0.4294
Assam	165	3,602	8.3	1,069	−0.1367	−0.3432	0.3672	0.4875
Bihar	551	23,596	13.3	4,797	1.0925	−0.0712	0.9149	0.0712
Gujarat	818	15,233	11.9	5,014	0.7697	0.2958	0.7347	0.1382
Haryana	180	5,092	9.9	1,363	−0.3794	0.1424	0.5932	0.2446
Himachal Pradesh	28	582	28.2	175	−1.2926	−0.9970	0.6866	−0.4090
Jammu and Kashmir	42	476	20.8	44	−2.5990	−0.0712	0.1821	0.9093
Karnataka	417	8,933	17.6	2,903	0.2192	0.2999	0.4434	0.2849
Kerala	320	5,693	12.4	1,389	−0.2079	−0.0102	0.7310	0.1366
Madhya Pradesh	412	13,518	14.7	2,653	0.5702	−0.1408	0.6700	0.2053
Maharashtra	1,652	30,787	13.4	14,140	1.4748	0.6276	0.3967	0.3369
Orissa	188	7,083	14.9	955	−0.0969	−0.4958	0.3534	0.4322
Punjab	211	7,182	10.1	1,237	−0.0587	−0.2753	0.8641	−0.0790
Rajasthan	217	8,003	13.6	1,402	0.0313	−0.2404	0.8400	0.2052
Tamil Nadu	992	15,812	12.8	4,975	0.8292	0.2286	0.3577	0.4920
Uttar Pradesh	612	20,638	13.9	3,744	0.9908	−0.2173	0.5348	0.1675
West Bengal	1,636	18,944	20.4	7,876	1.0795	0.4377	0.1971	0.4126
Delhi (union territory)	168	3,257	15.9	652	−0.7749	0.1995	—	—

— Not applicable.
Source: Same as table 6-1.

Table 6-4. *Ranking of States by Productivity and Urbanization*

State	*Rank by productivity* (λ) *1961*	*1971*	*1975*	*Rank by percent urban, 1971*
Andhra Pradesh	12	10	10	7
Assam	13	14	14	16
Bihar	14	9	8	14
Gujarat	3	3	3	3
Haryana	n.a.	8	6	9
Himachal Pradesh	2	1	16	17
Jammu and Kashmir	16	17	17	8
Karnataka	8	5	4	5
Kerala	7	6	7	12
Madhya Pradesh	9	15	9	11
Maharashtra	1	2	1	1
Orissa	15	16	15	15
Punjab	11	11	13	6
Rajasthan	10	12	12	10
Tamil Nadu	4	7	5	2
Uttar Pradesh	5	13	11	13
West Bengal	6	4	2	4

n.a. Not available.
Sources: Tables 6-1, 6-2, 6-3, 6-5, 6-6, and 6-7.

change for 1961, 1971, and 1975, the table lists the ranks for 1971 only. Except for Bihar, Himachal Pradesh, Jammu and Kashmir, Kerala, and Punjab, the ranks given by the productivity index show significant correspondence with the ranks by percent urban. The states with high ranks by the productivity index, such as Maharashtra, Gujarat, Tamil Nadu, Karnataka, and West Bengal also have high percents urban, whereas the low-ranked states have low percents urban. In Punjab the productivity of the manufacturing sector is low because Punjab is predominantly agricultural (in the next section Punjab is shown to have a very high agricultural productivity index) and the proportion of its manufacturing sector is relatively small compared with that of Maharashtra, Gujarat, or West Bengal. Since the agricultural sector in Punjab is highly mechanized, most of the secondary agricultural activities take place in small towns which are classified as urban, but in the standard classification these activities do not come under the modern manufacturing sector. The urban population in Jammu and Kashmir mainly serves the tourist industry and defense establishments, both of which are excluded from the manufacturing sector. Kerala ranks low in urban population, but it has the highest literacy rate in India, and this may be one reason for its high productivity. Himachal Pradesh is an exception: it is mostly hilly, with few urban centers, and a small manufacturing sector; its productivity index is therefore difficult to interpret. For all the remaining states, however, the correspondence between the ranks by productivity index and by percent urban is remarkable. Furthermore, the highest-ranked states—Gujarat, Maharashtra, Tamil Nadu, and West Bengal—together accounted for about 55 percent of the total value added and about 52 percent of manufacturing employment in the seventeen states and Delhi (Central Statistical Organization 1978b).

In the above discussion the only measure of urbanization is percent urban. The interspatial productivity index, however, is affected not only by a state's percent urban but also by the number and size of its urban agglomerations and the population of its large cities. Therefore, in the following regression analysis five measures of urbanization are used:

- Percent urban
- Average size of urban agglomerations (the total population of the urban agglomerations divided by number of urban agglomerations)
- Total population of Class I cities (cities with 100,000 or more inhabitants)
- Total population of cities with 100,000–499,999 inhabitants (hereafter referred to as medium-size cities)
- Total population of cities with 500,000 or more inhabitants (hereafter referred to as large cities).

Tables 6-5, 6-6, and 6-7 list all these indicators. To take into account the characteristics of the work force, urban literacy is included as a proxy for labor quality (the ASI does not classify labor by skill). It would also be desirable to differentiate states by their level of technological development, but, as mentioned earlier, data are not available on the age or kind of capital. The simple correlation matrix between total productivity and these alternative measures of urbanization is given in table 6-8. Productivity is more closely correlated with the total population of large cities than with percent urban.

To test the aforesaid hypothesis, the S economic entities in equation (6.4) may be assumed to differ in their urban agglomerations (UA), that is, $F(\ln Y, \ln X, UA) = 1$. Thus, the effect of the urban agglomeration economies can be seen when the total productivity index is regressed over the alternative measures of urbanization: $\lambda = f$(urban literacy, percent urban, total population of Class I cities, . . .). The results of regressing the total productivity index on the noncollinear combinations of the alternative definitions of urbanization are shown in tables 6-9 and 6-10.

Table 6-5. *Urbanization in States, 1961*

State	Percent urban	Average size of urban agglomerations (thousands)	Total population (millions) Class I cities	Medium-size cities[a]	Large cities[b]	Urban literacy (percent)
Andhra Pradesh	18.30	29.57	3.376	1.907	1.496	31.22
Assam	7.11	13.17	0.138	0.138	0.000	41.62
Bihar	8.43	29.70	3.497	1.874	1.623	29.89
Gujarat	25.78	31.86	1.336	1.336	0.000	34.54
Himachal Pradesh	6.40	21.47	0.000	0.000	0.000	14.57
Jammu and Kashmir	16.57	19.67	0.551	0.551	0.000	18.89
Karnataka	22.33	27.72	2.585	1.138	1.477	32.73
Kerala	14.92	24.06	0.933	0.933	0.000	40.49
Madhya Pradesh	14.30	22.26	2.136	2.136	0.000	29.74
Maharashtra	28.21	46.69	9.498	1.911	7.587	36.30
Orissa	6.32	18.50	0.239	0.239	0.000	26.97
Punjab	23.04	24.02	0.842	0.842	0.000	38.93
Rajasthan	16.27	23.26	1.713	1.713	0.000	27.24
Tamil Nadu	26.28	33.92	4.699	2.306	2.393	36.05
Uttar Pradesh	12.82	38.36	6.221	4.277	1.944	30.70
West Bengal	24.44	71.26	6.062	1.662	4.400	41.33
Delhi (union territory)	88.75	88.75	3.573	0.402	3.171	36.56

Note: Haryana was not a state in 1961.
a. Cities with 100,000–499,999 inhabitants.
b. Cities with 500,000 or more inhabitants.
Source: Computed from Government of India (1961).

In table 6-9, which presents the results of the regressions run for the pooled data of 1961–75, the coefficient of percent urban is significant at the 5 percent level in equations 1 and 3. The coefficients of both the total population of Class I cities and average size of urban agglomerations are significant in equations 1 and 5. As before, the Class I cities are divided into two categories: medium-size cities (100,000–499,999 inhabitants) and large cities (500,000 or more inhabitants). The coefficient of the medium-size cities is negative (though insignificant), whereas that of the large cities is positive and significant in equations 2 and 3. Since the area and population size of states differ, the total populations of Class I cities and its two subcategories are normalized by dividing by the total urban population of the respective states. The normalized variables thus represent the percentage share of Class I cities, medium-size cities, and large cities in the total urban population of the states. The simple correlation coefficient matrix between the total productivity index and the urbanization variables

Table 6-6. *Urbanization in States, 1971*

State	Percent urban	Average size of urban agglomerations (thousands)	Total population (millions)			Urban literacy (percent)
			Class I cities	Medium-size cities[a]	Large cities[b]	
Andhra Pradesh	17.41	40.58	4.056	2.260	1.796	46.91
Assam	6.49	17.97	0.168	0.168	0.000	58.84
Bihar	9.10	34.97	2.550	2.550	0.000	44.85
Gujarat	26.18	37.50	3.642	1.990	1.752	54.74
Haryana	15.76	27.23	0.451	0.451	0.000	51.15
Himachal Pradesh	5.09	6.85	0.556	0.556	0.000	60.99
Jammu and Kashmir	16.36	19.11	0.677	0.677	0.000	38.24
Karnataka	22.41	30.96	4.899	1.984	2.915	51.49
Kerala	14.34	39.43	1.822	1.822	0.000	66.27
Madhya Pradesh	14.39	29.22	3.049	1.955	1.094	49.76
Maharashtra	29.27	61.12	11.463	3.060	8.403	58.10
Orissa	6.51	23.59	1.032	0.559	0.000	48.96
Punjab	21.81	31.32	1.530	1.380	0.000	52.33
Rajasthan	15.73	30.06	1.899	1.263	0.636	43.10
Tamil Nadu	28.36	51.70	7.086	2.469	4.617	56.78
Uttar Pradesh	12.12	42.48	7.196	3.356	3.840	43.87
West Bengal	22.85	81.86	7.770	0.749	7.031	55.83
Delhi (union territory)	88.69	86.10	3.643	0.000	3.643	59.10

a. Cities with 100,000–499,999 inhabitants.
b. Cities with 500,000 or more inhabitants.
Source: Computed from Government of India (1971).

with and without normalization are shown in table 6-8, and the results of regressing the total productivity index on these normalized variables are shown in table 6-9. The coefficient of the percentage share of medium-size cities is again found to be negative. In equations 4, 5, and 7 of table 6-9 it varies between −0.0094 and −0.0145 and is significant at the 5 percent level. The coefficient of the percentage share of large cities, however, is positive and ranges between 0.0052 and 0.0146. Since the large cities are highly correlated with the percentage share of Class I cities, the coefficient of Class I cities becomes biased and has shown a wrong sign, whereas the same coefficient is positive when paired with medium-size cities in equation 7. The coefficient of urban literacy does not seem to be significantly correlated with the productivity index.

In table 6-10, the same regression equation is run for three separate years—1961, 1971, and 1975. The contribution of the percentage share of Class I cities to the productivity index is positive and significant in

Table 6-7. *Urbanization in States, 1975*

State	Percent urban	Average size of urban agglomerations (thousands)	Total population (millions): Class I cities	Medium-size cities[a]	Large cities[b]	Urban literacy (percent)
Andhra Pradesh	18.25	53.23	4.690	2.637	2.030	48.50
Assam	7.00	28.71	0.199	0.199	0.000	60.52
Bihar	7.46	37.42	3.247	1.974	1.393	47.50
Gujarat	26.08	47.98	4.889	1.928	2.961	58.34
Haryana	16.96	36.64	1.022	1.022	0.000	53.26
Himachal Pradesh	6.72	7.27	0.857	0.857	0.000	64.16
Jammu and Kashmir	16.00	22.31	0.899	0.899	0.000	40.74
Karnataka	23.91	30.60	5.583	2.407	3.176	53.95
Kerala	14.22	56.13	1.705	0.830	0.875	70.10
Madhya Pradesh	15.31	34.94	4.237	2.283	1.954	51.00
Maharashtra	30.03	84.48	13.334	3.335	9.999	60.21
Orissa	6.82	30.15	0.925	0.925	0.000	51.00
Punjab	22.72	34.48	1.656	1.162	0.594	53.21
Rajasthan	15.93	36.61	2.515	1.787	0.828	45.26
Tamil Nadu	27.98	65.01	8.991	2.573	5.918	59.31
Uttar Pradesh	13.00	30.30	8.706	4.068	4.638	44.00
West Bengal	21.50	99.99	9.422	1.334	8.098	60.00
Delhi (union territory)	87.00	92.00	4.678	0.000	4.678	60.15

a. Cities with 100,000–499,999 inhabitants.
b. Cities with 500,000 or more inhabitants.
Sources: Government of India (1981). The figures above relate to midyear census projections based on decadal growth rates. Central Statistical Organization (1978b) also provides midyear population projections.

both 1961 and 1975. The coefficient of the percentage share of medium-size cities is consistently negative, while that of the large cities is positive in all sets of equations in the three periods. Although the coefficients are not significant in 1971, they do have the expected sign, and the only significant variable shown for that year is urban literacy. Because some of the census definitions of cities, towns, migrants, and urban agglomerations changed in the 1971 census, there are inconsistencies in the data. Nevertheless, in other regressions (not shown here) percent urban and the total population of Class I cities were observed to have positive and significant coefficients.

There appears to be a slowdown in the contribution of large cities. The coefficient of the percentage share of large cities has declined from between 0.011 and 0.034 in 1961 to between 0.006 and 0.016 in 1975. A 1 percent increase in the share of large cities in 1975 would contribute 0.006 to 0.016 decimal points to the total productivity index. Similarly,

Table 6-8. *Simple Correlation Coefficient Matrix between Productivity and Alternative Measures of Urbanization*

	Variable						
Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*
1	1.00000						
2	0.32013	1.00000					
3	0.22313	0.61510	1.00000				
4	−0.44962	−0.44481	0.10716	1.00000			
5	0.47885	0.80191	0.75992	−0.56417	1.00000		
6	0.44879	0.70145	0.66943	−0.53194	0.90168	1.00000	
7	0.23889	0.15005	0.37843	0.00736	0.30861	0.36279	1.00000

Variable definitions
1 = productivity index
2 = percent urban
3 = ratio of total population of Class I cities to total urban population
4 = ratio of total population of medium-size cities (100,000–499,999 inhabitants) to total urban population
5 = ratio of total population of large cities (500,000 or more inhabitants) to total urban population
6 = average size of urban agglomerations
7 = urban literacy
Source: Author's calculations.

the adverse effect of the percentage share of medium-size cities has also declined over time, which suggests that medium-size cities are no longer as much of a barrier to productivity growth of the manufacturing sector as they were in 1971.

It can be concluded from the regressions in tables 6-9 and 6-10 that the greater the proportion of cities with at least half a million inhabitants, the more the contribution to productivity through agglomeration economies. There appears to exist a threshold level of city population size which is required to provide agglomeration economies. Cities with less than half a million inhabitants are just not large enough to exploit agglomeration economies.

The same method can be applied to estimate the intertemporal productivity index g if the space variables are replaced by time variables t and $t + n$ as follows:

$$(6.11) \qquad g_{t+n,t} = \tfrac{1}{2} \sum_i (R_i^{t+n} + R_i^t)(\ln Y_i^{t+n} - \ln Y_i^t) - \tfrac{1}{2} \sum_j (W_j^{t+n} + W_j^t)(\ln X_j^{t+n} - \ln X_j^t).$$

(*Text continues on page 122.*)

Table 6-9. *Interspatial Productivity of the Manufacturing Sector: Pooled Regression Results, 1961–75*

Ordinary least squares variable	*Equation*						
	1	*2*	*3*	*4*	*5*	*6*	*7*
Intercept	**−0.8081**	−0.4404	**−0.6117**	−0.2279	−0.2919	−0.2285	−0.2280
t-ratio	3.11	1.63	2.32	0.28	0.97	0.81	0.81
Percent urban	**0.0096**		**0.0075**				
t-ratio	2.22		1.70				
Average size of urban agglomerations					*0.0051*		
t-ratio					1.31		
Total population of Class I cities	**0.0692**						
t-ratio	3.53						
Total population of medium-size cities[a]		−0.0083					
t-ratio		1.16					
Total population of large cities[b]		**0.1256**	**0.1033**				
t-ratio		4.19	3.61				

Ratio of total population of Class I cities to total urban population						**−0.0094**	*0.0051*
t-ratio						2.05	1.63
Ratio of total population of medium-size cities[a] to total urban population				**−0.0094**	**−0.0104**		**−0.0145**
t-ratio				2.05	2.31		3.93
Ratio of total population of large cities[b] to total urban population				*0.0052*		**0.0146**	
t-ratio				1.63		3.93	
Urban literacy	*0.0068*	0.0044	0.0051	0.0072	0.0073	0.0071	0.0071
t-ratio	1.28	0.78	0.95	1.25	1.24	1.25	1.25
R^2	0.3782	0.3504	0.3848	0.2991	0.2859	0.2988	0.2991

Note: Coefficients in boldface are significant at the 5 percent level; those in italics are significant at the 10 percent level.

a. Cities with 100,000–499,999 inhabitants.

b. Cities with 500,000 or more inhabitants.

Source: Author's calculations.

Table 6-10. *Interspatial Productivity of the Manufacturing Sector: Regression Results, 1961, 1971, 1975*

	1975				*1971*		*1961*			
Ordinary least squares variable	*Equation 1*	*Equation 2*	*Equation 3*	*Equation 4*	*Equation 5*	*Equation 6*	*Equation 7*	*Equation 8*	*Equation 9*	*Equation 10*
Intercept	−0.1924	−0.2646	−0.2190	−0.2382	−2.7288	−2.9558	0.3615	0.2691	0.3593	0.3586
t-ratio	0.27	0.37	0.30	0.34	3.46	3.81	0.47	0.40	0.47	0.47
Percent urban						0.0016		**−0.0291**		
t-ratio						0.22		1.75		
Average size of urban agglomerations				0.0086						
t-ratio				2.00						
Ratio of total population of Class I cities to total urban population	**0.0061**	*−0.0091*				0.0010	*0.0114*			−0.0120
t-ratio	1.73	1.52				0.14	1.63			1.15

Ratio of total population of medium-size cities[a] to total urban population	**−0.0161**		*−0.0093*	*−0.0089*	−0.0030		**−0.0234**		−0.0119	
t-ratio	3.77		1.54	1.56	0.40		2.24		1.15	
Ratio of total population of large[b] cities to total urban population		**0.0160**	**0.0067**		0.0018			**0.0346**	*0.0114*	**0.0235**
t-ratio		3.76	1.73		0.37			2.62	1.64	2.25
Urban literacy	0.0038	0.0048	0.0042	0.00076	**0.0532**	**0.0548**	−0.0103	−0.0088	−0.0102	−0.0101
t-ratio	0.36	0.46	0.40	0.07	3.80	3.81	0.48	0.44	0.47	0.47
R^2	0.6287	0.6276	0.6287	0.6494	0.5686	0.5458	0.3191	0.3943	0.3198	0.3204

Note: Coefficients in boldface are significant at the 5 percent level; those in italics are significant at the 10 percent level.
a. Cities with 100,000–499,999 inhabitants.
b. Cities with 500,000 or more inhabitants.
Source: Author's calculations.

Since some states were reorganized after 1961, the intertemporal productivity index has been calculated only for the period 1971–75 (see the last two columns of table 6-3). The productivity index has increased most over time in states which were below the average. Except for Himachal Pradesh and Punjab, all states show increases in productivity through time—especially Andhra Pradesh, Assam, Jammu and Kashmir, and Orissa, where industrialization began rapidly after the fourth five-year plan. These four states, and Orissa and Madhya Pradesh in particular, were also the states in which many public projects and tribal area development plans were specially designed. Factor productivity, however, increased most in Bihar, followed by Punjab, Rajasthan, Gujarat, Kerala, Himachal Pradesh, and Madhya Pradesh. Despite this increase in productivity, these states were still below the average level in 1975.

From tables 6-1 to 6-7 it can be concluded that total productivity was consistently higher in states where either the percent urban was relatively high or the total population of large cities was relatively high. For example, states such as Maharashtra, Gujarat, Karnataka, Tamil Nadu, and West Bengal, which have a large number of Class I cities and therefore high percents urban, seem to have the most urban agglomeration economies. Hence, total productivity is higher in these states than in the less urbanized states.

Productivity and Migration

Let us examine how interspatial productivity differences are related to population migration to urban areas. There are no annual census data on migration rates. If we assume a two- to three-year lag in the response to migration, data from the National Sample Survey (NSS) can be used. The NSS 28th Round in 1973–74, for example, covered 129,000 rural households from 9,034 villages and 67,000 urban households from 5,034 urban blocks. A person was considered a migrant if he or she had normal residence one year previously at a place other than the place of current residence. The NSS also classified migration among and within states. The percentage distribution was as follows:

Flow	*Within states*	*Between states*	*Total*
Rural to rural	48	4	52
Rural to urban	13	3	16
Urban to rural	10	3	13
Urban to urban	16	3	19
All India	87	13	100

Movement within states accounted for 87 percent of the total, and movement between states for 13 percent. According to the NSS the search for work was the most important reason for the migration of males, followed by the opportunity to study. For females, however, the most important reason was marriage.[2]

It can be seen from table 6-11 that, except for Uttar Pradesh, states with a positive interspatial productivity index (which indicates a higher than average level of productivity) also have in-migration rates that are higher than the all-India average. It is also true that higher productivity leads to higher returns to labor (not shown here) and therefore higher rates of in-migration (Todaro 1969). Of course, there are many factors

Table 6-11. *Migration and Productivity in Urban Areas*

State	*Interspatial productivity index of the manufacturing sector, 1971*	*In-migration rate per 10,000 population in urban areas, 1973–74*[a]
Himachal Pradesh	0.8075	**305**
Maharashtra	0.6073	**221**
Gujarat	0.4858	**256**
Karnataka	0.3808	**184**
Kerala	0.3650	**373**
Uttar Pradesh	0.2097	143
West Bengal	0.1904	**185**
Tamil Nadu	0.1682	**276**
Haryana	0.0792	**290**
Bihar	−0.0680	82
Andhra Pradesh	−0.0888	**222**
Punjab	−0.1634	109
Rajasthan	−0.1927	**220**
Assam	−0.2510	83
Madhya Pradesh	−0.2889	150
Orissa	−0.5206	**441**
Jammu and Kashmir	−1.7123	101
Delhi (union territory)	0.4180	**530**
All India	—	180

— Not applicable.

Note: Figures in boldface are above the all-India average.

a. Rates of in-migration per 10,000 persons by sex and age are obtained by dividing the number of migrants within each group, sex, and age by the corresponding population and then multiplying by 10,000. The rates are calculated separately for rural and urban areas.

Sources: NSS (1974); Mukherji and Banerji (1978); table 6-2.

other than wages—such as housing, public amenities, and transportation—which affect migration, but a detailed analysis of migration lies beyond the scope of this chapter. The productivity level, however, does appear to have a significant impact on migration rates.

Interspatial Productivity in Agriculture

The method of Caves, Christensen, and Diewert is again used to estimate the interspatial productivity index of the agricultural sector. Unfortunately, the time series on capital assets in agriculture is not available, and thus the intertemporal index could not be calculated. Data on the agricultural labor force, however, were compiled from the 1971 census, which indicated that 70 percent of India's total work force was involved in cultivation and farming.

The data on capital assets in agriculture were taken from a report on the assets of rural households published by the Reserve Bank of India

Table 6-12. *Aggregate Value of Assets of All Rural Households, June 1971*

State	*Estimated number of households (thousands)*	*Aggregate value of assets (crores of rupees[a])*
Andhra Pradesh	6,690	5,406.0
Assam	1,937	1,517.0
Bihar	8,630	11,070.1
Gujarat	3,774	4,858.8
Haryana	1,058	2,871.0
Himachal Pradesh	460	1,042.7
Jammu and Kashmir	558	851.0
Karnataka	4,174	4,187.0
Kerala	2,604	3,024.5
Madhya Pradesh	5,915	6,222.3
Maharashtra	5,994	7,002.4
Orissa	3,672	2,211.8
Punjab	1,551	4,936.4
Rajasthan	3,131	3,993.1
Tamil Nadu	6,433	4,391.7
Uttar Pradesh	13,709	18,549.9
West Bengal	6,081	4,458.1
All India	77,035	87,131.6

a. One crore is equal to 10 million rupees.
Source: Reserve Bank of India (1976).

(1976). The only lacuna in these data is the exclusion of urban agriculture and mining and quarrying activities, which are otherwise listed in the agricultural sector. Since these categories account for hardly 2–3 percent of the total labor force, they can be ignored without causing serious error.

The composition of assets is as follows:

Asset	*Percentage of total assets*
Land	66.0
Housesite	0.4
Buildings	17.4
Livestock	6.5
Implements and machinery	2.7
Durable household goods	4.6
Financial assets	1.1
Cash	3.3
Kind	0.8

Data on each state's aggregate value of assets and number of rural households are presented in table 6-12. All we need is the share of labor in the total cost of agricultural production. Since data on production costs are not available, the share of labor is replaced by the income share of labor in total income originating in the agricultural sector. For income originating in the agricultural sector, state domestic product was divided into agriculture, manufacturing, and services at constant and current prices using the consumer price index for agricultural laborers (tables 6-13 and 6-14). The NSS publishes agricultural wages for male and female workers by different operations in agriculture (Satyanarayana 1981). An average wage rate for all agricultural operations was calculated, and it was multiplied by the male and female work force to get the total wage bill (tables 6-15 and 6-16).[3] The estimated income share of labor is the wage bill divided by the total income originating from the agricultural sector.

The interspatial factor productivity and total productivity indices for the agricultural sector in the seventeen states in this study are given in table 6-17. Interspatial factor productivity has increased most in Bihar and Uttar Pradesh, followed by Maharashtra, Madhya Pradesh, and Andhra Pradesh, while total productivity has increased most in West Bengal, Punjab, Haryana, and Orissa. All these states lie above the average. It is interesting that industrially backward states in which the total productivity index for manufacturing was either low or negative (see table 6-2) show a high and positive productivity index for the agricultural sector: Haryana, Punjab, Rajasthan, and Orissa are examples. Furthermore, the greater the wealth and asset holdings per household, the higher the inter-

Table 6-13. *State Domestic Product by Sector*
(millions of rupees at constant prices, 1960 = 100)

State and year	*Agriculture*	*Manufacturing*	*Services*
Andhra Pradesh			
1961	5,961.33	917.03	3,702.82
1971	6,913.60	2,229.50	3,746.90
Assam			
1961	1,096.00	661.00	803.00
1971	1,758.00	953.00	1,188.00
Bihar			
1961	6,822.00	2,688.00	3,870.00
1971	8,795.00	4,252.00	6,172.00
Gujarat			
1961	3,060.00	1,920.00	2,400.00
1971	4,940.00	2,870.00	3,680.00
Haryana			
1961	1,536.70	400.60	513.00
1971	2,582.10	812.20	961.00
Himachal Pradesh			
1961	n.a.	n.a.	n.a.
1971	679.89	200.55	340.73
Jammu and Kashmir			
1961	n.a.	n.a.	n.a.
1971	492.10	279.30	555.67
Karnataka			
1961	3,971.50	1,343.30	1,323.80
1971	5,254.00	1,930.03	3,275.56
Kerala			
1961	2,245.60	759.56	1,313.98
1971	2,275.50	1,252.30	2,362.00
Madhya Pradesh			
1961	5,280.00	1,243.00	1,800.00
1971	6,003.00	2,097.00	2,665.00
Maharashtra			
1961	6,642.50	4,621.60	5,070.10
1971	5,535.40	7,167.50	8,270.90
Orissa			
1961	2,293.60	535.90	943.00
1971	3,352.20	891.40	1,622.00
Punjab			
1961	2,180.70	630.90	1,226.10
1971	3,454.60	1,107.00	2,095.70
Rajasthan			
1961	2,915.70	1,022.90	2,496.50
1971	5,018.31	1,117.64	2,091.20
Tamil Nadu			
1961	5,780.20	1,955.60	3,382.70
1971	6,244.80	3,450.70	4,970.80

Table 6-13 *(continued)*

State and year	*Agriculture*	*Manufacturing*	*Services*
Uttar Pradesh			
1961	9,380.87	3,702.52	5,326.27
1971	10,587.50	5,113.60	7,429.90
West Bengal			
1961	5,659.10	3,636.40	4,929.60
1971	6,944.17	3,846.03	6,161.09

n.a. Not available.

Sources: State governments of India (1962–73); *Indian Labour Journal* (1978). The Centre for Monitoring Indian Economy, Bombay, recently released a consistent series of data on the state domestic product (SDP). The problem with using this series is that the SDP is at current prices and different states have used different base years. Furthermore, the sectoral composition of economic activities is not uniform among the states. I therefore prefer to formulate SDP from the individual state publications.

Table 6-14. *Income Originating from Agriculture*
(millions of rupees)

State	*Constant price (1960 = 100), 1970–71*	*Current price, 1970–71*
Andhra Pradesh	6,913.60	13,274.112
Assam	1,958.00	3,287.460
Bihar	8,795.00	16,270.750
Gujarat	4,940.00	8,793.200
Haryana	2,582.10	5,139.379
Himachal Pradesh	679.89	1,339.380
Jammu and Kashmir	492.10	871.017
Karnataka	5,254.00	10,140.856
Kerala	2,275.50	4,505.490
Madhya Pradesh	6,003.00	12,126.060
Maharashtra	5,535.40	10,406.550
Orissa	3,352.20	6,503.260
Punjab	3,454.60	6,805.560
Rajasthan	5,018.31	8,330.400
Tamil Nadu	6,244.80	11,115.740
Uttar Pradesh	10,587.50	20,539.750
West Bengal	6,944.17	12,846.710

Sources: Indian Labour Journal (1978); table 6-13.

Table 6-15. *Distribution of Workers by Sector, 1971*
(millions)

State	*Agriculture*	*Manufacturing*	*Services*
Andhra Pradesh	13.370	1.622	2.057
Assam	3.279	0.360	0.600
Bihar	14.701	0.884	1.902
Gujarat	5.703	1.006	1.684
Haryana	1.775	0.265	0.613
Himachal Pradesh	0.991	0.053	0.233
Jammu and Kashmir	0.984	0.092	0.327
Karnataka	7.307	1.045	1.826
Kerala	3.382	1.073	1.759
Madhya Pradesh	12.505	1.027	1.763
Maharashtra	12.274	2.404	3.714
Orissa	5.506	0.409	0.935
Punjab	2.490	0.442	0.979
Rajasthan	6.073	0.540	1.435
Tamil Nadu	9.546	1.964	3.230
Uttar Pradesh	21.328	1.994	4.009
West Bengal	7.779	1.740	2.922

Source: Compiled from Government of India (1971, general tables).

spatial productivity index in the agricultural sector (see tables 6-12 and 6-17).

It can also be confirmed that interspatial productivity in the agricultural sector is intimately related to the output-labor ratio (O/L) (table 6-18). It can be seen from table 6-19 that, with the exception of Kerala, every state that has an O/L value that is above the all-India average also has a positive interspatial productivity index. In the rest of the states the productivity index is either very low or negative and is associated with a less than average value of O/L.

Data from the NSS 28th Round on in-migration rates for rural areas are also presented in table 6-19. In every state except Maharashtra, Jammu and Kashmir, and Gujarat, the in-migration rate appears to be associated with either the output-labor ratio or the interspatial productivity index. The high migration rate in Jammu and Kashmir can be explained by the presence of defense establishments which, as mentioned earlier, are excluded from the manufacturing sector. It is unclear, however, why the interspatial productivity in Gujarat is positive although the migration rate there is unusually low.

In table 6-19, rural in-migration does not seem to be correlated with wage rates because agricultural wages are not fully adjusted for payments

Table 6-16. *Agricultural Wages by State, at Constant and Current Prices, 1970–71*
(rupees)

State	Ratio of female workers to male workers[a]	Current average daily agricultural wage	Agricultural consumer price index (1960 = 100)	Real daily wage
Andhra Pradesh	0.70	1.57	171	0.91
Assam	0.89	1.90	203	1.66
Bihar	0.85	1.90	206	0.92
Gujarat	0.82	2.29	173	1.32
Haryana	0.95	3.97	194	2.04
Himachal Pradesh	0.71	3.50	194	1.80
Jammu and Kashmir	0.93	3.41	167	2.04
Karnataka	0.78	2.83	188	1.50
Kerala	0.75	3.18	214	1.50
Madhya Pradesh	0.73	1.47	198	0.74
Maharashtra	0.68	1.87	192	0.97
Orissa	0.89	1.71	212	0.80
Punjab	0.98	5.00	194	2.57
Rajasthan	0.86	1.99	173	1.24
Tamil Nadu	0.77	2.09	174	1.14
Uttar Pradesh	0.89	2.48	183	1.14
West Bengal	0.91	2.15	206	1.14

a. Computed from Government of India (1971).
Sources: NSS (1971); Lal (1976); Satyanarayana (1981).

in kind—food, clothing, and shelter—although the payment in kind is converted into cash value at wholesale prices with a 10 percent markup for retail price adjustment. These wages, therefore, may represent cash value of food and clothing, but the rental value of shelter still varies among villages and farms in ways not fully accounted for in wages. In some instances farmers and landless workers are given free shelter by the employer, and this is not accounted for in the imputations. I therefore do not consider wage differences, which are small anyway, as the main incentive for migration. Productivity differentials and O/L values are more suitable than wages for explaining migration into rural areas.

It is also true that wage differentials encourage more rural to urban than rural to rural migration. Higher wages in one rural area may attract only temporary migrants from neighboring regions. Such migrants do not like to lose contact with their native villages. For instance, temporary migration (one or two months) into Haryana and Punjab—where wages

(Text continues on page 133.)

Table 6-17. *Interspatial Productivity of the Agricultural Sector, at Current Prices, 1970–71*

State	Millions of man-years worked	Aggregate value of capital assets (millions of rupees)	Income originating from agriculture (millions of rupees)	Total wage bill (millions of rupees)	Share of labor in total cost (percent)	Interspatial productivity index: Factor productivity (ρ)	Interspatial productivity index: Total productivity (λ)
Andhra Pradesh	3,208.90	54,060	13,274.11	2,920.11	42.2	0.5897	0.0500
Assam	787.08	15,170	3,287.46	1,306.56	74.3	−0.6727	0.0833
Bihar	3,528.47	110,701	16,270.75	3,528.69	40.1	1.0257	−0.1825
Gujarat	1,368.84	48,588	8,793.20	1,806.87	36.6	0.1508	0.0770
Haryana	426.01	28,710	5,138.38	869.07	33.7	−0.6355	0.3261
Himachal Pradesh	237.85	10,427	1,339.38	428.14	63.0	−1.5293	0.1247
Jammu and Kashmir	236.18	8,510	871.02	481.77	97.9	−1.6573	−0.4260
Karnataka	1,753.82	41,870	10,140.86	2,630.73	50.0	0.1849	0.1855
Kerala	811.74	30,245	4,505.49	1,217.62	53.5	0.3631	−0.8040
Madhya Pradesh	3,001.22	61,223	12,126.06	2,220.90	37.0	0.6272	−0.0780
Maharashtra	2,945.85	70,024	10,406.55	2,857.48	51.6	0.7031	−0.3068
Orissa	1,321.46	22,118	6,503.26	1,057.17	31.5	−0.3316	0.2577
Punjab	597.83	49,364	6,805.56	1,536.43	44.5	−0.4594	0.4310
Rajasthan	1,457.71	39,851	8,330.40	1,807.57	36.0	0.0204	0.1535
Tamil Nadu	2,291.20	43,917	11,115.74	2,611.97	41.8	0.2970	0.1652
Uttar Pradesh	5,118.94	185,499	20,539.75	5,835.59	55.1	1.0220	0.0542
West Bengal	1,867.15	44,851	12,846.71	2,240.58	32.3	0.1637	0.4432

Sources: Central Statistical Organization (1978b) and author's calculations.

Table 6-18. *Output-Labor Ratio by Sector, at Constant Prices, 1961–71*
(thousands of rupees)

State and year	*Agriculture*	*Manufacturing*	*Services*
Andhra Pradesh			
1961	445.400	400.37	1,239.23
1971	516.385	1,373.88	1,675.15
Assam			
1961	470.59	1,725.84	1,108.49
1971	536.06	2,644.87	1,979.57
Bihar			
1961	442.05	1,812.54	1,669.44
1971	598.22	4,805.28	3,244.75
Gujarat			
1961	520.77	1,611.07	1,705.86
1971	866.14	2,850.39	2,184.51
Haryana			
1961	n.a.	n.a.	n.a.
1971	1,454.64	3,063.28	1,565.73
Himachal Pradesh			
1961	n.a.	n.a.	n.a.
1971	606.01	376.90	1,462.36
Jammu and Kashmir			
1961	n.a.	n.a.	n.a.
1971	500.05	3,008.41	1,698.15
Karnataka			
1961	454.72	1,188.14	2,696.10
1971	719.02	1,820.49	1,793.41
Kerala			
1961	849.39	746.10	667.67
1971	672.76	1,165.05	1,342.25
Madhya Pradesh			
1961	379.34	1,055.18	983.06
1971	480.04	2,041.19	1,511.16
Madras			
1961	594.57	951.57	946.27
1971	n.a.	n.a.	n.a.
Maharashtra			
1961	486.47	2,163.16	1,606.71
1971	450.97	2,981.43	2,226.86
Orissa			
1961	396.13	868.56	750.80
1971	608.81	2,177.76	1,733.74
Punjab			
1961	473.38	704.98	766.52
1971	1,386.84	2,499.10	2,140.59
Rajasthan			
1961	382.65	1,328.07	2,092.62
1971	826.22	2,068.29	2,013.66

(Table continues on the following page.)

Table 6-18 (*continued*)

State and year	*Agriculture*	*Manufacturing*	*Services*
Tamil Nadu			
1971	654.13	1,756.69	1,538.48
Uttar Pradesh			
1961	429.11	1,422.63	1,214.37
1971	496.39	2,564.01	1,852.85
West Bengal			
1961	831.22	2,012.99	1,852.85
1971	892.58	2,209.75	2,108.35
All India			
1961	546.77	1,310.74	1,305.29
1971	600.31	2,397.84	1,974.12

n.a. Not available.
Sources: Tables 6-13 and 6-15.

Table 6-19. *Migration and Productivity in Rural Areas*

State	*Interspatial productivity index of the agricultural sector, 1971*	*In-migration rate per 10,000 population in rural areas, 1973–74*	*Output/labor ratio (thousands of rupees)*	*Real daily wages per worker, 1971 (rupees)*
West Bengal	0.4432	**91**	**892**	1.14
Punjab	0.4310	**109**	**1,386**	2.57
Orissa	0.2577	**122**	**608**	0.80
Karnataka	0.1855	**116**	**719**	1.50
Tamil Nadu	0.1652	**133**	**654**	1.14
Rajasthan	0.1535	**92**	**826**	1.24
Haryana	0.0772	**103**	**1,454**	2.04
Gujarat	0.0770	25	**866**	1.32
Uttar Pradesh	0.0542	56	496	1.14
Andhra Pradesh	0.0500	**147**	516	0.91
Madhya Pradesh	−0.0780	59	480	1.74
Assam	−0.0833	29	536	1.66
Kerala	−0.0840	**190**	**672**	1.50
Himachal Pradesh	−0.1247	87	606	1.80
Bihar	−0.1825	26	398	0.92
Maharashtra	−0.3068	**144**	480	0.97
Jammu and Kashmir	−0.4260	**134**	500	2.04
All India	n.a.	90	610	n.a.

n.a. Not available.
Note: Figures in boldface are above the all-India average.
Sources: NSS (1974); Mukherji and Banerji (1978).

are higher than in neighboring states—greatly increases during the harvest season, but once the harvest is over, the workers either return to their villages or seek work in nearby cities. People with even a small piece of land are reluctant to leave their village unless the land is completely barren or the destination is within 20–40 miles of their village. Should farmers or urbanites decide to migrate permanently (for at least one year, as defined by the NSS) into rural areas, they would do so with the intention of either establishing farms or setting up a small-scale agro-based factory or service unit. For such migrants the prime consideration would be the productivity comparison, not the wage differential. The productivity difference may be attributed to several other factors, such as better marketing facilities, abundance of capital and intermediate inputs, better weather conditions and less frequent floods and droughts, government rent and price controls, and taxes which cannot be completely capitalized into land prices. For landless migrants the better option is likely to be migration to urban areas.

Notwithstanding the fact that agriculture in India is still a gamble with the monsoons, it appears that higher assets and wealth holdings for households result in higher interspatial productivity, higher output-labor ratios, and higher in-migration rates. Higher asset and wealth holdings per household probably also reflect higher levels of technology and therefore higher levels of productivity, at least in Haryana, Punjab, and West Bengal.

Conclusions

Without data on capital stock by cities, it is difficult to quantify the contribution of urban agglomeration economies in raising manufacturing productivity in cities. Nonetheless, for the country as a whole, the percent urban and the population of Class I cities have significant effects on manufacturing productivity. When the population of Class I cities is divided into two subcategories of population—medium-size (100,000–499,999) and large (500,000 or more)—it becomes clear that large cities provide significant agglomeration economies. No matter how urbanization is defined, whether by percent urban, by population of Class I cities, or by population of large cities, the coefficients are significant.

The coefficient of large cities declined from 0.2315 in 1961 to 0.0822 in 1975, while the coefficient of medium-size cities increased from 0.2156 in 1971 to 0.2433 in 1975. This depicts an important phenomenon of Indian urbanization. The contribution of large cities to the productivity differential of the manufacturing sector is declining, whereas that of the medium-size cities is rising. In the mid-1970s large cities, such as Bombay, Calcutta, Delhi, Kanpur, and Madras, slowed in growth as compared with medium-size cities. Large cities seem to have exploited agglomeration economies during the 1960s (the coefficient of large cities is very large and significant for 1961), and they may have grown so big that their size has now become a disamenity. Rising rents and commuting costs have also discouraged migration to large cities. New factory establishments find it profitable to locate in the outer periphery of large cities. Medium-size cities located near large cities have greater growth potential than cities that are not near large cities.

If data on capital and disaggregated output are compiled by city, the scope of this productivity analysis can be extended to shed further light on the role of urban agglomeration economies in India's city size distribution. Many amenity variables—to take into account such factors as the availability of housing, school and hospital facilities, drinking water, power, and so on—can be included in the utility function of a city's population to analyze how a city resident or migrant balances amenities and disamenities with the higher returns available from growing cities.

When the productivity of the agricultural sector is compared among states, the theory that states with higher assets and wealth per household are more productive is confirmed. Extending the analysis is difficult because of the absence of data on production cost and capital stock in agriculture. Furthermore, because production in agriculture depends so much on the monsoon, a simple measure of productivity is likely to have a large error.

References

Aberg, Y. 1973. "Regional Productivity Differences in Swedish Manufacturing." *Regional and Urban Economics* no. 3: 131–57.

Alonso, William. 1968. "Urban and Regional Imbalances in Economic Development." *Economic Development and Cultural Change* no. 3: 1–14.

Caves, D. W., L. R. Christensen, and E. W. Diewert. 1982. "Multilateral Comparisons of Output, Input and Productivity Using Superlative Index Numbers." *Economic Journal* no. 92 (March).

Central Statistical Organization. 1978a. *Monthly Abstract of Statistics, 1978.* New Delhi: Ministry of Planning.

———. 1978b. *Statistical Abstract of India, 1978.* New Delhi: Ministry of Planning.

Denny, Michael, Melvyn Fuss, and J. D. May. 1981. "Intertemporal Changes in Regional Productivity in Canadian Manufacturing." *Canadian Journal of Economics* (August).

Diewert, Erwin W. 1976. "Exact and Superlative Index Numbers." *Journal of Econometrics* 4, no. 2.

Government of India. 1961. *Census of India, 1961.* Series 1. New Delhi: Census Commissioner of India.

———. 1971. *Census of India, 1971.* Series 1. New Delhi: Census Commissioner of India.

———. 1981. *Census of India, 1981.* Series 1, Provisional Population Tables. New Delhi: Census Commissioner of India.

Indian Labour Journal. 1978. Simla: Ministry of Labor.

Jorgenson, D. W., and Zvi Griliches, 1967. "The Explanation of Productivity Change." *Review of Economic Studies* 34.

Jorgenson, D. W., and Mieko Nishimizu. 1978. "U.S. and Japanese Economic Growth, 1952–74: An International Comparison." *Economic Journal* 88.

Lal, Deepak. 1976. "Money and Real Wages of Male Agriculture Laborers in India." *Economic and Political Weekly* (June).

Mera, Koichi. 1973. "On the Urban Agglomeration and Economic Efficiency." *Economic Development and Cultural Change* no. 21: 309–24.

Mukherji, S. R., and S. K. Banerji. 1978. "Some Aspects of Internal Annual Migration in India." *Barvekshna* (October). New Delhi.

NSS. 1971. *National Sample Survey, 25th Round: 1970–1971.* New Delhi: Ministry of Planning.

———. 1974. *National Sample Survey, 28th Round: 1973–1974.* New Delhi: Ministry of Planning.

Reserve Bank of India. 1976. "Report on Assets of Rural Households." Bombay.

Satyanarayana, Y. 1981. "Wage Trends in India: 1830–1976—An Analysis, Discussion and Compilation of Data." Washington, D.C.: World Bank, Employment and Rural Development Department, August. Processed.

Segal, David. 1976. "Are There Returns to Scale in City Size." *Review of Economics and Statistics* no. 58: 339–51.

Solow, Robert. 1957. "Technical Change and Aggregate Production Function." *Review of Economics and Statistics* no. 39.

State governments of India. 1962–73. *Statistical Outline of States.* Published individually by states by the state directorates of economics and statistics.

Sveikauskas, Leo. 1975. "The Productivity of Cities." *Quarterly Journal of Economics* no. 89: 392–413.

Todaro, Michael P. 1969. "A Model of Labour Migration and Urban Unemployment in LDC's." *American Economic Review* 59, no. 1.

Wheaton, William C., and Hisanobu Shishido. 1981. "Urban Concentration, Agglomeration Economies, and the Level of Economic Development." *Economic Development and Cultural Change* no. 30 (October): 16–30.

Notes

1. In census tables a town area is an area not classified as urban.
2. See also Mukherji and Banerji (1978).
3. The work force data were taken from the 1971 census. A standard 240 man-days a year was applied to derive the total working hours.

7

Indian Government Programs to Alter City Sizes

THIS CHAPTER IS CONCERNED with normative or government policy issues regarding the sizes of cities. The word "cities" is used here to mean generic urban areas, or "urban agglomerations," as they are called in India. It is widely believed in both developing and developed countries that large cities are too large and grow too rapidly. Governments in many countries have adopted a variety of programs to limit the growth or reduce the size of their large cities.[1] Most programs focus on the location of new employment opportunities, especially in manufacturing. Direct controls such as outright prohibitions on movement, stringent licensing requirements, land use regulations, location policies regarding government-owned businesses, and input rationing are often used. Indirect controls such as investment subsidies and tax concessions on business and housing investment at locations outside large cities are also common.

Some programs are aimed directly at population dispersal, others at both residential and production activities. Public housing may be withheld from large cities or concentrated in other places. Nonagricultural activities may be prohibited within greenbelts that are established around large cities.

There is much more government action to control the sizes of cities than serious economic analysis of the desirability and effects of such programs. Theoretical models of city sizes generated by market interactions are abstract and artificial. The pertinent literature includes several statistical studies of actual city sizes similar to the analysis presented in chapter 4, but there is nothing in the literature to distinguish between actual and desirable city size distributions. Fortunately, careful studies in recent years have begun to present theoretical and empirical analyses of specific phenomena that may prevent cities from reaching or staying at their optimal size.[2]

In this situation any discussion of programs to control city sizes must be tentative. Governments may have a variety of motives for attempting to control the sizes of their large cities, and it is difficult to know whether the goals are appropriate—not to mention whether the programs are the most efficacious ways to achieve the goals or whether the social costs of such programs outweigh the social benefits. In many cases it is difficult to ascertain how strenuously control programs have been applied. Inevitably, it is difficult to measure their effects.

The final point to be made in this introduction is that city sizes are not arbitrary or capricious, and governments should not attempt to alter them without carefully studying the problems and alternative solutions. City sizes are determined by millions of decisions, government and private, by people and firms as to where to locate, work, and produce. Decisions to move to large cities are made as carefully as possible by households in pursuit of higher living standards. Decisions to locate production facilities in large cities are made by firms seeking the most production and employment creation from valuable investments. Such decisions lead to a specific set of sizes and locations of cities.

The result is a set of relative city sizes that has shown extraordinary persistence and stability over long historical periods. As we saw in chapter 4, this persistence and stability is borne out in twentieth-century data from India and in nineteenth- and twentieth-century data from other countries. That relative city sizes are persistent does not prove that they represent a socially efficient set of input and output locations. That, indeed, is the subject of this chapter. But it does mean that attempts to alter the city size distribution are as important and difficult as almost any program that governments may undertake. The reasons for attempting to alter the sizes of cities and the range of alternative programs available to do so should be studied carefully. More commonly, governments simply decide to discourage locations in large cities by means of a variety of ad hoc programs and without study of the likely success or costs of such efforts.

Careful attempts to interfere with the process of city size and location determination can improve things, but misguided attempts to reduce the size or growth rates of large cities can do great harm, insofar as they are successful. The greatest danger is that desperately needed industrial production will be prevented from growing to its full potential because factories and businesses are forced or induced to locate in the wrong place. Government programs to alter city sizes should be studied with the greatest care as to their desirability, alternative means of achieving the goal, and the benefits and costs of doing so.

Altering city sizes through government programs is not easy. Despite some three decades of effort, Indian governments have had no measurable effect on the city size distribution. The first requirement is a better understanding of the likely causes of undesirable city sizes. Then programs can be designed with the causes of the problems in mind. We will undertake these two tasks after a brief survey of government programs to control the sizes of Indian cities.

Programs to Control the Size of Large Cities

In principle, almost every tax, expenditure, or regulatory action taken by a government affects the location of some economic activities and therefore the sizes of some cities relative to others. But most such actions have other goals, and their effects on city sizes are small and incidental. Furthermore, the effects of many government actions on city sizes may not be consistent or easy to predict. Thus, only those government programs explicitly designed to affect relative city sizes will be discussed in this section.[3]

Ambiguity exists even within the set of actions intended to have locational effects. Some programs are intended to shift locations of economic activities to small cities, some to shift them to rural areas, some to shift them to poor states or districts, and some to achieve combinations of these goals. All these goals militate against locations in large cities, but each may do so to a different degree. In this section, programs with any combination of these goals will be discussed.

There is no comprehensive document that describes Indian government policies on city size. Policies are contained in laws, government reports, and five-year plans. The earliest national government programs to influence the location of industry made use of industrial licenses. A law enacted in 1951 made such licenses a requirement for all new or expanded industrial investment, and an industrial policy resolution passed in 1956 stated the government's intention to use licensing as a means to achieve balanced regional development (Sekhar 1983, p. 15; Agarwal 1976, chap. 27). In 1977 it was decided that industrial licenses would be denied in large metropolitan areas and urban agglomerations (Sekhar 1983, p. 21).[4]

Direct investment in government-owned enterprises is a second means through which the government of India has tried to restrict the growth and size of large cities. Since 1954 most heavy industry has had to be government-owned (Sekhar 1983, p. 22; Agarwal 1976, chap. 28). Successive five-year plans have stated that preference in locations of government-owned industry was to be given to low-income areas, rural areas, or small cities and towns.

A third industrial location policy pursued by the government of India has been much discussed. It has long been national policy to equalize among regions the delivered prices (net of state sales taxes) of such basic products as cement, steel, and coal (Sekhar 1983, p. 31; Agarwal 1976, chap. 28). This policy has clearly led to excessive consumption of the commodities in regions distant from the places of production and deficient consumption in regions close to them. It also leads to excessive transport of the commodities, because users have no incentive to buy from the nearest source. The effects on city sizes are less clear. Most of the regulated products are resource-based, so the production sites tend to be far from major population centers. As a result, such commodities are likely to be underpriced in large urban areas, which presumably leads to the further expansion of already large cities. Government freight-pooling policies, however, also influence the location of production sites. Since location affects both production and consumption, it is difficult to conjecture the net effect on city sizes of the freight-pooling policy.

A fourth locational program is government encouragement of small-scale industry (Sekhar 1983, pp. 37–43; Agarwal 1976, chap. 25). Because small-scale industry in India is not a major source of employment and is located for the most part in small towns and rural areas, encouragement of small-scale industry inevitably promotes the development of activities not normally located in large cities. In addition, an explicit objective of the Indian government program is to encourage small-scale industry to locate in small towns and rural areas. A large number of specific instruments is used in this program: exemption from licensing requirements, prohibition of production of specific products in factories, concessionary finance, input allocations, training programs, technical assistance, marketing assistance, and industrial estate development. Some parts of the program date from the early 1950s, whereas other parts are of much more recent origin.

A fifth government of India locational program has been the creation of industrial estates (Sekhar 1983, pp. 43–49). An industrial estate consists of infrastructure and other capital provided by the government, usually for small industrial enterprises. Since 1955 the government has established industrial estates with two goals in mind: encouragement of small-scale industry and dispersal of industry from metropolitan areas to small towns and rural areas. The program has by no means had an entirely rural orientation, but it has been government policy not to build industrial estates in large cities.

The sixth locational program to be discussed in this section is metropolitan planning (Sekhar 1983, pp. 56–58). Master plans have been made for all large Indian cities and several large metropolitan regions. Most plans attempt to restrict the growth of manufacturing employment within their own jurisdiction. The effect of such plans is to induce manufacturing production to locate on the outskirts of urban agglomerations. Most large-scale manufacturing plants would probably locate there in any case, but the location of small-scale manufacturing sites may be influenced by such plans.

The final locational program to be discussed is a program to develop backward districts. The definition of backward districts has been expanded gradually since the program was started in 1970. Today it encompasses districts containing about 60 percent of the population. Among the tools employed in the program are subsidies for capital investment and transport, income tax concessions, and concessional loans (Sekhar 1983, pp. 59–68). State governments also provide incentives for industries to locate in backward districts. Definitions of backward districts and of eligible industries vary from one component of the program to another.

Thus, India's national and state governments have employed a wide variety of programs to limit the size and growth of large cities. Most of the programs have been intended to disperse industry from large cities to small towns and rural areas or to poor states and districts.

What effects have these programs had on relative city sizes? By the standards of verification preferred by modern social scientists the answer must be that no one knows or could know the answer. There is no usable theoretical model of the city size distribution—no model that could be estimated with data and used to predict changes in relative city sizes in the absence of government intervention, and then compared with actual Indian data.[5] All that is known, as was pointed out in chapter 4, is that the international tendency is for city sizes to follow an increasingly even distribution as the economy develops. Nevertheless, several considerations can be brought to bear that will create a presumption about the likely effects of Indian government policies on city sizes.

First, the programs have focused almost exclusively on manufacturing employment, which accounts for less than 10 percent of the nation's total employment and less than 30 percent of its urban employment. Most of the programs have been even narrower in scope, focusing on factory employment, which is less than 6 percent of total employment and less than 25 percent of urban employment. Large effects would be needed in these relatively small sectors to have much effect on overall city sizes. It is unclear whether the government programs have affected manufacturing employment locations by city size. Sekhar (1983) concludes that the government programs to disperse industry have had relatively little effect on manufacturing location. Data presented at the end of chapter 4, however, suggested that government programs might have dispersed manufacturing from the largest manufacturing centers. But the analysis of the determinants of city growth in chapter 4 also suggested that manufacturing employment is not a major determinant of city size in India. If this is true, modest effects of government programs to disperse manufacturing would have little effect on the size of cities.

Second, programs have been introduced at a variety of times since independence. Goals have been changed as programs developed. Programs have operated with different and sometimes conflicting goals. In some cases programs seem to have been employed more for other, perhaps higher priority goals than industrial dispersal. The goal of dispersal has frequently, and justifiably, been subservient to the goal of establishing viable government and private industries.

Third, the data analyzed in chapter 4 suggest, but do not prove, that government has had little effect on relative city sizes. During the first six decades of this century, Indian cities became irregularly but substantially more even in relative size. Since 1961, the period during which government programs could have had significant effects on relative city sizes, the cities have become slightly but steadily more unequal. There are many ways to measure relative city sizes, and other measures may yield different conclusions. At most, our measures show that government programs have not made relative city sizes more nearly equal. It cannot be known what would have happened without the programs, but the data strongly suggest that government programs have had no more than a minimal effect on relative city sizes. These considerations create a presumption that Indian government policies to disperse industry have had little or no effect on relative city sizes.

Analysis of Distortions in City Size Distributions

Why should governments intervene to alter the city size distribution? The starting point for this discussion must be the basic theorem of welfare economics.[6] The theorem states that if consumers maximize well-behaved preference functions and firms maximize well-behaved profit functions, both at fixed input and output prices, then prices that equate supply and demand provide a socially efficient allocation of resources.[7] Absence of scale economies, externalities, and other sources of market failure must also be assumed.

The theorem is no less applicable in a spatial than a nonspatial context. In a formal analysis it would be necessary to distinguish among commodities produced in different locations, and a variable for transportation to move commodities from the place of production to the place of consumption would have to be introduced into the model. Under the assumptions indicated in the previous paragraph, firms and households locate where profit and utility levels are as large as possible. The aggregate of such location decisions produces a socially efficient set of equilibrium locations. In particular, the size and spatial distributions of cities are socially efficient in equilibrium. Of course, in a changing economy neither locations nor production and consumption are in complete equilibrium. But markets provide powerful incentives to move toward equilibrium. There is no reason to believe that government intervention can accelerate the process of adjustment to disequilibrium. The reverse is more typical: governments slow down needed adjustments. Thus, provided that equilibrium is itself socially efficient, the fact that the economy is out of equilibrium is not a justification for government intervention in the adjustment process.

There are, of course, several reasons the conditions of this theorem are not satisfied in practice. Those reasons are the subject of this section. Before turning to them, however, we should clear away some misunderstandings concerning the conditions that might or might not violate the theorem about socially efficient resource allocation through decentralized decisionmaking.

First, the fact that Indian governments participate in production does not mean that the theorem breaks down. In India, much of the heavy industry is government-owned, and government participation in factory investment through loans and in other ways is substantial. An important motive for government participation is to avoid the distortions in production, pricing, and location that are thought to result from private monopoly power and other market defects. To that extent, government participation brings about the same production and location patterns as would efficient competitive markets. The fact that government participates in industrial production and location decisions creates no presumption that the normal criteria for efficiency are not employed. In practice, governments are guided by a variety of motives in choosing the locations of production facilities. It is not known to what extent the results differ from those that would be produced by decisions based on economic efficiency.

Second, the relative prices of some products differ in urban areas, especially large cities, from those found elsewhere. The larger the city, the higher the land values and, consequently, the price of housing. The result is that housing space per resident is smaller. People substitute high- for low-density housing and other goods and services for housing (Gupta 1982). That is a natural consequence in large cities. It does not imply that welfare levels are lower in large cities, and it does not imply market failure. High land values are the normal consequence of high productivity levels in large cities, and they provide an important deterrent to excessive movement of people and jobs to large cities.

Many ideas have been put forward to explain why the city size distribution may not be socially efficient and, in particular, why large cities may be too large. Most such claims, and, we believe, all the important ones, can be grouped in three categories: the disamenities of large cities, the effects of government programs whose locational effects are incidental, and the misallocation of infrastructure investment. Each will be analyzed in turn. Equity issues related to city sizes will be discussed in the subsequent section.

Disamenities of large cities

Large cities are thought to generate many disamenities. Pollution and congestion are the most important. Our analysis is conducted in the context of pollution, but it applies with minor modifications to any disamenity that is positively correlated with city size.

The air and water environments can absorb limited amounts of wastes without deteriorating significantly. If discharges are heavy, however, the environment deteriorates and people's health and welfare suffer. Discharges tend to increase in total and per unit of land area as a function of city size; thus large cities have lower-quality environments than do small cities. Air pollution is probably not a serious problem in most Indian cities, but water pollution is a serious health hazard. Not only are more people exposed to pollution in a large than in a small urban area, but also the larger the city, the lower and more health-impairing is the environmental quality likely to be. Since normal market forces are not adequate to deter private and government-owned firms from discharging pollutants, government intervention is necessary to prevent excessive environmental damage.

It follows from the analysis in the previous paragraph that people's welfare would improve if production facilities were moved from large to small cities. Indeed, if polluting discharges per unit of production were technically fixed, government programs to limit the sites of polluting activities in large cities would be justified.

But the noxiousness of industrial discharges depends not only on the volume of wastes discharged, but also on the organization of production. Organic wastes can be treated before discharge, waste materials can be reused instead of discharged, clean fuels can be substituted for dirty fuels, and so on. The best way to reduce pollution is for governments to mount programs that induce or force the offending parties to reduce their discharges. Or governments could reduce discharges directly by means of government-owned treatment facilities and other methods. Discharge standards may have to be stricter in large than in small cities. Such pollution abatement programs would almost certainly be an inducement to reorganize production to reduce polluting discharges per unit of output. The greater stringency of discharge standards in large cities, however, might induce some movement of activities from large to small cities, especially if the costs of abatement are levied on producers and consumers. Such movement would be incidental and probably of minor magnitude. The main point is that the environmental quality would be improved in both large and small cities.

To control pollution by moving activities from large to small cities is to start at the wrong end of the problem. The main effect would be to move pollution from large to small cities, not to improve the environment overall. To set standards for polluting discharges would pervasively improve environmental quality and would have only minor effects on city sizes.

The moral of this analysis is that government programs to abate pollution should concentrate on abating pollution, not on altering the sizes of cities through controls on industrial location and other activities. This simple moral applies to all disamenities. For example, the best way to deal with urban congestion is to provide better transportation management, not to limit city size. As with environmental quality, congestion is more serious in large than in small cities, so transportation managers may have to make transportation more expensive in the large cities. That may have some effect in reducing the sizes of large relative to small cities, but probably not much. The main effect would be a more efficient transportation system in both large and small cities. There is no perfect or easy way to reduce congestion or pollution or any other disamenity, but programs that attack the disamenities have a much better chance of succeeding than do programs to control city sizes.

Counteracting the effects of other government programs

Governments undertake many programs that may have incidental though predictable effects on city size. Three examples will be presented here to illustrate the issues involved and appropriate ways of dealing with them.

First, governments of many countries protect their manufacturing industries from foreign competition with tariffs, quotas, and other measures. Since manufacturing tends to be concentrated in large cities, some people claim that the resulting growth of the manufacturing sector makes large cities too large and that this justifies the imposition of controls to ensure that a proportion of manufacturing locates elsewhere. But manufacturing locates in large cities only if that is the most efficient location. If it is worthwhile to stimulate manufacturing growth by applying import controls, it is worthwhile to permit manufacturing to locate where it would be most efficient. Otherwise, the protection will not lead to the largest possible increment in manufacturing production and employment. It is not good government policy to stimulate domestic manufacturing with import controls and then prevent it from locating where it can generate the most jobs and maximum production.

Minimum wages are a second example. In many countries governments stipulate wage rates below which it is illegal to employ workers or certain categories of workers. Inevitably, in the absence of minimum wage controls, money wages are higher in large than in small cities. Land values and thus living costs are higher in large than in small cities, so workers must be paid more to induce them to live and work in large cities. In many countries real wages are also higher in large than in small cities, but that is not relevant to the present analysis. The fact that money wages are higher in large than in small cities means that it is easier to pay the legal minimum wages in large cities. Thus, low-productivity workers may be able to find jobs at the minimum wage in large cities but not in small cities. This, it is claimed, causes large cities to be larger than otherwise.

The situation does not justify controls on large cities. Low-productivity workers who find jobs in large cities might otherwise have been unemployed in small cities, in which case the increased size of large cities that results from the minimum wage law is a desirable result and should not be undone by other programs. If, however, it is believed that unemployment caused by the minimum wage law is not important, then the minimum wage should perhaps be set higher in large cities. This would prevent the distortion of city sizes through the effects of the minimum wage law. In either case, programs to control the size of large cities are unjustified.

The third example is rent control, which is employed in India and in many other countries. If rents are controlled at levels above competitive equilibrium, the control has no effect on the housing supply. If they are controlled below competitive equilibrium, the supply of housing will be reduced unless government makes up the deficiency. If rents are regulated at the same below-competitive levels for housing of given types regardless of city size, the controls will keep down the size of large cities more than of small cities, since competitive housing prices are higher in large cities. If controls permit rents high enough to cover costs when a dwelling is new, but restrict increases to uniform percentage rates regardless of city size, they curtail relative growth rates of large cities only to the extent that competitive rents rise faster in large cities. Once again the effects on the relative sizes of cities are incidental, and the controls should be judged on their own merits.

As a general rule, then, programs whose natural effect is to increase the relative sizes of large cities should be judged on their merits. If a program is desirable, so is the resulting effect on city size. Alternatively, modification of a program can prevent an undesirable increase in the size of large cities relative to others.

Infrastructure and local government services

"Infrastructure" refers to government-owned capital used to produce government services. Such services are important to people and firms; the capital is one among several inputs. The services provided include transportation, water supply, waste disposal, education, health, and sometimes gas and electricity.

The fact that such services are important to people implies that provision of large amounts of high-quality local government services attracts people and firms to a city, whereas provision of only a few low-quality services deters people and firms from locating there. High-quality government services are neither necessary nor sufficient for growth, however. Industrial estates and related infrastructure have been provided in some small towns and rural areas in India and have remained unused for years. At the same time, some Indian cities with only minimal local government services have experienced considerable growth. But no one would deny that government service provision tends to cause growth, especially when comparisons are made among cities that are otherwise similar.[8]

Such local government service provision tends to be extremely controversial. Some people claim that most local government services would be unnecessary in the absence of urbanization. That is partially true for certain services: water supply and waste disposal can be provided much more cheaply in rural than urban areas if rural densities are low enough so that natural systems can be used to deliver water and carry wastes away. But urbanization does not create the need for education, health services, and power supply. These services are required for economic growth and human welfare. Urbanization generates population and employment densities that make feasible the provision of most local services for ordinary people (Linn 1982).

More relevant to the issues in this chapter is the claim that large cities become too large if local government services are excessive in relation to those provided in smaller cities. That claim seems indisputable. In principle, infrastructure investments and government services should be provided where benefits are greatest relative to costs. Much effort is devoted to estimating the benefits and costs of individual projects, and most estimates are based on the assumption that the population in the city where the project is planned will grow at a rate that is not affected by the project.

To estimate effects of infrastructure investment on city growth it would be necessary to have estimates of the stock of infrastructure capital (and preferably of its components) for a sample of cities so that the infrastructure could be related to the cities' growth rates. This issue was discussed in chapter 5, where preliminary data for such an estimation were presented. As with the amenity issues discussed above, the appropriate government response to misallocation of infrastructure is to improve the allocation procedures, not to control city sizes.

In India, as elsewhere, the larger the city, the higher the local government expenditures per capita. But it is unclear whether this is attributable to the more extensive provision of services or to the higher unit cost of services provided (Datta 1981).

Although the Indian data needed to ascertain whether the amounts spent by large or small cities on local public services are equitable or socially efficient are not yet publicly available, an important procedural matter can be analyzed. As has already been pointed out, studies analyzing cross-country differences in city size distributions have concluded that larger, higher-income countries have more nearly even distributions. But other things being equal, countries with federal systems of government have much more nearly even distributions than do countries with unitary systems (Henderson 1982; Wheaton and Shishido 1981). In federal systems, states or provinces are constitutionally empowered to levy taxes to pay for the provision of local government services judged to be in the interests of local residents. In unitary systems, however, national governments set most of the tax levels, although some taxes may be collected by state or provincial or local governments. The central governments also make most spending decisions, and they specify within narrow bounds the kinds and amounts of expenditures local governments can make.

In unitary systems, there is an irresistible tendency to pour money into infrastructure and local government services in the national capital, to the neglect of other cities. The result is that more people and businesses flock to the national capital. In federal systems, local governments compete for residents and firms to some extent and can tailor local taxes and expenditures to satisfy the needs and wishes of residents. As a result, many cities develop and the size distribution is dispersed.

It is important not to overstate this proposition. Considerable amounts of local government revenues are provided by national governments even in federal systems. But the fact that India, the United States, and Australia have federal systems probably has a lot to do with the dispersed distribution of city sizes in those countries. Unitary governments, in contrast, try to undo the urban concentrations they have created in their capitals by imposing rules that make it difficult for people and firms to locate there.

The foregoing does not settle the issue of whether local government service provision is excessively concentrated in large Indian cities. Overinvestment in large cities is certainly possible, even in a federal system. But given the autonomy of India's state and local governments and India's dispersed system of city sizes, it seems unlikely that infrastructure investment has caused much distortion.

In India, as elsewhere in recent years, an increasing fraction of local government revenues has come from grants from the national government. If the trend continues, the ways in which local governments function will change radically. In particular, the link between local-government-provided services and the wishes and tax-paying ability of local residents will be loosened. The test of local government officials will increasingly become their ability to obtain money from Delhi rather than their ability to propose tax and spending programs that please local residents. The experience of other countries suggests, but does not prove, that one result will be a greater concentration of government spending in Delhi and perhaps in other large cities, unless there is a determined effort to decentralize spending.

Equity Issues

The analysis in the preceding section strongly suggested that the relative sizes of India's large cities should not be an important government policy issue. There is not much reason to believe that government programs to reduce the size of India's large cities would significantly improve social efficiency or living standards. The argument until this point has been couched in terms of social efficiency and economic growth. Many government and private documents, however, claim that equity should be an important reason for restricting the size and growth of large cities.

The argument is straightforward. Industrial jobs and other employment generated by industrial production pay good wages in India and other developing countries. Incomes are higher in states and cities with large manufacturing sectors than they are elsewhere. Therefore, it is claimed, efforts should be made to locate manufacturing in small towns, rural areas, and low-income states and districts so as to raise incomes in such areas.

Much has been written about India's income distribution and actual and proposed government policies to alter the distribution. We will review the literature in the next chapter. Here the issue is much narrower: whether restrictions to curb the growth of large cities are an effective way to increase the incomes of poor people. Unfortunately, none of the fine discussions of India's redistribution policies seems to have addressed the redistribution of city sizes.[9] Even in discussions of policies to raise low rural incomes, the movement of industry from large cities to rural areas is hardly mentioned as a realistic possibility.[10] Thus, in this section we can do little more than identify issues and indicate possibilities.

In all countries industry thrives better in some regions than in others and better in cities of a certain size than in others. Industrialization always entails large migrations of people to areas of industrial growth. Furthermore, large sums of money have been wasted in many countries on efforts to locate industry where it cannot survive or develop.

These conditions and dangers are no less applicable to India than to other countries. The great danger in India is that desperately needed industrialization will be slowed because of government insistence that industry locate where it cannot survive or expand, at least not without government controls that force consumers to pay excessive prices. It is quite possible that the widely discussed decreases in Indian industrial capital productivity during the 1960s and 1970s resulted in part from government attempts to disperse industrial investment from large cities (Wolf 1978).

The danger is especially great with attempts to limit the growth of large cities. Large cities offer very special advantages of market size, labor and other input supplies, port and other transportation facilities, and so forth. Government policies court disaster if they attempt to force industries that would normally locate in large cities to locate in small towns or rural areas instead. Some Indian government programs are designed to discourage industrial location in all cities with more than a million inhabitants, or even in all Class I cities. That cannot possibly be sound policy.

It would be much more prudent to attempt to shift the location of industry only among cities differing much less in size—among cities that differ by no more than, say, a factor of two. Of course, size is by no means the only characteristic of a city that is important to successful industrial investment. Government policy must attempt to locate industries only in cities that satisfy the most requirements for success, including size. Whether government programs can be designed with the needed complexity and flexibility is an open question. It is also an open question whether the resulting equity benefits would be substantial. There are no Indian data on income distribution by city size. Nor are there data on the effect on income distribution of attempts to locate industries in cities somewhat smaller than those in which they would normally locate in the absence of government location policies.

Policy content is also important. Policies that motivate location are much better than policies that mandate location. Examples of the first are concessionary loans for investment in small cities and somewhat higher taxes in large cities. Licensing is the prime example of the second category.

Policies that affect locational incentives are better for two reasons. First, although mistakes are likely to entail wasted money, they are unlikely to curtail industrial growth. If concessionary loans are not taken up because the concessions are not great enough, or if they are taken up even though the concession is greater than needed, then justifiable investments occur, whether or not the program has worked as well as it could have. But a license that is usable only in a small town may not be taken up, and the investment may not be made. Second, concessionary policies facilitate evaluation. It is easy to keep track of concessions used to affect industrial location, and it is possible to estimate how much investment was affected. But it is impossible to know how much investment is forgone because licenses are not given for locations in large cities.

Bad policies are costly and good policies are beneficial. The point of the previous paragraph is that inducements are quantitative and can be ignored, whereas prohibitions are absolute and cannot be ignored.

State or national government subsidization of infrastructure is one possible inducement to location in cities of moderate size. Infrastructure subsidies should focus both on services to industry and on services to residents. Given the scarcity of infrastructure in most Indian cities, the money would hardly be wasted even if little industry was thereby induced to locate in the subsidized cities.

Our judgment is that concessional finance for industry and infrastructure subsidization is the best program to induce growth of moderate-size, as opposed to large, cities. We have grave doubts that social efficiency is enhanced by any program that places direct controls on the growth of large cities. Subsidization of industrial growth in selected medium-size cities might entail modest redistribution of investment from large cities and might entail gains in social efficiency. It is unlikely that programs to disperse industry from large cities would have measurable effects on income distribution. The main effect is likely to be a reduction in the amount and productivity of investment.

Conclusions

A variety of government programs to disperse industry from large cities has been employed in India since independence. It is unclear whether such programs have had much effect on manufacturing locations, but it is unlikely that they have had measurable effects on the relative sizes of cities.

Many reasons are given for the belief that government programs to reduce the relative sizes of large cities will improve equity or social efficiency. Some are based on misunderstandings about the characteristics of large cities and about the symptoms of inefficiency in resource allocation by city size. Pollution, congestion, and other disamenities are certainly worse in large than in small cities, but they should be remedied by programs to improve the organization of production to reduce the disamenities, not by controls on city size. Infrastructure investment can cause distortions of city sizes, but it is unlikely that it has been a significant problem in Indian cities. It is also unlikely that programs to disperse industry from large cities can significantly improve personal income distribution.

References

Agarwal, A. N. 1976. *Indian Economy.* New Delhi: Vikas Publishing House.

Dandekar, V. M., and N. Rath. 1971. *Poverty in India.* Poona: Indian School of Political Economy.

Datta, Abhijit. 1981. "Municipal Finances in India." Washington, D.C.: World Bank, Urban and Regional Economics Division. Processed.

Debreu, Gerard. 1959. *Theory of Value.* New York: John Wiley.

Gupta, Devendra. 1982. "Some Aspects of Urban Housing in India." Washington, D.C.: World Bank, Urban and Regional Economics Division. Processed.

Henderson, J. Vernon. 1982. "The Impact of Government Policies on Urban Concentration." *Journal of Urban Economics* 12, no. 3 (November): 280–303.

Just, Richard, Darrel Hueth, and Andrew Schmitz. 1982. *Applied Welfare Economics and Public Policy.* Englewood Cliffs, N.J.: Prentice-Hall.

Linn, Johannes. 1982. "The Costs of Urbanization in Developing Countries." *Economic Development and Cultural Change* 30 (April): 625–48.

Renaud, Bertrand M. 1981. *National Urbanization Policy in Developing Countries.* New York: Oxford University Press.

Sekhar, Uday. 1983. *Industrial Location Policy: The Indian Experience.* World Bank Staff Working Paper no. 620. Washington, D.C.

Swanson, Joseph, Kenneth Smith, and Jeffrey Williamson. 1974. "The Size Distribution of Cities and Optimal City Size." *Journal of Urban Economics* 4:395–409.

Tolley, George, Philip Graves, and John Gardener. 1979. *Urban Growth in a Market Economy.* New York: Academic Press.

Wheaton, William C., and Hisanobu Shishido. 1981. "Urban Concentration, Agglomeration Economies, and the Level of Economic Development." *Economic Development and Cultural Change* no. 30 (October): 17–30.

Wolf, Martin. 1978. "Capital and Growth in India: 1950–71." In *India: Occasional Papers.* World Bank Staff Working Paper no. 279. Washington, D.C.

Notes

1. The best book describing such programs is Datta (1981). For a fine study of urbanization policies in developing countries, see Renaud (1981).
2. The best such study is Tolley, Graves, and Gardener (1979).
3. This section relies heavily on Sekhar (1983).
4. Various exceptions from license requirements have been established in recent years.
5. Empirical studies, with no normative content, that regress indices of relative city sizes on measures of economic development and related variables in a cross-country analysis invariably obtain R^2's that are less than 0.25 (See Linn 1982). See Renaud (1981) for a fine discussion of normative issues related to city sizes.
6. The definitive statement is Debreu (1959). A fine recent exposition is Just, Hueth, and Schmitz (1982).
7. A socially efficient allocation of resources is such that no other allocation can make anyone better off without making others worse off.
8. On the subject of this subsection see Linn (1982) and Swanson, Smith, and Williamson (1974).
9. See Agarwal (1976) and Dandekar and Rath (1971), for example.
10. Many papers on this subject appear in *Economic and Political Weekly.*

8

Income Distribution in India

THE DISTRIBUTION OF INCOME in India has been the subject of numerous in-depth studies (see Kumar 1974 for a survey). Statistics on India abound and are probably more detailed than those for any other developing economy. In this chapter we examine the salient features of India's income distribution, concentrating on the relationship between urbanization and inequality. We then consider some of the determinants of interstate inequality.

Recent Trends

There is no data set from which a definitive index of income distribution in India can be derived. Problems of data availability and reliability make assessment of India's personal income distribution difficult. Indian censuses do not include household income data. Sample surveys, however, are conducted by research institutions. The most recent survey for which published results are available was conducted by the National Council of Applied Economic Research (NCAER) for the agricultural year 1975–76 (see NCAER 1980). The survey covered 5,125 households—roughly 0.005 percent of all households in the country—which were scientifically chosen from 237 rural villages and 150 cities and towns. Results from this nationwide survey

Note: This chapter has benefited from comments by John R. Hansen, Koichi Mera, and Felipe Morris. We are grateful to Lynne Sherburne for her assistance with the computations and to William Knight for a literature search.

Table 8-1. *Household Income by Decile,* NCAER *and Myrdal Surveys*
(percent of total income)

	NCAER			Myrdal	
Decile[a]	*1975–76*	*1967–68*	*1964–65*	*1960*	*1953–57*
All households (urban and rural)					
0–10	2.27	1.8	2.68	1.1	3.1
11–20	3.52	2.9	3.92	3.0	4.8
21–30	4.45	3.7	4.88	4.2	5.7
31–40	5.47	4.7	5.83	5.3	6.6
41–50	6.49	5.8	6.72	6.4	7.5
51–60	7.71	7.2	7.65	7.7	8.6
61–70	9.28	9.0	8.75	9.2	9.8
71–80	11.49	11.8	10.37	11.4	11.5
81–90	15.44	16.4	14.09	15.0	14.2
91–100	33.88	36.7	35.11	36.7	28.2
Gini coefficient	0.42	0.48	0.41	0.47	0.34
Urban households					
0–10	2.26	2.0	2.26	1.3	2.9
11–20	3.56	3.2	3.26	2.7	4.5
21–30	4.47	4.2	4.13	3.5	5.3
31–40	5.36	5.0	4.83	4.5	5.9
41–50	6.43	6.1	5.83	5.4	6.7
51–60	7.57	7.3	7.02	6.7	7.5
61–70	9.37	8.9	8.71	8.3	8.5
71–80	11.53	11.2	9.97	10.5	9.8
81–90	15.85	15.0	14.11	14.2	11.8
91–100	33.60	37.1	39.88	42.9	37.1
Gini coefficient	0.42	0.46	0.46	0.53	0.42
Rural households					
0–10	2.50	1.9	3.01	0.8	4.0
11–20	3.81	2.8	4.38	3.1	5.1
21–30	4.81	3.7	5.39	4.4	6.0
31–40	5.82	4.6	6.40	5.5	6.7
41–50	6.90	5.8	7.42	6.8	7.6
51–60	8.11	7.2	8.34	8.1	8.7
61–70	9.62	9.1	9.56	9.8	9.9
71–80	11.78	11.8	10.85	11.9	11.6
81–90	15.12	16.6	13.75	15.5	14.4
91–100	31.53	36.5	30.90	34.1	26.0
Gini coefficient	0.39	0.48	0.35	0.45	0.31

a. Arranged from lowest to highest income.
Sources: Jain (1975) and Bardhan (1974).

on the household income distribution, as well as results from earlier surveys by the NCAER and by Gunnar Myrdal, are presented in table 8-1.

No uninterrupted trends in the national income distribution can be observed in table 8-1. There are too few comparable data sets for a reliable time series on the household income distribution.

Although the NCAER is the only direct source of income estimates, other studies exist that are based in part on consumption expenditure data from the National Sample Survey (NSS). The shortcomings of both the NCAER and NSS surveys are given detailed attention in Bardhan (1974) and need not be repeated here. Bardhan also discusses the procedures some authors have used in generating income distribution estimates from consumption data. Summary statistics from some of the studies that use NSS data appear in table 8-2.

If it is assumed that trends in per capita household income, household income, personal income, and other units over which distributions are

Table 8-2. *Income Distribution Parameters Generated from* NSS *Surveys*

Author and year of survey	*Income distribution unit*	*Inequality measure (Gini coefficient, except as noted)*		
		Urban areas	*Rural areas*	*All-India*
Ahmed and Bhattacharya				
1956–57	Household per	0.399	0.326	n.a.
1960–61	capita expen-	0.350	0.321	n.a.
1963–64	ditures	0.360	0.297	n.a.
Ahmed and Bhattacharya				
1956–57	Personal	n.a.	n.a.	0.408
1960–61	pretax	n.a.	n.a.	0.382
1963–64	income	n.a.	n.a.	0.361
Iyengar and Jain				
1961–62	Per capita	n.a.	n.a.	0.399
1964–65	household income	n.a.	n.a.	0.334
Ojha and Bhatt				
1953–55	Personal	0.40	0.34	0.38
1963–65	income	0.46	0.32	0.38
Jain				
1961–62	Per capita	0.493[a]	0.437[a]	n.a.
1963–64	household	0.747[a]	0.346[a]	n.a.
1964–65	expenditures	0.511[a]	0.305[a]	n.a.
Joshi				
1972–73	Personal expenditures	0.657[b]	0.425[b]	n.a.

Table 8-2 (*continued*)

Author and year of survey	Income distribution unit	Inequality measure (Gini coefficient, except as noted): Urban areas	Rural area	All-India
Chatterjee and Bhattacharya				
April–June, 1951	Per capita household consumer expenditures	n.a.	0.362	n.a.
August–November, 1951		0.384	0.334	n.a.
1952		0.365	0.340	n.a.
1952–53		0.397	0.330	n.a.
1953		0.364	0.330	n.a.
1953–54		0.371	0.334	n.a.
1954–55		0.390	0.350	n.a.
1955		0.371	0.335	n.a.
1955–56		0.368	0.344	n.a.
1956–57		0.402	0.319	n.a.
1957		0.394	0.332	n.a.
1957–58		0.359	0.334	n.a.
1958–59		0.348	0.340	n.a.
1959–60		0.357	0.314	n.a.
1960–61		0.350	0.321	n.a.
1961–62		0.357	0.312	n.a.
1963–64		0.360	0.297	n.a.
1964–65		0.349	0.294	n.a.
1967–68		0.345	0.293	n.a.

n.a. Not available.

a. σ: log variance estimated from a displaced lognormal distribution.

b. σ.

Sources: Srinivasan and Bardhan (1974) for Ahmed and Bhattacharya, Ojha and Bhatt, and Chatterjee and Bhattacharya; Iyengar and Jain (1974); Jain (1977); Joshi (1979).

defined move together, some patterns emerge. It is virtually certain that the income shares of the bottom groups increased during the early 1960s and again during the period 1967–68 to 1975–76. But this does not necessarily indicate a long-term trend toward greater income equality. The NCAER estimates reveal a rise in inequality between 1964–65 and 1967–68 that more than offset the subsequent decline in inequality. This rise is consistent with the substantial increase in rural inequality that Bhatty (1974) found in data from the NCAER surveys. Yet Chatterjee and Bhattacharya (1974) do not find a similar rise in urban or rural inequality for the mid-1960s in the NSS data. Nevertheless, comparison of the Jain (1977) and Joshi (1979) studies of urban and rural expenditure distributions indicates that inequality does not steadily decline in the NSS samples either.

The trend in income distribution during the late 1950s is also subject to dispute. Myrdal's figures (cited in Jain 1975) suggest that there was a rise in inequality, as do, implicitly (given the trend toward increased equality in the early 1960s), Ojha and Bhatt (1974). In contrast, there is no trend in the pattern of inequality indicated by the NSS surveys reported in Chatterjee and Bhattacharya. Furthermore, Ahmed and Bhattacharya's (1974) findings suggest a decline in overall inequality.

Clearly, the choice of initial and terminal dates makes a difference in an analysis of Indian inequality. By choosing years selectively, one can present evidence for increasing, decreasing, or stable inequality. The year-to-year variation in inequality statistics is large.

All this uncertainty about income distribution trends is reinforced by data problems. The sample sizes—especially in the NCAER surveys—are not large, given India's size and diversity.[1] Bardhan (1974) discusses other problems, including the undersampling of the rich by the NSS and the omission of homeless household groups from both the NCAER and NSS samples. Furthermore, much of the income imputed to rural households, such as that generated from animal husbandry and small-scale industry, is estimated arbitrarily. Mukherjee (1972, pp. 84–100) estimates that about 40 percent of India's net domestic product (NDP) is based on subjective valuations.

One must also make strong and somewhat arbitrary assumptions about savings to derive an income distribution from NSS studies of consumption expenditures. Many studies have analyzed the inference of income from expenditure data and the choice of functional form appropriate to the sample incomes or expenditures. Other statistical problems are that nonresponse rates have not been evaluated and there are inconsistencies between NSS figures and government production estimates. Moreover, if one tries to compare income distribution estimates over time, one is confronted by dissimilarities—in reference period, geographical coverage, and valuation procedures—between the NSS surveys before round 14 (1958–59) and those after. The fact that different researchers work with different units of observation and use different techniques for inferring income from NSS consumption surveys and for estimating a distribution from the sample further limits the comparability of their findings.

Despite these limitations two additional observations can be made. First, inequality is more pronounced within India's urban sector than within its rural sector. The conclusion is the same whether one examines summary inequality statistics, such as the Gini coefficient, or the shares of the lower deciles, and it holds true in every study based on NSS data and in all the NCAER surveys except 1967–68.

Second, trends in urban and rural income distribution mostly follow national trends. This is evident in table 8-1, although the urban Gini coefficient does not increase between 1964–65 and 1967–68, as do the rural and national Ginis. Nevertheless, the share of the bottom urban quintile declined from 5.5 to 5.2 percent during this period; the second, third, and fourth quintiles, however, gained at the expense of the top and bottom quintiles. Similarly, the NSS surveys indicate that urban inequality increased slightly during the period 1960–61 to 1963–64, whereas rural and national inequality declined substantially.

It should be noted that national inequality may increase even if both urban inequality and rural inequality decrease, if the percentage of the population living in the sector with greater inequality increases (Robinson 1976). This seems not to have happened in India.

We turn now to an examination of the Indian distribution patterns. We consider the relation between inequality and production trends, examine the effect of income and inequality patterns on the real earnings of the bottom groups, and compare Indian inequality with that found in other countries. The oscillating trends in Indian income distribution have been accompanied by slow, oscillating growth in per capita GNP and by even slower growth in per capita private consumption. As can be seen from table 8-3, the rise in the proportion of resources devoted to investment is primarily responsible for the relatively slow growth of private consumption.

In postindependence India, per capita GNP has grown at an average annual rate of 1.25 percent to 2 percent, depending on the choice of initial and terminal years. If one picks years with comparable rainfall indices (such as 1960–61 and 1977–78), the growth rate of per capita GNP tends to be roughly 1.5 percent.[2] By contrast, annual per capita consumption growth has averaged slightly less than 1 percent.

Table 8-4 shows the growth in average real household income by income class for three periods for which household income distribution surveys are available. The figures for 1960 are suspect, given the wide variation in income distribution estimates for that year; however, even if inequality is substantially less than estimated by Myrdal, the combination of oscillating degrees of inequality and slow but positive per capita income growth does not appear to have led to the impoverishment of the bottom groups. To the contrary, if 1975–76 is used as the terminal date, the data suggest that the bottom groups fared better than average in the 1960s and 1970s. Per household income grew by 67 percent for the bottom quintile; it rose by at least 23 percent for the second quintile and 19 percent for the middle quintile. In contrast, the top quintile enjoyed a rise in real per household income of only 12.6 percent.

Unfortunately, these findings probably convey a rosier picture than war-

Table 8-3. *Trends in Per Capita Production and Consumption, 1950–80*
(constant 1970–71 rupees)

Year	GNP (1)	Private consumption (2)	Private consumption ÷ total resources[a] (percent) (3)	General government consumption ÷ total resources[a] (percent) (4)	Gross domestic investment ÷ total resources[a] (percent) (5)	GDP (6)	Final private consumption[b] (7)	NNP (8)
1950–51	512	434	84.6	8.0	7.3	512	482	463
1955–56	563	431	76.4	13.9	9.6	564	521	503
1960–61	621	474	74.2	11.1	14.7	624	488	577
1965–66	652	471	70.3	12.4	17.3	656	489	556
1966–67	634	487	74.7	6.8	18.5	n.a.	n.a.	n.a.
1967–68	671	510	74.0	8.4	17.5	n.a.	n.a.	n.a.
1968–69	678	516	75.1	8.7	16.2	684	518	583
1969–70	702	516	72.7	10.2	17.1	n.a.	n.a.	n.a.
1970–71	734	537	72.3	9.3	18.5	735	545	625
1971–72	734	535	71.7	10.3	18.0	n.a.	n.a.	n.a.
1972–73	714	521	72.2	10.0	17.9	n.a.	n.a.	n.a.
1973–74	727	542	73.7	7.1	19.1	726	529	616
1974–75	713	511	70.7	8.7	20.6	n.a.	n.a.	n.a.
1975–76	763	518	66.9	13.1	20.0	n.a.	n.a.	n.a.
1976–77	760	528	69.2	9.5	21.3	764	535	653
1977–78	797	551	70.4	9.5	20.1	810	580	698
1978–79	n.a.	n.a.	n.a.	n.a.	n.a.	850	597	724
1979–80	n.a.	n.a.	n.a.	n.a.	n.a.	793	557	674

Annual growth rate (percent)				
1950–77	1.65	0.89	1.71	0.69
1950–78	n.a.	n.a.	1.83	0.77
1950–79	n.a.	n.a.	1.52	0.50
1950–55 to 1972–77	1.50	0.92	n.a.	n.a.
1960–77	1.48	0.89	1.55	1.02
1960–78	n.a.	n.a.	1.73	1.13
1960–79	n.a.	n.a.	1.27	0.70
1960 to 1972–77	1.27	0.75	n.a.	n.a.
1965 to 1972–77	1.43	1.22	n.a.	n.a.
1966 to 1972–77	1.93	0.97	n.a.	n.a.

n.a. Not available.

a. "Total resources" is defined as GDP plus net imports of goods and services.

b. "Final private consumption" equals disposable income less gross domestic savings, public consumption, and a statistical discrepancy.

Sources: Columns (1) through (5): World Bank (1980b, pp. 102–103); columns (6) through (8): World Bank data.

Table 8-4. *Average Real Per Household Income by Quintile*

Quintile	*Myrdal, 1960*		NCAER, *1964–65*		NCAER, *1967–68*		NCAER, *1975–76*		*Income growth rate (percent)*		
	Income share (percent)	*Rupees*	*Income share (percent)*	*Rupees*	*Income share (percent)*	*Rupees*	*Income share (percent)*	*Rupees*	*1960–61 to 1964–65*	*1964–65 to 1967–68*	*1967–68 t 1975–76*
Bottom	4.1	434	6.6	758	4.7	520	5.8	725	13.2	−11.8	4.2
Second	9.5	1,005	10.7	1,229	8.4	929	9.9	1,237	4.6	−8.9	3.6
Third	14.1	1,492	14.4	1,655	13.0	1,437	14.2	1,774	2.3	−4.6	2.7
Fourth	20.6	2,179	19.1	2,195	20.8	2,299	20.8	2,599	0.2	1.6	1.5
Top	51.7	5,470	49.2	5,653	53.1	5,870	49.3	6,160	0.7	1.3	0.6
Per capita NNP (1970–71 rupees)	—	557	—	605	—	582	—	658	1.9	−1.3	1.5
Per household income (1970–71 rupees)	—	2,116	—	2,298	—	2,211	—	2,499	—	—	—

— Not applicable.

Note: Rupee values were obtained as follows: We started with a series of NNP per capita figures in constant 1970–71 rupees. The ratio of pe household income (as given by the 1964–65 survey in 1964–65 rupees, which we converted to 1970–71 rupees by multiplying by the GDP deflato index, 1.445) to per capita NNP was assumed to be constant (and equal to 2,298 ÷ 605). We multiplied the per capita NNP figures for other year by this constant to obtain the estimates of per household income. Note that the choice of per capita income level does not affect the growth rat of per household income by quintile. Per household income was also available for the 1975–76 survey; the deflated estimate of 1975–76 per househol income was 2,890 rupees, 15.6 percent larger than the figure derived when the 1964–65 ratio of per household income to per capita income wa used as a constant.

Sources: Jain (1975, p. 50); NCAER (1980, p. 7); Bardhan (1974, p. 105); World Bank data.

ranted. First, 1975–76 was a year with an abnormally egalitarian income distribution, so the gains by low-income households that appear in table 8-4 may have been transitory. Indeed, the bottom quintile's income peaked in 1964–65, and even the second quintile had a 1975–76 per household income that was barely greater than its 1964–65 level. Moreover, our assumption of a constant ratio of per household income to per capita net national product (NNP) has probably biased the 1975–76 figures upward, since the share of business-retained earnings has almost certainly grown.

Chatterjee and Bhattacharya (1974), Bardhan (1974), and others have pointed out that use of a single deflator is likely to lead to biased results in intertemporal comparisons, since the appropriate price index for low-income households has risen more rapidly than the index for high-income households throughout much of the postindependence era. Thus, although the relative nominal income differentials may have narrowed (as between 1960–61 and 1964–65), the gaps in relative living standards may not have decreased much, if at all. The NDP deflator used in table 8-4 rose 72.6 percent between 1960–61 and 1967–68 (the figure given by Bardhan for the national income deflator is 70 percent). This is considerably less than the 96 percent rise in the agricultural laborer's consumer price index (CPI). (The CPI is comparable to a price index constructed by Bardhan from NSS rural expenditure data and to another based on wholesale price indices and NSS consumption weights.) Use of this higher price index reduces the bottom quintile's income gain (during the period 1960–61 to 1967–68) from 20 to 5.5 percent. What is more striking, this suggests that mild declines in real per household incomes are much greater for the second and middle quintiles (declines for these quintiles go from 8 to 4 percent and 19 to 15 percent, respectively).

Although Indian income distribution statistics reveal no time trends, it is conceivable that they are affected by patterns of economic performance. Lack of a consistent data set prevents us from undertaking a formal statistical analysis. Since regional income variance may be viewed as a component of all-India rural inequality, we expected to find a strong correlation between the standard deviation of regional rainfall levels and the all-India rural Gini coefficients; however, we were unable to uncover any association between the two time series.

We also expected that national agricultural fluctuations would be connected with the relatively low level of rural and national income inequality recorded in 1975–76. Indeed, 1975–76 had the highest rainfall index in the period 1960–61 to 1977–78 and was a year of record cereal yields (Ahluwalia and others 1978, p. 26; Sanderson and Roy 1977, pp. 24–25). Moreover, weather indices for three of the five preceding years were favorable too. The inegalitarian year 1967–68, however, also enjoyed moderately favorable weather, although the two years preceding it had been catastrophic. The trend toward increasing income equality in the early 1960s was not accompanied by a shift in weather patterns or in real per capita product in agriculture.

Unusually favorable weather patterns are likely to affect income distribution in two ways. First, in rural areas demands for productive factors will increase, and the effects on income distribution will depend on technology and on the operation of factor markets. If the surplus-labor model in which labor receives an institutionally fixed subsistence wage is accepted, one would expect that weather-related shocks would have the greatest impact on land and capital shares. Data for the 1960s indicate that this was indeed the case (Dholakia 1976, p. 310; Ahluwalia and others 1978, p. 24). Nevertheless, sharecropping arrangements may imply that the rainfall pattern has a distributionally neutral impact on per capita income, despite its strong effects on factor shares. We have found no evidence that the Gini coefficients reported in tables 8-1 and 8-2 for rural areas vary with weather patterns.

The second anticipated impact on national inequality is that income differences between the urban and rural sectors may become smaller, depending on technology and demand patterns. A rise in agricultural productivity could increase the differential, but this is unlikely. Government procurement policies and India's partial integration into the international food market prevent dramatic declines in food prices in unusually good crop years. Government price controls may also keep the prices of urban goods from rising as fast as they might otherwise. Consequently, in years of favorable weather, rural incomes are likely to grow more rapidly than urban incomes. Since rural incomes are smaller, the effect is to reduce inequality.

Inequality and Urbanization

Several studies have examined international trends in inequality (see, for example, Adelman and Morris 1973; Chenery and Syrquin 1975; Kuznets 1955). Ahluwalia (1976) presents perhaps the most complete analysis. Among the equations he estimates are the following:

$$\begin{aligned} (8.1)\qquad Q_1 &= \underset{(4.93)}{27.31} - \underset{(3.71)}{16.97y} + \underset{(3.74)}{3.06y^2} + \underset{(8.28)}{5.54S} \qquad R^2 = 0.54 \end{aligned}$$

$$\begin{aligned} (8.2)\qquad Q_1 &= \underset{(5.31)}{34.52} - \underset{(4.32)}{21.00y} + \underset{(3.91)}{3.32y^2} + \underset{(1.77)}{0.027P_u} \\ &\quad + \underset{(2.01)}{0.022\text{LIT}} - \underset{(1.44)}{0.336\dot{P}} + \underset{(6.10)}{4.685S} \qquad R^2 = 0.60 \end{aligned}$$

where

Q_1 = income share of the bottom quintile
y = log (base 10) of per capita GNP (1970 U.S. dollars)
S = dummy variable that indicates whether the country is socialist
P_u = percentage of the population living in urban areas
LIT = literacy rate
$\dot{P}$ = population growth rate.

The figures in parentheses are t-statistics. Sample comparability problems abound for several independent variables but particularly for the income distribution statistics.[3] Still, it is difficult to believe that the strong relationships found are entirely spurious.

Equations (8.1) and (8.2) both imply a U-shaped curve: that is, inequality initially worsens (Q_1 falls) as development proceeds and y rises, but eventually the positive quadratic term dominates and Q_1 rises with y. This U curve suggests that India may experience increasing income inequality as development proceeds, unless population growth rates systematically decline or the literacy rate and urbanization increase.

The estimated inequality curve can be used to place India's inequality in perspective. If a per capita income for India of $97.87 (in 1970 U.S. dollars) is used in equation (8.1), the predicted share of the bottom quintile is 5.66 percent. The predicted share rises to 5.84 percent if the coefficients estimated from a sample restricted to developing countries are used. When the 1970 Indian values of y, P_u, $\dot{P}$, and LIT are substituted in equation (8.2) the predicted share of the bottom quintile is 6.30 percent.[4] Ahluwalia employs several types of income recipient units in his sample of income distribution shares. This limits the comparability of predicted and actual shares, since some recipient units have systematically greater inequality than others. His observation for India's bottom quintile is 5.0 percent for 1963–65, which is substantially less than any of the predicted values.[5] This is the same as the 5 percent share recorded in Ojha and Bhatt (1974, p. 164) as the share of the bottom quintile in the distribution of personal income. However, the bottom quintile in the individual disposable income distribution gets 7 percent of disposable income; the bottom quintile in the individual expenditure distribution makes 10 percent of all expenditures. The comparable figures for 1953–55 are 7 percent and 10 percent, respectively.

The bottom quintile shares in the surveys in table 8-1 are 5.8 percent for 1975–76, 4.7 percent for 1967–68, 6.6 percent for 1964–65, 4.1 percent for 1960, and 7.9 percent for 1953–57. Ahmed and Bhattacharya (1974, table 10, p. 179) average two fits and report that the bottom quintile appears to receive a share comparable to that received by bottom groups in countries at a similar level of development. Thus, at times the bottom group fares worse and at other times better than its predicted share.

One can predict the trend in India's inequality from equation (8.2), though predictions of brief time series based on cross-country parameter estimates should be taken with more than a grain of salt. The predicted share of the bottom quintile is 6.92 percent, if 1960 values for y, P_u, $\dot{P}$, and LIT are used. The predicted 1970 share is only 6.30 percent, while the predicted share for the mid-1970s is 6.44 percent (World Bank 1980b and table 8-3). Compared with 1960, increased urbanization adds 0.05 percent to the bottom quintile's 1970 predicted share and 0.08 percent to the mid-1970s share. Corresponding increases in the literacy rate add 0.11 percent to the 1970 share and 0.18 percent to the mid-1970s share. Increases in the birth rate reduce the predicted 1970 share by 0.20 percent and the mid-1970s share by 0.03 percent. The U-curve effect generated by rising per capita income reduces the predicted share by 0.54 percent in 1970 and 0.70 percent in the mid-1970s. This U-curve effect dominates the trend in the predicted value of the bottom share between 1960 and 1970, but declines in the birth rate and increases in the literacy rate more than offset the U-curve effect from 1970 to the mid-1970s. Rather surprisingly, the predicted quintile share trends based on equation (8.2) mirror the trends in the three NCAER studies reported in table 8-1. The correspondence disappears, however, if the Myrdal figure for 1960–61 replaces the NCAER value for 1964–65. Although the correspondence is probably fortuitous, it draws attention to the need for further study of the link between Indian income distribution characteristics and variables such as fertility and literacy rates.

To conclude this section we compare India's income distribution statistics with those for countries at similar levels of development and for which methodologically similar surveys are available. The income distribution statistics in table 8-5 cover surveys with recipient unit definitions similar to those in the NCAER surveys reported in table 8-1. Most of the surveys are from the late 1960s, and Ahluwalia uses many of them in his U-curve estimates. While India may not be far off Ahluwalia's estimated curve, countries on or below the curve differ considerably from those with abnormally egalitarian distributions.

India's neighbors Pakistan, Bangladesh, and Sri Lanka enjoy more egalitarian distributions. Rapidly growing Thailand and Malawi have inequality levels comparable to India's. Korea has experienced a far more egalitarian income distribution than has India, but India's distribution is slightly more even than those of the Philippines, Malaysia, and Egypt. The income distribution is noticeably less unequal in India than in the poor countries of Latin America, such as Honduras, and of Africa, such as Tanzania and Sierra Leone. In Tanzania and Sierra Leone, however, distorted internal prices (specifically, low prices of agricultural products) are likely to make inequality appear far greater than if world prices were used in computing peasant incomes.

India emerges from the international comparisons as a country with an income distribution about typical for its stage of development. Countries typically viewed as experiencing successful growth with redistribution strategies are more egalitarian than India. The share of India's bottom quintile is similar to the corresponding shares in several countries undergoing major structural transformations. Unfortunately, India has not enjoyed rapid growth. Moreover, India's income distribution appears to fluctuate widely and does not display distinct time trends. In the absence of evidence that dramatic movements in relative prices have had particularly adverse impacts on the poor,[6] a corollary is that India's bottom classes share the fruits of growth, however meager they may be.

Factor Shares and the Sectoral Distribution of Income

In this and the following section we analyze forces that explain inequality in India. In principle, every force that affects relative factor and product prices alters the income distribution. Since the set of such variables is enormous, we restrict our discussion to major determinants and trends in a few key components of the Indian economy.

Determinants of inequality can be analyzed by dividing an aggregate measure into components. The log variance of the national income distribution, σ^2, can be shown to be a function of sectoral income distribution log variances σ_a^2 and σ_b^2, of the squared difference of the log means, y^2, and of the proportion of the population working in sector a, P_a (see Robinson 1976):

$$\sigma^2 = P_a \sigma_a^2 + (1 - P_a)\sigma_b^2 + P_a(1 - P_a)y^2. \tag{8.3}$$

First we divide India into agricultural (A) and nonagricultural (NA) sectors. Then equation (8.3) can be used to describe national inequality as an increasing function of sectoral inequality and of the intersectoral income gap. For constant values of σ_A, σ_{NA}, and y^2, aggregate inequality increases as P_A declines, until it reaches $(\sigma_A^2 - \sigma_{NA}^2 + y^2)/2y^2$.

The first trend to be considered is that of agricultural and nonagricultural shares in GDP. Agriculture accounted for 50.1 percent of GDP at factor cost in 1950–51; a decade later its share was 48 percent.[7] The slow decline continued during the 1960s: agriculture's share fell to 44.5 percent during the poor harvest of 1965–66, but recovered to 45.7 percent by 1970–71.

Table 8-5. *Bottom Quintile Income Shares in Selected Developing Countries*

Country	*Year*	*Income share of bottom quintile (percent)*	*Gini coefficient*	*Per capita* GNP[a] *(U.S. dollars)*	*Recipient unit*	*Original data source*
South Asia						
India	1964–65	6.8	0.41	192.2	Households	NCAER
	1967–68	4.7	0.48	192.6	Households	NCAER
	1975–76	5.8	0.42	217.6	Households	NCAER
Pakistan	1968–69	8.2	0.34	259.0	Households	Ministry of Finance, Planning and Development, Pakistan
	1968–69	12.9	0.21	259.0	Per capita household	Ministry of Finance, Planning and Development, Pakistan
	1966–67	7.6	0.36	243.0	Households	Central Statistical Office (CSO), Pakistan
	1966–67	11.9	0.22	243.0	Per capita household	CSO, Pakistan
Bangladesh	1966–67	7.9	0.34	119.5	Households	CSO, Pakistan
	1966–67	13.4	0.17	119.5	Per capita household	CSO, Pakistan
Sri Lanka	1969–70	6.9	0.38	212.5	Households	Department of Census and Statistics, Sri Lanka
Thailand	1962	5.7	0.51	318	Households	Harry Oshima, *Malayan Economic Review,* October 1970, p. 13

Other						
Egypt, Arab Rep. of	1964–65	4.6	0.43	339.7	Households	Institute of National Planning, Egypt
Honduras	1967–68	1.6	0.62	597.8	Households	Directorate of Statistics anc the Census, Honduras
Korea, Rep. of	1968	8.6	0.30	713.9	Households	Economic Planning Board, Korea
Malawi	1969	5.7	0.47	156	Households	Christian Morrison, Development Research Center, World Bank
Malaysia	1970	3.5	0.52	982.1	Households	Department of Statistics, Malaysia
Philippines	1965	3.6	0.47	493.7	Households	Bureau of Census and Statistics, Philippines
	1971	3.9	0.49	555.3	Households	Bureau of Census and Statistics, Philippines
Sierra Leone	1968–69	1.1	0.61	307.5	Households	Central Statistical Office, Sierra Leone
Tanzania	1967	5.1	0.50	247	Households	Christian Morrison, Development Research Center, World Bank
	1969	2.3	0.60	253	Households	Ministry of Economic Affairs and Development Planning, Tanzania

a. GNP per head (midyear population) in 1980 market prices.

Sources: All figures except per capita GNP are from Jain (1975); per capita income figures are from International Monetary Fund (1984). Survey with recipient units most similar to those in NCAER and NSS surveys were chosen.

Dramatic declines in agriculture's share, however, occurred in the mid- and late 1970s. It fell from 48.1 percent in 1973–74 to 38.5 percent three years later and to 33.0 percent in 1979–80. Constant rupee GDP series show an equally large but less abrupt decline in agriculture's share during the postindependence era. Relative price fluctuations have partially offset variations in agricultural yields, thereby making the variance of agricultural income small.

Several sectors gained at agriculture's expense. Registered manufacturing grew from 6 to 15 percent of GDP from 1950–51 to 1979–80. Small-scale unregistered manufacturing suffered declines in its share until 1970; in the 1970s its share rose modestly. Fledgling sectors such as electricity, gas, and water supply and banking and insurance have enjoyed rapid share increases. Increases have also been recorded by construction and other services. Unregistered manufacturing, public administration, and defense have suffered the strongest declines in the terms of trade.

Again with reference to equation (8.3), if inequality within the A and NA sectors did not change, one would expect to find an increase over time—particularly during the 1970s—in income inequality for two reasons. First, as table 1-5 in chapter 1 indicates, P_A has fallen in recent years (although definitional changes limit comparability), dropping from 69.5 percent in 1961 and 69.8 percent in 1971 to 66.7 percent in 1981. Second, because NA output relative to A output has grown more rapidly than relative labor force use, y^2 should also be growing (since per capita output in NA was larger initially than A's per capita output as well). As $\partial\sigma^2/\partial y^2 > 0$, the result of growing y^2 also should be growing inequality. Since moderate followed by rapid expenditure and personal income inequality increases have not been the pattern for India, two possibilities arise. First, it is conceivable that movements in one or more of the sectoral log variances have prevented secular trends in inequality from being observed. Second, as we will discuss presently, it also appears that the share of retained earnings has risen particularly rapidly in the NA sector. Consequently, the personal income growth rate may not differ much among sectors. However, these patterns do change somewhat if we redefine A to encompass all economic activity outside the organized sector.[8]

India's structural transformation has been accompanied by a modest shift in labor force activities. Measured in constant 1970–71 prices, the output of the registered manufacturing sector grew at an annual rate of 5.6 percent from 1960–61 to 1979–80. According to World Bank data, employment in the organized manufacturing sector increased at a rate of only 2.9 percent during the same period (the figure is 2.0 percent for the private component of organized manufacturing). Comparable labor force growth for the period 1965–66 to 1979–80 was even slower. Total organized manufacturing employment grew at an annual rate that was 1.8 percent less than the labor force growth rate. For the entire organized sector, employment grew at an annual rate of 4.4 percent between 1960–61 and 1979–80 (1.9 percent for the private component) and 2.5 percent between 1965–66 and 1979–80.

These figures imply that transformations in the employment structure of the sort expected by Robinson, Kuznets, and others to generate an inequality-income U curve have not yet occurred in India. The driving force in Robinson's exposition is an increase in the share of the labor force employed in high-paying jobs, primarily in the organized nonagricultural sectors. The use of urbanization rates and per capita incomes in regression analysis is justified on the grounds that these variables are correlated with the labor force distribution (Mazumdar 1979).

The sectoral share trends also imply that the share of land in national income is likely to fall with agriculture's share, since land's share is highest in agriculture. In fact, Dholakia (1976) found that land's share in net domestic product fell substantially, dropping from 20.3 percent in 1948–49 to 15.3 percent in 1968–69. Labor's NDP share gained roughly 3 percent, and capital the remaining 2 percent.

Dholakia also divided the economy into agricultural and nonagricultural sectors. Not only did land's share decline, because agriculture declined relative to nonagriculture, but also land's share fell within agriculture (from 31.9 to 27.1 percent) and nonagriculture (from 3.8 to 1.5 percent). When Dholakia (1976, p. 302) regressed the shares on a time trend, he found:

$$\text{Land's share in agriculture} = 33.5 - \underset{(5.29)}{0.38}\ \text{time} \qquad R^2 = 0.58 \tag{8.4}$$

$$\text{Land's share in nonagriculture} = 3.9 - \underset{(34.12)}{0.12}\ \text{time} \qquad R^2 = 0.98 \tag{8.5}$$

$$\text{Labor's share in agriculture} = 50.9 + \underset{(4.03)}{0.41}\ \text{time} \qquad R^2 = 0.45 \tag{8.6}$$

$$\text{Labor's share in nonagriculture} = 67.6 - \underset{(5.86)}{0.13}\ \text{time} \qquad R^2 = 0.63 \tag{8.7}$$

$$\text{Capital's share in agriculture} = 15.6 - \underset{(0.96)}{0.04}\ \text{time} \qquad R^2 = 0.04 \tag{8.8}$$

$$\text{Capital's share in nonagriculture} = 28.6 + \underset{(11.1)}{0.25}\ \text{time.} \qquad R^2 = 0.86 \tag{8.9}$$

The figures in parentheses are t-statistics.

These are strong trends, except for the constancy of capital's share in agriculture. Since labor income is probably more evenly distributed than land or capital income, shifts in factor shares may increase equality in both the rural (predominantly agriculture) and the national income distributions, in view of equation (8.11)—that is, σ_A^2 may exhibit a declining time trend. The falling labor share in nonagriculture may have generated a decline in disposable income relative to value added and hence may have prevented a rise in y^2 that would have increased overall income inequality.

The trend in national income factor distribution is likely to be toward a higher share for labor, given its strong growth coefficient in the agricultural sector and the shift in production to nonagriculture in which labor's weight is greater. For similar reasons one would expect a rising share for capital and a declining share for land in national income. Dholakia's regressions indicate such patterns:

(8.10) Land's share in national income $= 21.2 - 0.34$ time
$(7.50) \qquad R^2 = 0.74$

(8.11) Labor's share in national income $= 57.9 + 0.20$ time
$(3.54) \qquad R^2 = 0.39$

(8.12) Capital's share in national income $= 20.1 + 0.14$ time.
$(5.20) \qquad R^2 = 0.57$

Nevertheless, the level of aggregation is too high to permit an uncomplicated mapping from factor share trends to income distribution trends. Also, the trends in factor and product shares do not have an unambiguous effect on inequality, since they affect its components in equation (8.3) in a variety of ways. It is not even certain that a rising labor share has contributed to lower sectoral inequality (σ_A^2 and σ_{NA}^2), since much of the gain in labor's share appears to have been captured by skilled and other highly paid urban labor. Moreover, the shift from land to capital is likely to worsen the agricultural income distribution, since land is probably more equally distributed in India than is capital ownership.

In summary, it is clear from equation (8.3) that the small decline in P_A was one reason that no distinct time trend in national income inequality emerged between the 1950s and 1970s. Increases in y^2 probably were minor also. As is shown in the following section, the ratio of urban to rural per capita incomes—which is highly correlated with the ratio of nonagricultural to agricultural incomes—shows no clear time trend. Declines in sectoral inequality owing perhaps to labor's growing share in both sectors were evidently sufficient to overcome the effect of the rising differential (y^2).

Many explanations can be given for the factor share trends. Several are discussed by Dholakia; we mention three here. First, the Green Revolution has generated land-augmenting technological progress, which, with elasticities of substitution presumably below unity, has led to a reduction in land's share. Second, investment rates have increased. When the possibilities for substitution are limited, such growth increases the scarcity and shares of labor—particularly, skilled labor. Third, growth in the labor-intensive service sector is likely to raise labor's share in national income.

Dholakia's study covers the period 1948–49 to 1968–69. The patterns he found probably persisted into the 1970s. When Sarma (1980) examined labor income data from the NCAER surveys, he found that its share in national income rose from 33 to 41 percent during the period 1967–68 to 1975–76. Although labor income as defined by the NCAER does not include all labor income (much of it is included in "mixed labor and capital income"), the aggregate trend is probably one of growth.

The Urban-Rural Gap

Of the forces generating India's inequality, the gap between urban and rural living standards provokes the greatest controversy. Social scientists concerned with India tend to dislike the urban sector, which is often viewed as parasitically siphoning off a large portion of the nation's surplus to support its own extravagant consumption propensities. The presence of a pervasive urban bias in national development policy is analyzed in Lipton (1977). We do not study this issue here. Instead, we focus on the factual issue—differences in household earnings and expenditures between urban and rural areas—and we present a brief survey of explanatory theories.

For all households, the urban-rural ratio of per capita household expenditures varies between roughly 1.30 and 1.65 (see table 8-6). Disposable income differentials are slightly higher, ranging from 1.35 to 1.98 (see table 8-7). Neither of these ranges is unusual, nor is a differential undesirable. Urban areas contain a higher proportion of skilled workers and major-asset owners than do rural areas. To some extent, then, the differential reflects the fact that urban and rural demographic compositions differ and not that labor of identical quality receives a higher reward in urban areas.

In addition, the differential reflects higher prices in urban than rural areas. Chatterjee and Bhattacharya (1974) found the price gap to be about 15 percent in 1963–64. They also found that the urban-rural price index was higher for the top quintile (about 1.21) than for the bottom quintile

Table 8-6. *Ratio of Urban-Rural Per Capita Expenditures,* NSS *Surveys*

NSS *round*	*Year*	*All households*
3	1951	1.43
4	1952	1.34
5	1952	1.46
6	1953	1.42
7	1953–54	1.38
8	1954–55	1.65
9	1955	1.56
10	1955–56	1.43
11	1956–57	1.48
12	1957	1.53
13	1957–58	1.38
14	1958–59	1.39
15	1959–60	1.38
16	1960–61	1.38
17	1961–62	1.44
18	1963–64	1.48[a]
19	1964–65	1.33
22	1967–68	1.36
27	1972–73	1.50

a. Chatterjee (1976, table 4, p. 560) puts the ratio for all households at 1.52 and gives the following breakdown by quintile: bottom quintile, 1.27; second quintile, 1.29; third quintile, 1.33; fourth quintile, 1.42; top quintile, 1.70.

Sources: For NSS rounds 3–22: Srinivasan and Bardhan (1974, table 5, p. 194); for NSS round 27: Joshi (1979, tables 3 and 4, p. 7). For round 27 the figures are for the observed distribution. In all cases, the numbers given are for the combined samples.

(about 1.12). Thus, the real income differential in 1963–64 was only 1.32, whereas the nominal income differential was 1.52.

Unfortunately, a complete time series on urban-rural price differentials is not available. For the bottom decile, Chatterjee and Bhattacharya (1974, p. 198) estimate that the 1956–57 differential was only about 5.5 percent. It is unlikely, however, that the gap has widened since 1963–64. Between 1955–56 and 1973–74 the consumer price index for agricultural workers grew by 79 percent; it grew by only 72 percent for industrial workers and 63 percent for urban nonmanual employees (Krishna 1978; World Bank data). Since 1973–74, however, urban prices appear to have caught up. In the absence of an apparent time trend, we assume that the price differential has remained constant, although we recognize that annual variations may account for some of the variation in nominal urban-rural expenditure and income ratios.

Urban-rural expenditure and income (except for 1967–68) differentials

Table 8-7. *Urban-Rural Differentials in Average Disposable Income Per Household,* NCAER *Surveys*

	Sector	Year	Disposable income per household (current rupees)					
			Bottom quintile	*Second quintile*	*Third quintile*	*Fourth quintile*	*Top quintile*	*All households*
(1)	Urban	1960	383	787	1,142	1,776	5,214	1,862
(2)	Rural	1962	375	665	940	1,367	3,101	1,293
(3)	Urban	1964–65	732	1,222	1,746	2,506	7,216	2,696
(4)	Rural	1964–65	506	803	1,075	1,390	3,042	1,364
(5)	Urban	1967–68	899	1,587	2,327	3,417	9,107	3,467
(6)	Urban	1967–68	901	1,594	2,322	3,482	9,026	3,465
(7)	Rural	1967–68	602	1,062	1,664	2,675	6,797	2,560
(8)	Urban	1975–76	2,064	3,481	4,974	7,346	17,490	7,074
(9)	Rural	1975–76	1,230	2,069	2,922	4,170	9,090	3,897
Urban-rural ratio								
	(1) ÷ (2)	1960–62[a]	1.14	1.32	1.35	1.45	1.87	1.61
	(3) ÷ (4)	1964–65	1.45	1.52	1.62	1.80	2.37	1.98
	(5) ÷ (7)	1967–68	1.49	1.49	1.40	1.28	1.34	1.35
	(6) ÷ (7)	1967–68	1.50	1.50	1.40	1.30	1.33	1.35
	(8) ÷ (9)	1976–76	1.68	1.68	1.70	1.76	1.92	1.82

n.a. Not available.

a. The denominator is divided by the index of nominal NNP growth (1960 = 1.00; 1962 = 1.115). Given the interyear income variability in India, however, as well as intersectoral variations from year to year, this ratio is far less meaningful than the others. The ratio for all households in 1959–60 is 1.67.

Sources: Rows (1), (2), and (5) are from Sarma (1980, pp. 29–30); rows (3) and (4) are from Srinivasan and Bardhan (1974, p. 105); rows (6) and (7) are from Jain (1975, pp. 52, 54); rows (8) and (9) are from Sarma (1980, p. 34). The data for 1959–60 are from Lipton (1977). The figures from Jain are obtained from the expenditure shares fitted to the sample and thus differ slightly from the figures presented in row (5). Unfortunately, comparable "unfitted" data were not available for the rural sector in 1967–68.

are greater for the upper classes. As Chatterjee and Bhattacharya (1974) noted, this pattern reflects the greater inequality of the urban income distribution. The differences in the differential across income classes are reduced but not eliminated after decile-specific urban-rural consumer deflators are imposed. The deflated ratios of urban-rural per capita expenditures in 1963–64 rose from about 1.17 for the bottom decile to 1.58 for the top decile.

If we assume that the gap in urban-rural real per capita expenditures or disposable income is 15 to 30 percent for the bottom half of the distribution, what might account for it? In equilibrium, real wages for labor of comparable skills are expected to be equal across sectors.

Accounting practices might bias the differential, and there are attributes of urban and rural life other than income that matter to a potential migrant (see Lipton 1977, chap. 5, for a detailed list). Several factors enhance the attractiveness of urban life: public services (whose value may not be completely capitalized into land rents), including schools, health clinics, and water taps, are more available; low infant mortality rates mean that urban women spend a smaller proportion of their lives in pregnancy than rural women; and many consumer goods are more readily available in urban areas. Other elements increase the attraction of rural life: sanitation problems are less severe, and landowning or sharecropping peasants are probably more secure against famine than they would be if they were unskilled urban workers (although agricultural laborers are probably the least secure; Sen 1977).

The income differential also reflects imperfect substitutes and nonpriced goods. Urban housing is more expensive than rural housing. Some of this expenditure difference is included in price indices. In practice, it is difficult to generate indices that accurately distinguish between differences in price and quality, and the variation in quality between urban and rural housing is substantial. The recorded housing price differential is likely to be arbitrary. It may also be overstated if public service attributes are capitalized into urban property values. Furthermore, the value of open space in rural areas and the cost of urban congestion are neglected in calculations of income differentials.

The fact that per capita comparisons omit considerations of scale economies in expenditures reduces the effective gap (Lipton 1977, p. 147). Because rural households are, on average, larger than urban ones, economies of scale in dwelling space, heating, and cooking should make per capita costs slightly less in rural areas. It is, however, true that underreporting of income and expenditures (believed to increase with income in India) is probably more prevalent in urban than in rural areas.

Overall, there seems to be little reason to believe that a correct consideration of all intangibles and an ideal accounting procedure would render rural life relatively more attractive. This is not to say that a particular index does not contain biases. For example, comparison of per capita consumption expenditures understates the urban-rural gap by omitting savings, which are distributed more unevenly than consumption. The NCAER definition of personal disposable income per capita seems to be the most desirable measure of household welfare, although even it omits differentials in leisure time.

Theories of urban-rural income gaps abound. Although we mention here only a few models, the 15 to 30 percent urban-rural gap is consistent with a wide variety of theories. Furthermore, the desirability of such a gap depends on the model one accepts.

One theory is that a large portion of urban employment is in protected sectors, where favored industrial workers are paid more than their reservation wage. This hypothesis, combined with a queuing model of labor turnover (such as Sabot's [1979] for Tanzania), explains higher average earnings both in the protected industry and in other urban sectors. Protectionist pricing policies by government on behalf of these industries may ensure high capital incomes in those sectors.

Wage gaps can arise even in a competitive situation. Stiglitz's (1974) labor turnover model posits greater hiring and training costs in urban than in rural occupations. If search activities are inversely related to the wage paid, the competitive outcome includes a premium to workers in costly turnover industries.

Mazumdar (1973) finds that wages for unskilled labor in the urban formal sectors exceeded wages for similar labor in the unorganized urban sectors and in rural areas before the advent of labor unions and government intervention. In his analysis of the Bombay labor market, Mazumdar (1979) provides a model similar to that of Stiglitz. He suggests that the supply price of stable labor is higher than that of transitory labor, since stable workers incur the expense of maintaining families in the urban environment. Given the low labor turnover rates observed in Bombay, a higher reservation wage is a more likely cause than search activity for the higher wages paid by industries with low turnover rates.

There are also human capital and demographic theories. Mazumdar (1979) finds that human capital variables are significant determinants of wage variation within Bombay, although sector of work (casual small-scale enterprises or factories) is more important. It is uncertain how much of the urban-rural household per capita income gap can be explained by human capital differences, but it seems almost certain that some can be.

In addition, there are fewer females per male in urban than rural India. In 1981 the total population contained 340 million males and just over 318 million females—a sex ratio of 0.936. Yet the urban population contained 73 million females and 83 million males—a sex ratio of only 0.880 (Government of India 1981).[9] This differential suggests that the ratio of prime-aged male workers to population is higher in urban than rural areas. The proportion of males in prime working years to all males is also likely to be higher in urban than in rural areas. Both of these demographic patterns imply higher per capita incomes and expenditures in urban areas.

Not only does urban India have a relatively skilled population and a high proportion of prime-aged males, but also it has a small proportion of residents in scheduled castes and tribes. In 1971 only 8.8 percent (9.5 million out of 109.1 million people) of urban India belonged to scheduled castes, and 1.2 percent (1.3 million out of 109.1 million) belonged to scheduled tribes. In contrast, more than 18 percent of rural India belonged to scheduled castes (80 million out of 439 million), and 8.4 percent (36.7 million out of 439 million) belonged to scheduled tribes. Thus, more than one-quarter of rural India belonged to disadvantaged groups, but only one-tenth of the urban population fell into these categories (Government of India 1977, table 11). To the extent that these groups suffer systematic labor market discrimination, then, the outcome may be reduced average earnings, particularly in rural areas.

Finally, the incidence of debilitating disease, such as malaria, is almost certainly higher in rural than in urban areas. In summary, cross tabulations suggest factors that might explain a moderate gap between urban and rural real per capita household incomes. A task that awaits researchers is evaluation of the relative importance of the various human capital variables in determining income differentials. A second task involves the testing of implications of various models of income differentials. Normative conclusions regarding the urban-rural gap depend on the economic model of the gap that one chooses.

We now turn from our study of the determinants of the gap to the task of placing India's urban-rural income differential in international perspective. The Indian gap is small by international standards.[10] As compared with an income differential in India ranging from 1.35 to 2, many developing countries have gaps that range from 1.4 to 2.5, and even above 3.0. Of countries for which data are available, only Cyprus, Pakistan, Panama, and Sri Lanka have differentials as small as India's.[11] African countries appear to have very high differentials (although here we run into severe problems of data comparability), as do Latin American countries such as Brazil, Colombia, Costa Rica, Honduras, and Peru. Ratios of 2 to 2.5 are also recorded for the Philippines and Thailand; Malaysia's differential also tends to be somewhat greater than India's.

The Indian urban-rural income differential, although it varies considerably from year to year, has no apparent time trend. Both NCAER income figures and NSS expenditure figures agree that the differential grew in the early 1960s, shrank in 1967–68, but picked up again in the early and mid-1970s. With respect to equation (8.3), it appears that there were no secular trends in agricultural-nonagricultural mean incomes (strongly associated with urban-rural household per capita expenditures) that might create a strong time trend in total inequality.

If the average disposable income per household of the bottom urban quintile is compared with the average for the bottom rural quintile, a secular pattern emerges. Unfortunately, the trend is toward an increasing gap for the bottom quintiles and possibly for the second quintiles (see table 8-7). Little confidence, however, can be placed in these results, especially given the uncertainty surrounding the 1960–62 figure. The models of urban-rural income differentials discussed above do not immediately suggest that there should be a strong time trend in the differential between incomes of the bottom groups.

Patterns in the urban-rural average per capita expenditure differentials by state differ strikingly from the national average (see table 8-8).[12] The sixteen major states can be classified into three groups, two with low expenditure differentials and one with high differentials.

The seven states that border northern and western India make up the first group, which is characterized by low to nonexistent urban-rural expenditure differentials. In Haryana, Jammu and Kashmir, and Punjab, rural per capita consumer expenditures frequently exceed urban per capita expenditures. The urban-rural differential has also been small in Rajasthan (typically just above 1.2), Gujarat (1.12), and Uttar Pradesh (ranging between 1.04 and 1.35). The differential has been slightly greater in Bihar, but except for the NSS 17th Round (1961–62), its differential has always been below the national level.

The second group of states with low differentials comprises the low-income southern states of Andhra Pradesh, Karnataka, and Kerala. These states have slightly higher differentials than do those in the first group, but rarely do their gaps exceed the national level. The differential for Andhra Pradesh ranges from 1.17 to 1.46; for Karnataka it is typically about 1.25; the range for Kerala is 1.13 to 1.40.

Six states with large urban-rural differentials remain. This group includes states with the largest metropolitan areas—Maharashtra, Tamil Nadu, and West Bengal. Maharashtra and West Bengal have differentials of nearly 2.0; Tamil Nadu's is in the range of 1.35 to 1.5. The low-income central states of Madhya Pradesh and Orissa also have above-average differentials, as does the relatively well-off eastern state of Assam. The differential ranges from about 1.25 to 1.5 for Madhya Pradesh, 1.5 to 2 for Orissa, and 1.4 to 1.75 for Assam.

Another pattern evident from table 8-8 is that since the NSS 17th Round only about one-third of India's states have had differentials at least as great as the national rate. In fact, the weighted (by 1971 population) average of the 27th Round state differentials is only 1.43, considerably less than the observed national urban-rural expenditure differential of 1.50. Since the weighted average would equal (ignoring nonstate entities such as Delhi) the national rate if state urban-rural per capita expenditure and urbanization rates did not vary, part of the national differential is attributable to higher urbanization rates in relatively prosperous areas.

It is important to ask whether much of the variation in the state differentials stems from variations in urban-rural differences in the cost of living. Unfortunately, data series on urban and rural prices by state are unavail-

Table 8-8. *Average Per Capita Consumer Expenditures, by State,* NSS *Surveys*

State	NSS *round* 13 (1957–58)	14 (1958–59)	15 (1959–60)	16 (1960–61)	17 (1961–62)	18 (1963–64)	19 (1964–65)	22 (1967–68)	27 (1972–73)
Andhra Pradesh									
Urban	22.60	24.13	24.80	27.75	25.58	28.02	30.66	40.71	56.32
Rural	16.92	16.51	18.32	19.54	20.12	20.75	26.26	30.64	39.79
U ÷ R	1.34	1.46	1.35	1.42	1.27	1.35	1.17	1.33	1.42
Assam									
Urban	33.90	38.44	33.71	37.49	39.62	46.80	41.11	61.51	60.75
Rural	25.47	24.76	24.36	24.05	22.06	26.33	29.15	41.62	41.67
U ÷ R	1.33	**1.55**	**1.38**	**1.56**	**1.80**	**1.78**	**1.41**	**1.48**	1.46
Bihar									
Urban	23.88	23.29	21.45	27.45	35.26	29.99	31.36	44.32	59.91
Rural	17.47	18.92	17.31	25.55	18.83	21.42	26.52	32.97	41.20
U ÷ R	1.37	1.23	1.24	1.07	**1.87**	1.40	1.18	1.34	1.45
Bombay[a]									(Gujarat) (Maharashtra)
Urban	27.22	30.87	32.03	31.88	35.57	35.93	39.02	47.94	(57.58) (74.84)
Rural	17.10	18.52	19.93	20.61	20.76	22.09	25.73	30.97	(51.70) (41.55)
U ÷ R	**1.59**	**1.67**	**1.61**	**1.55**	**1.71**	**1.63**	**1.52**	**1.55**	(1.12) **(1.80)**
Jammu and Kashmir									
Urban	22.34	27.78	25.35	25.09	24.96	31.64	28.92	39.80	49.38
Rural	22.68	23.75	26.77	25.76	24.44	27.89	28.10	37.63	48.14
U ÷ R	0.99	1.17	0.95	0.97	1.02	1.13	1.03	1.06	1.03

Karnataka[b]										
Urban	20.29	24.13	24.80	30.11	27.46	25.71	31.36	37.99	57.89	
Rural	21.01	19.32	19.73	21.47	25.31	20.52	24.95	31.64	44.53	
U ÷ R	0.97	1.25	1.26	**1.40**	1.08	1.25	1.26	1.20	1.30	
Kerala										
Urban	23.11	21.61	22.84	23.62	26.52	27.36	29.27	34.37	58.27	
Rural	16.55	16.91	17.31	20.82	20.98	20.30	22.06	28.30	42.19	
U ÷ R	**1.40**	1.28	1.32	1.13	1.26	1.35	**1.33**	1.21	1.38	
Madhya Pradesh										
Urban	27.22	29.46	28.70	27.45	26.83	29.99	33.45	40.25	61.88	
Rural	18.59	20.53	19.32	18.03	21.41	23.43	26.00	31.30	40.72	
U ÷ R	**1.46**	**1.43**	**1.49**	**1.52**	1.25	1.28	1.29	1.29	**1.52**	
Orissa										
Urban	24.65	28.06	23.12	21.54	34.01	31.64	30.66	47.04	62.35	
Rural	13.38	14.29	16.71	14.60	17.30	19.41	20.48	30.30	34.96	
U ÷ R	**1.84**	**1.96**	**1.38**	**1.47**	**1.97**	**1.63**	**1.50**	**1.55**	**1.78**	
Punjab[c]									(Punjab)	(Haryana)
Urban	25.94	30.30	28.70	33.65	29.64	34.28	35.54	47.94	77.88	69.88
Rural	27.14	34.02	28.38	31.35	32.66	28.56	37.29	44.96	74.62	70.07
U ÷ R	0.96	0.89	1.01	1.06	0.91	1.20	0.95	1.07	1.04	1.00
Rajasthan										
Urban	24.91	30.02	23.68	29.22	29.95	32.63	33.10	49.75	63.87	
Rural	23.98	25.36	28.58	23.19	23.36	23.20	30.46	38.63	51.98	
U ÷ R	1.04	1.18	0.83	1.26	1.28	1.41	1.09	1.29	1.23	

(*Table continues on following page.*)

Table 8-8 (*continued*)

State	NSS *round* 13 (*1957–58*)	14 (*1958–59*)	15 (*1959–60*)	16 (*1960–61*)	17 (*1961–62*)	18 (*1963–64*)	19 (*1964–65*)	22 (*1967–68*)	27 (*1972–73*)
Tamil Nadu									
Urban	24.65	24.69	25.35	26.27	29.95	31.64	33.10	41.61	34.02
Rural	14.87	16.90	16.91	18.68	21.63	23.43	24.42	29.64	37.70
U ÷ R	**1.66**	**1.46**	**1.50**	**1.41**	1.38	1.35	**1.36**	**1.40**	1.43
Uttar Pradesh									
Urban	24.34	23.00	22.57	21.83	25.58	29.00	28.92	43.87	53.55
Rural	17.85	21.34	21.77	20.18	22.71	21.42	26.79	34.63	42.12
U ÷ R	1.25	1.07	1.04	1.08	1.13	1.35	1.08	1.27	1.27
West Bengal									
Urban	31.07	36.20	38.73	36.01	38.69	41.53	39.72	51.56	68.23
Rural	18.96	21.34	20.53	22.76	20.76	23.65	23.11	29.64	38.45
U ÷ R	**1.64**	**1.70**	**1.89**	**1.58**	**1.86**	**1.76**	**1.72**	**1.74**	**1.77**

Notes: All figures are average per capita expenditures per thirty days. Numbers in boldface denote ratio of urban-rural per capita expenditures greater than or equal to the national ratio.

a. The state of Bombay was renamed Maharashtra in 1960; at the same time, Gujarat state split off from Bombay.

b. The early NSS rounds are for the former state of Mysore, which was absorbed into Karnataka in 1973.

c. The state of Punjab was divided into two states, Punjab and Haryana, in the early 1970s.

Sources: Chatterjee and Bhattacharya (1974, tables 5 and 9, pp. 194, 203–206); Joshi (1979, tables 3 and 4, p. 17).

able. It is known that the main cities in the states with the highest differentials (Bombay in Maharashtra, Calcutta in West Bengal, Madras in Tamil Nadu) are the cities with the highest prices (see Gopalkrishna 1968). Yet cities such as Gauhati (Assam), Gwalior and Bhopal (both in Madhya Pradesh), and Cuttack (Orissa) are not particularly expensive, although Indore and Jabalpur (both in Madhya Pradesh) are. Certainly, much of the variation in interstate urban-rural income differentials cannot be accounted for by city-to-city variations in the cost of living. Furthermore, evidence exists that rural price patterns have caused an understatement of variation in the urban-rural differentials across states. Rural CPIs constructed by Chatterjee and Bhattacharya (1974, table 11, p. 210) for the 18th Round (1963–64) indicate that the states with above-average urban-rural per capita expenditure differentials also have rural price levels above the all-India rural price level. Such states are Assam, Maharashtra, and especially West Bengal; only Orissa does not fall into that category. In contrast, the states with the lowest differentials tended to have rural price levels around the Indian average. In summary, from the scanty evidence available it does not appear that interstate patterns in urban-rural expenditure differentials result from systematic regional differences in prices.

Interstate Inequality

By the standards of all but the most developed countries, interregional inequality in India is small. In this section we discuss a few recent contributions to the literature and present some notable characteristics of regional inequality. Relevant data appear in table 8-9.

The reliability of the various indices of well-being is questionable. Of the major states, Bihar had the lowest 1960–61 per capita state NDP. Yet its rural sector per capita consumer expenditure is the third highest (in 1960 all states were overwhelmingly rural). Although usually NSS consumer expenditure surveys and state NDP results are not so inconsistent, conclusions depend on the statistical source.

There is also considerable year-to-year variation in the rankings of well-being. In the nine NSS rounds reported in table 8-8 the rural sectors of two states (Orissa and Kerala) were always among the poorest, and Punjab always had the richest rural sector. Yet six states (Bihar, Jammu and Kashmir, Karnataka, Kerala, Orissa, and Uttar Pradesh) at some time had the poorest urban sector, and three (Assam, Maharashtra, West Bengal) had the wealthiest. To some extent, the large variability in urban rankings probably reflects small sample sizes in the early NSS rounds.

The numbers are also different when regional price indices are used to deflate the figures (table 8-9, column 12). Unfortunately, the only regional price index available to us is by Chatterjee and Bhattacharya (1974) for rural India for 1963–64. A comparison of deflated and undeflated figures for that year suggests that real expenditures for some of the wealthier, relatively urbanized areas (such as Assam, Gujarat, Maharashtra, Tamil Nadu, and West Bengal, but also Kerala) are overstated by current figures. In contrast, several relatively backward states—Madhya Pradesh, Orissa, and Uttar Pradesh—do much better when deflated prices are used.

Perhaps the most distressing indicator of inequality is the diversity in per capita calorie intake in rural India. Even if we allow for the possibility that demographic characteristics account for a portion of the disparity, the variations in both food consumption and malnourishment are striking.[13] Since income elasticities of food demand are usually estimated to be well below unity and the range of state per capita incomes is fairly small, and since price differences are not large, it is surprising to find such consumption variation. Calorie intake in northwestern India, as well as in Madhya Pradesh and Uttar Pradesh, is quite high (3,711 calories a day in Punjab). Intakes in these wealthy states are 33 to 80 percent greater than in Kerala, Tamil Nadu, and West Bengal, where undernutrition is greatest.

Indian regional inequality is surprisingly low, given the size and diversity of the country. In 1960–61, the range between poor states such as Bihar (215 rupees per capita) and Orissa (216 rupees) and the richest states such as Maharashtra (409 rupees) was less than 2 to 1. The range widened by the late 1970s, as successful Green Revolution areas emerged in northwestern India. States such as Punjab (625 1960–61 rupees), Haryana (482 rupees), and Maharashtra (527 rupees) came to enjoy per capita NDPs more than twice the level of Bihar's (237 rupees). Still, the level of inequality remained relatively low even after the Green Revolution took hold in Punjab but not in the poorer states. Per capita consumer expenditures by urban and rural sectors rarely vary by a factor greater than 2. These figures contrast with those for other developing countries, where the variation is often much greater (though price differences may also be greater).[14] The ratio of richest area to poorest commonly exceeds 2; capital regions may have per capita products ten to twenty times that of the poorest region.

Table 8-9 indicates that in the richest states—except for Punjab and Haryana—per capita NDP grew at rates below the national average between 1960–61 and the late 1970s. But long-term trends are difficult to discern. Using the weighted coefficient of variation (CV_W) as his measure of regional inequality, Majumdar (1978) found that regional inequality declined between 1950–51 and 1955–56, remained roughly constant for the next five

(Text continues on page 188.)

Table 8-9. *Indicators of Interstate Inequality*

State	Per capita NDP[a] Rupees 1960–61 (1)	Per capita NDP[a] Rupees 1977–78 (2)	Growth rate[b] (percent) (3)	Ratio to per capita NDP of poorest state[c] 1960–61 (4)	Ratio to per capita NDP of poorest state[c] 1977–78 (5)
Large states					
Andhra Pradesh	275	322	0.93	1.28	1.39
Assam	315	302	−0.25	1.47	1.31
Bihar	215	231	0.45	1.00	1.00
Bombay	n.a.	n.a.	n.a.	n.a.	n.a.
Gujarat	362	408	0.71	1.68	1.77
Haryana	327	482	2.45	1.52	2.09
Himachal Pradesh	314	359	1.35	1.46	1.55
Jammu and Kashmir	269	355	1.65	1.25	1.54
Karnataka	296	366	1.34	1.38	1.58
Kerala	259	291	0.73	1.20	1.26
Madhra Pradesh	260	270	0.22	1.21	1.17
Maharashtra	409	527	1.50	1.90	2.28
Orissa	216	286	1.89	1.00	1.24
Punjab	366	625	3.20	1.70	2.71
Punjab and Haryana	n.a.	n.a.	n.a.	n.a.	n.a.
Tamil Nadu	334	360	0.44	1.55	1.56
Uttar Pradesh	252	275	0.52	1.17	1.19
West Bengal	390	387	−0.05	1.81	1.68
Small states					
Delhi (union territory)	785	1240	2.90	3.65	5.37
Manipur	154	274	3.45	0.72	1.19
Meghalaya	n.a.	n.a.	n.a.	n.a.	n.a.
Tripura	249	374	2.75	1.16	1.62
All-India	307	385	1.34	1.43	1.67

(*Table continues on following pages.*)

n.a. Not available.

Notes: See notes to table 8-8 for changes in state definitions.

a. Per capita NDP at factor cost at 1960–61 prices.

b. Average annual compound growth rate of per capita NDP.

c. Ratio of state per capita NDP at factor cost at 1960–61 prices to per capita NDP of the poorest major state.

Table 8-9 *(continued)*

	Index of consumer					
	Urban				*Rural*	
State	*1960–61* (6)	*1963–64* (7)	*1967–68* (8)	*1972–73* (9)	*1960–61* (10)	*1963–64* (11)
Large states						
Andhra Pradesh	1.29	1.09	1.18	1.05	1.34	1.07
Assam	1.74	1.82	1.79	1.13	1.65	1.36
Bihar	1.27	1.17	1.29	1.12	1.75	1.10
Bombay	1.48	1.40	1.39	n.a.	1.41	1.14
Gujarat	n.a.	n.a.	n.a.	1.08	n.a.	1.17
Haryana	n.a.	n.a.	n.a.	1.30	n.a.	n.a.
Himachal Pradesh	n.a.	n.a.	n.a.	2.03	n.a.	n.a.
Jammu and Kashmir	1.16	1.23	1.16	0.92	1.76	1.44
Karnataka	1.40	1.00	1.11	1.08	1.47	1.06
Kerala	1.10	1.06	1.00	1.09	1.43	1.05
Madhra Pradesh	1.27	1.17	1.17	1.16	1.23	1.21
Maharashtra	n.a.	n.a.	n.a.	1.40	n.a.	1.12
Orissa	1.00	1.23	1.37	1.16	1.00	1.00
Punjab	n.a.	n.a.	n.a.	1.45	n.a.	n.a.
Punjab and Haryana	1.56	1.33	1.39	n.a.	1.47	1.44
Tamil Nadu	1.22	1.23	1.21	1.01	1.28	1.21
Uttar Pradesh	1.01	1.13	1.28	1.00	1.38	1.10
West Bengal	1.67	1.62	1.50	1.27	1.56	1.22
Small states						
Delhi (union territory)	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Manipur	n.a.	n.a.	n.a.	1.10	n.a.	n.a.
Meghalaya	n.a.	n.a.	n.a.	1.85	n.a.	n.a.
Tripura	n.a.	n.a.	n.a.	1.39	n.a.	n.a.
All-India	1.37	1.28	1.32	1.18	1.47	1.15

d. Ratio of state average per capita consumer expenditures to value for the poorest major state (NSS surveys).

e. Deflated, based on state per capita expenditures (rural sector only) converted into all-India prices.

f. Average daily intake of calories, rural sector, 1971–72.

g. Percentage of consumer units with daily intake of less than 1,960 calories a day, 1971–72.

h. Share of food items in total consumer expenditure, 1973–74.

Sources: World Bank data. State income figures differ slightly in base year for constant price series used, source material, and methodology used in generation, and thus have some

expenditure[d] Rural 1963–64[e] (12)	1967–68 (13)	1972–73 (14)	Calorie intake[f] (15)	Calorie deficiencies[g] (percent) (16)	Food shares[h] (percent) (17)	State
						Large states
1.07	1.08	1.14	2,666	25.6	73.0	Andhra Pradesh
1.27	1.47	1.19	2,665	13.6	80.0	Assam
1.10	1.17	1.18	2,732	22.7	81.0	Bihar
n.a.	1.09	n.a.	n.a.	n.a.	n.a.	Bombay
1.06	n.a.	1.48	2,822	20.4	81.75	Gujarat
n.a.	n.a.	2.00	3,652	4.4	77.3	Haryana
n.a.	n.a.	2.01	3,190	7.0	79.7	Himachal Pradesh
1.45	1.33	1.38	3,490	2.4	75.9	Jammu and Kashmir
1.08	1.12	1.27	2,839	21.1	75.9	Karnataka
1.00	1.00	1.21	2,023	41.4	71.7	Kerala
1.28	1.11	1.16	3,536	12.7	75.85	Madhra Pradesh
1.07	n.a.	1.19	2,567	22.7	73.65	Maharashtra
1.04	1.07	1.00	2,533	20.8	79.4	Orissa
n.a.	n.a.	2.13	3,711	5.0	72.1	Punjab
1.59	n.a.	n.a.	n.a.	n.a.	n.a.	Punjab and Haryana
1.13	1.05	1.08	2,394	36.7	75.1	Tamil Nadu
1.19	1.22	1.20	3,198	10.5	78.4	Uttar Pradesh
1.07	1.05	1.10	2,311	29.9	85.0	West Bengal
						Small states
n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	Delhi (union territory)
n.a.	1.57	n.a.	n.a.	n.a.	n.a.	Manipur
n.a.	1.55	n.a.	n.a.	n.a.	n.a.	Meghalaya
n.a.	1.35	n.a.	n.a.	n.a.	n.a.	Tripura
1.17	1.18	1.21	2,724	19.0	77.3	All-India

comparability problems. For several states (Assam, Bihar, Madhya Pradesh, Manipur, Orissa, Tamil Nadu, and Uttar Pradesh) the figures in column (2) were converted to 1960–61 prices from 1970–71 prices using the implicit price deflator of all-India NDP at factor cost. In column (1) the figure for Himachal Pradesh is for 1967–68; in column (2) the figures for Haryana, Karnataka, Kerala, West Bengal, and Delhi are for 1976–77, while the figures for Tripura and Orissa are for 1975–76. Data sources for columns (1) and (2) are as in table 8-8. Figures in column (12) are derived from Chatterjee and Bhattacharya (1974). Figures in columns (15–17) are from Sawant (1982). The figure he reports for Maharashtra in column (17) is 33.65; we assume this is a misprint.

years, then rose to its initial level in 1967–68.[15] In a recent comprehensive study, Mathur (1983, table 1) reaches similar conclusions, but also finds a secular rise in regional inequality between 1964–65 and 1975–76.

Mathur's conclusion is that India's regional inequality measures are following a U-shaped pattern (declining, then increasing), in contrast to the inverted-U pattern observed in international cross-sectional studies (see Vyas 1973). In large part, Indian national trends in CV_W reflect the dominance of the primary and tertiary sectors. Both sectors exhibited initial declines in the regional disparity indices during the 1950s, and then increases in the 1960s. The secondary sector followed a reverse pattern, with regional variation increasing in the 1950s and then declining between 1960–61 and 1975–76. Throughout the postindependence period, variation in state per capita incomes generated by the primary sector has been far lower than CV_W's for the secondary and tertiary sectors.

Mathur provides a detailed decomposition of the sources of regional income inequality. He first decomposes per capita income variation into variation in labor force participation rates and in labor productivity (as well as the usual covariance terms); he concludes that participation rate variation is small. He then decomposes productivity variation to determine the contributions of intersectoral and intrasectoral productivity variation. He finds (p. 493) the intersectoral productivity variation has been more influential (and increasingly so) than variation in productivity across states within the primary, secondary, and tertiary sectors.[16] Only intrasectoral variation in the primary sector is a major contributor to interstate aggregate productivity variation.

Mathur further finds that the decline in variation within the primary sector prior to 1956 and its increase since then have had a substantial effect on the overall productivity variation. His explanation is that productivity advances during the 1950s involved little new technology, but rather were attributable to expanded irrigation and the extension of land under cultivation. By contrast, the Green Revolution of the 1960s and 1970s typically required special conditions for implementation and thus was introduced selectively. One requirement was a capability for financing inputs and capital equipment; consequently, wealthy areas tended to be the first to adopt high-yielding crop varieties.

The secondary sector, in contrast, contributed to an increase in productivity variation in the 1950s and early 1960s and a decrease thereafter. Mathur (1983) and Majumdar (1978) note the tendency of private sector manufacturing investments to be located in high-income and relatively industrial areas, which thus contributes to increased variation. Government efforts to decentralize investments are viewed by Mathur as generating the trend toward deconcentration.[17]

The major contributor to the decreased variation in average productivities in the late 1960s appears to have been a convergence of sectoral productivities. Intersectoral productivity trends continued to narrow regional productivity variation in the 1970–75 period (Mathur, 1983, table 10, p. 493). Prior to 1965, however, relatively rapid productivity growth in the secondary and tertiary sectors had the effect of increasing total variation.

The time trends revealed by the decomposition analysis are interesting, but they are extremely sensitive to choice of years. Green Revolution effects are certainly overstated if 1965–66 is used as a base year and 1970–71 as a terminal year. Cereal yields per hectare in India rose 38 percent during this period, but weather-adjusted yields grew by only 13 percent (Sanderson and Roy 1977, pp. 24–25). The role of weather should be considered in determining interstate variations in productivity and income during any given year: 1975–76, for example, experienced exceptionally favorable weather.

Vyas (1973) attempted to avoid these problems of annual fluctuations in agricultural output. He compared variation in output between averages from the 1959–61 triennium and the 1969–71 triennium, thereby eliminating much weather-induced variation. The coefficient of variation of yield per hectare was considerably higher in the second triennium, but this increase nearly disappeared when Punjab was removed from the sample. Vyas's analysis of the growth in food-grain production between the periods attributed the dramatic output rise in Punjab in particular not only to an increase in the yields of "type A" crops (wheat, bajra, maize, and rice; "type B" crops include jowar and other cereals and pulses), but also to an increase in the area under cultivation for type A crops. Growth in type A yields was also important for other states with above-average growth in food-grain production.

The worst performance was recorded by Maharashtra (where per capita production, which was already below average, fell an astonishing 31 percent), Madhya Pradesh (down 13 percent), and Andhra Pradesh (down 12 percent). Maharashtra and Andhra Pradesh both experienced marked falls in type B yields, along with declines in their shares in the national total of type A crop areas. These shares also fell in Madhya Pradesh, which experienced stagnant yields for both type A and type B crops.

From Majumdar's regressions and Mathur's decompositions it is apparent that state per capita incomes and productivities depend heavily on output structures. Although agricultural or rural productivities in a few of the northwestern states exceed nonagricultural or urban productivities elsewhere, the size of the manufacturing sector is positively correlated with per capita income. Indeed, it would be most surprising if this were not the case. Because the importance of the manufacturing sector is positively correlated with the percentage of the population living in urban areas, state per capita incomes and percents urban are also positively correlated. Majumdar (1978, pp. 347–48) also found that incomes are inversely related to an index of land pressure and positively related to the state literacy rate.

Urban Inequality at the National and State Levels

Researchers and policymakers studying the income distribution in India have devoted most of their attention to the problems of rural poverty. The urban sector has generated less attention, although it is recognized that considerable inequality exists in that sector. In this section we put India's urban inequality in an international context and then examine the pattern of urban inequality across states.

It has recently been shown that urban as well as national income inequality measures tend to follow a Kuznets inverted-U curve across countries.[18] Defining $Q1$ to be the income share of the bottom quintile of households in the urban income distribution, G to be the urban distribution's Gini coefficient, and GDP*khs* to be the natural logarithm of "Kravis-adjusted" per capita GDP in 1970 U.S. dollars (see Kravis, Heston, and Summers 1978), Becker found:

$$(8.13) \qquad Q1_i = 0.498 - 0.1404\text{GDP}khs_i + 0.011(\text{GDP}khs_i)^2 \qquad R^2 = 0.16$$

and

$$(8.14) \qquad G_i = -1.5758 + 0.64533\text{GDP}khs_i - 0.0513(\text{GDP}khs_i)^2. \qquad R^2 = 0.25$$

All coefficients are significant at the 95 percent level. In equation (8.13) the bottom quintile's share reaches a minimum of 5.0 percent at a Kravis-adjusted per capita GDP of \$591 (in 1970 U.S. dollars), compared with a value of \$267 for India in 1967–68; the Gini coefficient peaks at 0.453 at a Kravis-adjusted per capita GDP of \$539.

The income distribution surveys included in the sample were restricted to surveys of household income (as defined in Jain 1975, the principal data source). Despite the caveats mentioned earlier regarding the comparability of household and household per capita distributions, we assume that the NCAER surveys are comparable to the others used in these regressions. The regressions above provide a predicted bottom quintile share (for 1967–68) of 5.7 percent and a Gini coefficient of 0.43 for urban India. These estimates indicate that the predicted urban distribution is slightly more egalitarian than that observed in table 8-6. The observed distribution for 1975–76, however, is somewhat more evenly distributed than predicted.

Superior results were obtained in the cross-country urban sector regressions when a proxy for the logarithm of urban sector real per capita income (UGDPkhs) was used as the independent variable in place of national per capita income.[19] With all variables significant at the 99 percent level, the curves fitted were:

$$(8.15) \qquad Q1_i = 907 - 0.2532 U\text{GDP}khs_i + 0.0187(U\text{GDP}khs_i)^2 \qquad R^2 = 0.29$$

and

$$(8.16) \qquad G_i = -3.0527 + 1.04846 U\text{GDP}khs_i - 0.07832(U\text{GDP}khs_i)^2. \qquad R^2 = 0.37$$

If we assume an urban-rural income gap of 2 to 1 for all countries, including India, the predicted share of India's bottom urban quintile in total urban income is 5.8 percent; the corresponding Gini coefficient estimate is 0.43.[20] The higher predicted share from equation (8.15) relative to (8.13) reflects the greater curvature of the first curve. If we recognize that the urban-rural income ratio in the 1967–68 NCAER survey was only 1.35, and if we adjust the UGDPkhs figures accordingly, the predicted bottom urban share rises to 6.7 percent and the predicted Gini coefficient falls to 0.40.

India's urban sector thus seems slightly less egalitarian than predicted by regressions based on cross-country samples. This conclusion should be taken cautiously, however, because even samples of a similar income unit contain some methodological differences, and there is variability of the distributions from year to year.

These urban U curves also lead to the prediction that India's urban sector distribution will worsen as development proceeds. As the economy develops, increasing proportions of the urban sector will be employed in the modern economy: registered manufacturing, government, and capital-intensive services. At the same time, an increasing proportion of the labor force will become skilled. Both the sectoral and skill transformations generate income-inequality U curves.[21] So far, India's urban sector has not undergone a steady worsening of its income distribution. Counteracting forces, such as a narrowing of skilled-unskilled wage differentials and modern-traditional economy productivity differentials, may keep India from following the cross-sectionally observed path of initially declining and subsequently rising income shares of the bottom urban groups. On a more pessimistic note, however, the absence of a trend may merely reflect the combination of the slow pace of India's employment transformation and the rapidly rising capital-output ratios of its modern sector (see Wolf 1978, pp. 229–52).

We turn now to an examination of urban sector per capita expenditure inequality across Indian states. As is obvious from table 8-10, the Gini coefficients from 1957–58 and 1963–64 are generated from small samples. Consequently, one can have but limited confidence in those estimates. There is also a nonlinear relationship between σ, the displaced log standard deviation reported for 1972–73, and the Gini coefficients.

Great differences exist between the states with high and those with low urban inequality. In 1972–73 Kerala had by far the most inegalitarian urban household per capita expenditure distribution, followed by Meghalaya. The other small northern and eastern states (Himachal Pradesh, Jammu and Kashmir, Manipur, and Tripura) all had relatively low inequality parameters. With the exception of Rajasthan and Meghalaya, all states with displaced log standard deviations above 0.8 were in a contiguous region in south and central India (encompassing Karnataka, Kerala, Madhya Pradesh, Maharashtra, and Tamil Nadu).

There is no association between inequality and observed mean per capita expenditures, of either a monotonic or a quadratic sort. Wealthy urban Maharashtra is highly unequal; so are poor Tamil Nadu and several states in between. Since no specific patterns were perceptible, we turned to regression analysis and used as a dependent variable the 1972–73 displaced log standard deviation shown in table 8-10. We should, however, make it clear that these regressions were performed in order to uncover statistical regularities. No direct causal relationship derived from a formal model has been tested, although we mention possible explanations for the relationships uncovered.

By far the most robust variable in explaining variation in σ is the 1971 state (urban and rural; figures for the urban sector only were unavailable) literacy rate ($LR\,71$). The estimated coefficient is insensitive to changes in the sample and to the inclusion of additional variables. The simple ordinary least squares (OLS) regression of σ on $LR\,71$ is

$$\sigma_i = 0.579 + \underset{(3.31)}{0.696} LR\,71_i + u_i. \tag{8.17}$$

$R^2 = 0.388$; $F = 10.757$

number of observations = 19

The figure in parentheses is a t-statistic. An increase in $LR\,71$ by its standard deviation (0.096) leads to a rise in σ of 0.067, some 8.5 percent of its mean value.[22] States with high literacy rates and inequality indices include Kerala, Maharashtra, and Tamil Nadu, although Gujarat had the fourth-highest 1971 literacy rate and the lowest value for σ.

The link between literacy rates and σ is consistent with the labor force structural transformations mentioned above. States with high literacy rates presumably have a relatively large proportion of the urban labor force in high-paying skilled occupations. Until human capital becomes commonplace, increments to the pool of skilled workers increase inequality. The extent of the increases in inequality depends on the distribution of education and on the possibilities for substitution among labor of different qualities.

To some extent, however, relatively high literacy rates can be expected in states with greater inequality. Education costs include a large private component beyond the means of most poor Indians. Among states with a given low mean income, those with greater inequality are likely to have a larger proportion of their population able to afford education; hence, the inegalitarian state will have a higher literacy rate.

Mean income had little impact. We included observed sample mean per capita expenditures for both the urban and rural sectors. The coefficients were not always significant, even at low significance levels. Nor were they robust: the signs and significance levels changed as we altered the sample size and the set of explanatory variables. When the literacy rate variable is removed, the income values become moderately significant:

$$\text{(8.18)} \quad \sigma_i = \underset{(2.39)}{0.542 + 0.598LR71_i} - \underset{(0.49)}{0.00238\mathrm{PGR}71_i} - \underset{(1.00)}{0.00245MR27_i} + \underset{(0.92)}{0.00195MU27_i} + \underset{(0.64)}{0.148\sigma_i^R} + \underset{(0.61)}{0.226UR71_i} + u_i$$

$R^2 = 0.502$; $F = 2.016$
number of observations = 19

and

$$\text{(8.19)} \quad \sigma_i = \underset{(1.57)}{0.492 + 0.558UR71_i} + \underset{(1.13)}{0.276\sigma_i^R} + \underset{(1.32)}{0.00307MU27_i} - \underset{(1.20)}{0.00324MR27_i} + u_i.$$

$R^2 = 0.265$; $F = 1.260$
number of observations = 19

where

- $\mathrm{PGR}71$ = 1961–71 state population growth rate
- $MR27$ = rural sector's mean per capita expenditure in the NSS 27th Round
- $MU27$ = urban sector's mean per capita expenditure in the NSS 27th Round
- σ^R = rural sector's displaced log standard deviation in the NSS 27th Round
- $UR71$ = percentage of the state's population living in urban areas.

Table 8-10. *Urban Sector Expenditure Inequality Statistics, by State,* NSS *Surveys*

State	27th Round (1972–73) Displaced log standard deviation (σ)[a]	27th Round (1972–73) Sample size	18th Round (1963–64) Gini coefficient	18th Round (1963–64) Sample size	13th Round (1957–58) Gini coefficient	13th Round (1957–58) Sample size
Andhra Pradesh	0.6998	4,654	0.322	359	0.309	n.a.
Assam	0.7949	1,500	n.a.	n.a.	0.246	n.a.
Assam, Manipur, and Tripura[b]	n.a.	n.a.	0.312	202	n.a.	n.a.
Bihar	0.7808	3,693	0.320	264	0.366	n.a.
Bombay	n.a.	n.a.	n.a.	n.a.	0.347	n.a.
Gujarat	0.6116	1,990	0.321	175	n.a.	n.a.
Haryana	0.7282	1,676	n.a.	n.a.	n.a.	n.a.
Himachal Pradesh	0.6872	702	n.a.	n.a.	n.a.	n.a.
Jammu and Kashmir	0.6845	2,688	0.285	125	0.287	n.a.
Karnataka[c]	0.8039	2,466	0.345	214	0.302	n.a.
Kerala	1.1055	1,407	0.368	113	0.349	n.a.
Madhya Pradesh	0.8274	3,213	0.341	308	0.342	n.a.
Maharashtra	0.8795	6,181	0.369	542	n.a.	n.a.
Manipur	0.6561	442	n.a.	n.a.	n.a.	n.a.
Meghalaya	0.9273	502	n.a.	n.a.	n.a.	n.a.
Orissa	0.7022	1,861	0.330	132	0.381	n.a.
Punjab	0.7776	1,697	n.a.	n.a.	n.a.	n.a.
Punjab and Haryana	n.a.	n.a.	0.381	175	0.335	n.a.
Rajasthan	0.8551	2,392	0.350	205	0.343	n.a.
Tamil Nadu	0.8634	3,541	0.334	311	0.405	n.a.
Tripura	0.6933	647	n.a.	n.a.	n.a.	n.a.
Uttar Pradesh	0.7592	5,982	0.378	625	0.335	n.a.
West Bengal	0.7734	4,465	0.310	437	0.385	n.a.
Union territories	n.a.	n.a.	n.a.	n.a.	0.323	n.a.
All India	0.8119	52,820	0.359	4,301	0.359	3,583

n.a. Not available.

a. Parameters were estimated assuming a displaced lognormal distribution of per capita consumer expenditures. The random variable X is said to have a displaced lognormal distribution [that is, $X \sim LN_d(x/\tau, \mu, \sigma^2)$; $\tau < x < \infty$] if X' is distributed lognormally with mean M and standard deviation and $X' = X - \tau$. The parameter τ determines a lower bound to the range of values of the variate X. See Joshi (1979, pp. 4–5). Joshi (1979) and Jain (1977) conclude that India's per capita expenditure distributions are better fitted by estimates of displaced lognormal distributions than by unadjusted lognormal distributions.

b. Only aggregate figures are available for the small, far eastern states of Assam, Manipur, and Tripura.

c. Includes the former state of Mysore.

(Notes continue on the following page.)

The figures in parentheses here and below are absolute values of t-statistics. In these regressions and others not reported, income variables rarely have a strong impact on urban inequality.

Demographic variables such as the urbanization rate and state population growth rate were typically insignificant as well. Urban inequality, however, was consistently positively related to the degree of rural inequality, as we found above. This relationship is robust and is stronger when some variables are removed:

$$(8.20)\qquad \sigma_i = \underset{}{0.390} + \underset{(3.87)}{0.653LR71_i} + \underset{(1.84)}{0.337\sigma_i^R} + u_i.$$

$R^2 = 0.54$; $F = 8.808$
number of observations = 18

Note that the estimated coefficient on σ^R is well below unity: increases in rural inequality are associated with less than matching increases in urban inequality. The relationship between σ and σ^R is not surprising. The most likely explanation is that the urban and rural sectors within a state have partially (if not fully) integrated labor and capital markets, whereas market integration across states is incomplete. Factor income shares influence to some extent the urban and rural inequality within a state. These factor shares are in turn influenced by factor prices, which are correlated within the urban and rural areas of a state. Hence, as long as the influence of factor prices on factor shares is similar in both regions (aggregate factor substitution elasticities are greater or less than unity in both urban and rural parts of the state), the events that determine state factor prices will affect urban and rural inequality in the same direction.

A third variable that has a significant relationship with urban inequality is the growth rate of real per capita state NDP, PCGROW (see table 8-9). The coefficient estimates are consistently negative:

$$(8.21)\qquad \sigma_i = 0.616 + \underset{(4.66)}{0.746LR71_i} - \underset{(3.35)}{0.0469\text{PCGROW}_i} + u_i$$

$R^2 = 0.678$; $F = 15.798$
number of observations = 18

Table 8-10 (*notes continued*)

Sources: Joshi (1979) for the 27th Round and Chatterjee and Bhattacharya (1974) for the 13th and 18th Rounds. All values refer to combined subsamples. Unfortunately, estimates of σ and the Gini coefficients from earlier rounds (the method of calculation of the Gini coefficients is described by Chatterjee and Bhattacharya [1974, pp. 564–65]) are not easily comparable. The nonlinear relationship between σ and the Gini coefficient (based on a displaced lognormal distribution) is provided in Jain (1975).

and

$$(8.22)\quad \sigma_i = \underset{}{0.699} + \underset{(5.41)}{0.812LR71_i} - \underset{(3.25)}{0.0580\text{PCGROW}_i} - \underset{(0.77)}{0.00219\text{PGR71}_i}$$
$$- \underset{(1.74)}{0.00315MU27_i} + \underset{(1.72)}{0.00361MR27_i} + u_i.$$

$R^2 = 0.769$; $F = 8.008$
number of observations = 18

Growth appears to generate equality rather than inequality, but we must remember that association with growth and egalitarian urban sector expenditure distributions need not imply similar associations between state product growth and rural or all-state distributions.

Furthermore, as stated previously, finding patterns is not the same as accepting hypotheses generated by a model. Inequality may cause slow growth, or a third variable (or set of variables) may be responsible for both the growth rate and urban inequality. Nevertheless, there seem to be grounds for less concern over the presence of a growth-inequality or income-inequality tradeoff, at least as far as the Indian urban sector is concerned.

Future Research Needs

We have surveyed the literature pertaining to several key topics in Indian income distribution, but we deliberately omitted a section on rural inequality and poverty. The literature on rural inequality is sufficiently vast to require a separate study. Rural inequality is also a far more frequent topic of high-quality research and thorough surveys than is urban inequality.

One cannot but be impressed at the range of data generated by well-conducted surveys that are available even to researchers dependent on secondary sources. These sources are augmented by frequent NSS publications, the NSS journal *Margin,* and the excellent NCAER journal *Sarvekshana.* Indeed, given the wealth of data available, there seems to be a dearth of research on several topics with regard to India's income distribution.

As discussed above, there is uncertainty over several time series. To examine the presence of time trends, as well as patterns across states, frequent large samples of household incomes and expenditures are needed. Consistent regional price indices would also be most valuable. To obtain further understanding of the determinants of India's income distribution, more comprehensive microeconomic surveys must be prepared and utilized. Panel studies would be extremely useful.

Current understanding of India's factor share trends remains largely conjectural. Both an updating of Dholakia's work and construction of testable models are needed. Priority should also be given to analysis and explanation of the differences in rural and urban income levels. Decomposition analysis could be fruitfully employed in explaining regional variation in income differentials, and microeconomic labor market and migration studies should help to explain wage differentials.

Research on regional variation in inequality can provide information as to what may lie ahead as India develops and what may be avoided. The regressions we have reported are only preliminary in this regard; more extensive data sets (readily obtainable for interregional rural inequality studies) need to be constructed and used to test hypotheses generated from formal models.

Our survey indicates that Indian inequality is comparable to the inequality found in other countries at similar development levels. Yet several sources of inequality that account for much of the inequality elsewhere are comparatively unimportant in India. Interstate variation in per capita product is not great, nor is the gap between urban and rural per capita expenditures. As we saw in chapter 2, India is not highly urbanized for its stage of development. Hence, it is necessarily true (divide the economy into urban and rural sectors in equation [8.3]) that urban or rural sectoral inequality must be relatively great in India in comparison with that in other developing countries, since the urban-rural gap is small. Regressions (8.13)–(8.16) confirm that urban India is slightly less egalitarian than predicted.

These observations suggest that there is a critical need for analysis of the sources of urban and rural inequality. Given the substantial variation in urban and rural inequality across states, what is necessary is an examination of the urban and rural income distributions of particular regions. Further cross-state analysis (extending the work reported in the penultimate section of this chapter) may also prove useful.

References

Adelman, Irma, and Cyntha Taft Morris. 1973. *Economic Growth and Social Equity in Developing Countries.* Stanford, Calif.: Stanford University Press.

Ahluwalia, Montek. 1976. "Inequality, Poverty and Development." *Journal of Development Economics* 3:307–42.

Ahluwalia, Montek, and others. 1978. *India: Occasional Papers.* World Bank Staff Working Paper no. 279. Washington, D.C., November.

Ahmed, Mahfooz, and Nikhilesh Bhattacharya. 1974. "Size Distribution of Per Capita Personal Income in India: 1955–56, 1960–61 and 1963–64." In Srinivasan and Bardhan (1974), pp. 167–182.

Bardhan, P. K. 1974. "The Pattern of Income Distribution in India: A Review." In Srinivasan and Bardhan (1974), pp. 103–38.

Becker, Charles. Forthcoming. "Urban Sector Income Distribution and Economic Development." *Journal of Urban Economics.*

Bhatty, I. Z. 1974. "Inequality and Poverty in Rural India." In Srinivasan and Bardhan (1974), pp. 291–336.

Chatterjee, G. S. 1976. "Disparities and Per Capita Household Consumption in India: A Note." *Economic and Political Weekly* (April 10).

Chatterjee, G. S., and Nikhilesh Bhattacharya. 1974. "On Disparities in Per Capita Household Consumption in India." In Srinivasan and Bardhan (1974), pp. 183–214.

Chenery, Hollis, and Moises Syrquin. 1975. *Patterns of Development: 1950–1970.* London: Oxford University Press.

Cobeljic, Nikola. 1975. *Privreda Jugoslavije: Rast, Struktura i Funkcionisanje* [The Yugoslav economy: Growth, structure and performance]. Vol. 2. Belgrade: Savremena Administracija.

Dholakia, Bakul. 1976. "Behavior of Income Shares in a Developing Economy—The Indian Experience." *Indian Economic Journal* 23, no. 4 (April–June): 295–310.

Frank, Charles, and Richard Webb, eds. 1977. *Income Distribution and Growth in Less Developed Countries.* Washington, D.C.: Brookings Institution.

Gopalkrishna, G. C. 1968. "Ranking of Cities according to Costliness." *Indian Economic Journal* 15, no. 5 (April–June): 618–23.

Government of India. 1977. *Statistical Abstract: India, 1977.* New Delhi.

———. 1981. *Census of India, 1981.* Series 1, paper 2. New Delhi: Census Commissioner of India.

International Monetary Fund. 1984. *International Financial Statistics.* Supplement on Output Statistics. Supplement series no. 8. Washington, D.C.

Iyengar, N. S., and L. R. Jain. 1974. "A Method of Estimating Income Distribution." *Economic and Political Weekly* (December 21): 2,103–109.

Jain, L. R. 1977. "On Fitting a Three-Parameter Log-Normal Distribution to Consumer Expenditure Data." *Sankhya* 39, series C, part 2: 61–73.

Jain, Shail. 1975. *Size Distribution of Income: Compilation of Data.* World Bank Staff Working Paper no. 190. Washington, D.C.

Joshi, P. D. 1979. "On the Suitability of Displaced Lognormal Distribution in Graduating the Size Distribution of Total Consumer Expenditure." *Sarvekshana* (July): 4–8.

Kravis, I. B., A. Heston, and R. Summers. 1978. "Real GDP Per Capita for More Than One Hundred Countries." *Economic Journal* 88, no. 350 (June): 215–42.

Krishna, Sidlov. 1978. "Inflation and Distribution of Income." *Economic and Political Weekly* (January 14): 63–67.

Kumar, Dharma. 1974. "Changes in Income Distribution and Poverty in India: A Review of the Literature." *World Development* 2 (January): 31–41.

Kuznets, Simon. 1955. "Economic Growth and Income Inequality." *American Economic Review* 45 (March): 1–28.

Lipton, Michael. 1977. *Why the Poor Stay Poor: A Study of Urban Bias in World Development.* Cambridge, Mass.: Harvard University Press.

Majumdar, Madhavi. 1978. "Regional Disparities in India: 1950–51 to 1967–68." *Urban Studies* 15:343–49.

Mathur, Ashok. 1983. "Regional Migration in India: A Sectoral Analysis." *Economic Development and Cultural Change* 31, no. 3 (April): 475–505.

Mazumdar, Dipak. 1973. "Labor Supply in Early Industrialization: The Case of the Bombay Textile Industry." *Economic History Review* 26 (August).

———. 1979. *Paradigms in the Study of Urban Labor Markets in LDCs: A Reassessment in the Light of an Empirical Survey in Bombay City.* World Bank Staff Working Paper no. 366. Washington, D.C., November.

Mohan, Rakesh. 1977. "Urban Land Policy, Income Distribution and the Urban Poor." In Frank and Webb (1977).

Mukherjee, M. 1972. "National Income Statistics." In Rao (1972).

NCAER (National Council of Applied Economic Research). 1980. *Household Income and Its Disposition.* New Delhi: Government of India.

Ojha, P. D., and V. V. Bhatt. 1974. "Patterns of Income Distribution in India: 1953–55 to 1963–65." In Srinivasan and Bardhan (1974), pp. 163–66.

Rao, C. R., ed. 1972. *Data Base of Indian Economy: Review and Reappraisal.* Vol. 1. Calcutta: Indian Publishing Society.

Robinson, Sherman. 1976. "A Note on the U Hypothesis Relating Income

Inequality and Economic Development." *American Economic Review* (June): 433–40.

Sabot, Richard. 1979. *Economic Development and Urban Migration: Tanzania 1900–1971.* Oxford: Clarendon Press.

Sanderson, Fred, and Shyamal Roy. 1977. *Food Grains and Prospects in India.* Washington, D.C.: Brookings Institution.

Sarma, I. R. K. 1980. "Household Income Structure and Distribution." *Margin* 12, no. 3 (April).

Sawant, S. C. 1982. "Incidence of Undernutrition in India: An Interregional Perspective." *Indian Economic Journal* 29, no. 4 (April–June): 19–48.

Sekhar, A. Uday. 1983. *Industrial Location Policy: The Indian Experience.* World Bank Staff Working Paper no. 620. Washington, D.C., October.

Sen, A. K. 1977. "Starvation and Exchange Entitlements: A General Approach and Its Application to the Great Bengal Famine." *Cambridge Journal of Economics* 1:35–59.

Srinivasan, T. N., and P. K. Bardhan, eds. 1974. *Poverty and Income Distribution in India.* Calcutta: Statistical Publishing Society.

Stiglitz, Joseph. 1974. "Alternative Theories of Wage Determination and Unemployment in LDC's: The Labor Turnover Model." *Quarterly Journal of Economics* 88, no. 2 (May): 194–227.

Thomas, Vinod. 1978. *The Measurement of Spatial Differences in Poverty: The Case of Peru.* World Bank Staff Working Paper no. 273. Washington, D.C., January.

United Republic of Tanzania. 1970. *Tanzanian Second Five-Year Plan.* Vol. 3, *Regional Perspectives.* Dar es Salaam: The Government Printer.

Vyas, V. S. 1973. "Regional Unbalances in Foodgrains Production in Last Decade." *Economic and Political Weekly, Review of Agriculture* (December).

Wolf, Martin. 1978. "Capital and Growth in India, 1950–71." In Ahluwalia and others (1978).

World Bank. 1980a. "Social Indicators/Data Sheet." Washington, D.C.: Economic Analysis and Projections Department, April. Processed.

———. 1980b. *World Tables: The Second Edition [1980].* Baltimore, Md.: Johns Hopkins University Press.

Notes

1. This discussion is based on Bardhan (1974). Bardhan also links sample differences to the consistently greater inequality found in NCAER surveys than in studies based on NSS surveys.

2. An index of rainfall is available in Sanderson and Roy (1977, pp. 24–25); regional rainfall data are available in annual statistical abstracts, such as Government of India (1977).

3. Ahluwalia (1976) designates as socialist only those countries in which government owns almost all the production facilities. India is nonsocialist in his classification. One reason for the positive coefficient of S is that data for socialist countries cover only workers' earned incomes, which are usually more evenly distributed than household incomes.

4. The predicted values differ slightly if the per capita income figure of $110.30 given by Ahluwalia (1976, p. 340) is used. The estimates are then 5.41 percent and 5.48 percent, respectively, for the bottom quintile shares from the alternate versions of (8.1), and 5.95 percent for the bottom share as predicted by (8.2).

5. Ahluwalia's figures come from P. D. Ojha and V. V. Bhatt, as recorded in Jain (1975), but they differ slightly from Jain's: Jain's quintile shares are obtained from a Lorenz curve fitted to the original data, whereas Ahluwalia's figures are read off a Lorenz curve plotted through individual points recorded in the original source (see Ahluwalia 1976, p. 339).

6. According to World Bank data, between 1960 and 1979 the consumer price index for agricultural laborers and urban industrial workers rose 260 percent, while the current price NNP series grew 272 percent more rapidly than did the 1970–71 price NNP series during the same period.

7. The national accounts figures in this section are based on World Bank data.

8. The "organized sector" is defined to include all public establishments, plus private establishments with twenty-five or more workers.

9. The implied rural sex ratio is 0.953.

10. Comparisons of urban-rural income ratios appear in Lipton (1977, table 5.1, p. 430), in Mohan (1977, p. 438), and in Becker (forthcoming).

11. Unfortunately, the data are not entirely comparable. Most figures are comparisons of average household incomes (a distinction that is not likely to alter our conclusions). But some figures are for earnings per employee, income per income recipient, or incomes of the economically active population or are for agricultural and nonagricultural sectors. Our basic conclusions, though, do not appear to be sensitive to the income unit and sectoral definitions used.

12. Time trends in the state analysis are suspect, owing to small samples in the early NSS surveys. By the 27th Round, the smallest cell was urban Manipur, with 442 sample households (the largest, rural Uttar Pradesh, had 7,985). But nine years earlier, the 18th Round sampled only 113 households in urban Kerala and 4,301 households in all of urban India. The smallest round (15th) had only 2,201 urban and 7,700 rural households. See Chatterjee and Bhattacharya (1974, p. 187); Chatterjee (1976, pp. 557); and Joshi (1979, p. 6). Investigation of the relationship over time between state per capita expenditure differentials and state urban-rural CPIs (proxied by agricultural labor and urban manual employee price indices) would be an interesting exercise. Unfortunately, the appropriate data set was not available to the authors.

13. Nutrition figures are taken from Sawant (1982). Consumer units with an average daily intake of less than 1,960 calories were defined as undernourished. Unfortunately, Sawant's article does not specify whether the units were households (and values were reported on a household per capita basis) or individuals. It is more likely that they were households, and this restricts the interpretability of the data. For example, large households with many children may have low observed per capita intakes without being malnourished. As long as the household size distribution varies little across states, the ranking of states by percentage malnourished is likely to be unaffected; however, the absolute percentage malnourished may be somewhat different from that reported.

14. For example, in Tanzania the per capita gross regional product in 1967 was 5,020

Tanzanian shillings (Tsh.) in Dar es Salaam (the capital region) and Tsh. 695 in Tanga, but only Tsh. 225 in Singida, although the variation did narrow somewhat in 1973 (United Republic of Tanzania 1970, p. 11; World Bank data). In Peru, 1971 per capita expenditure ranged from 16,462 soles in Lima (capital), to 9,732 in the Central Coast region, to 4,302 in the North Sierra, although deflated values showed considerably less variation (Thomas 1978, p. 52). Finally, in Yugoslavia, 1970 per capita social product was twice as high in Macedonia (4,223 1966 dinars) as in Kosovo (2,111 dinars). Yet per capita social product in Croatia (7,463 dinars) was 77 percent greater than in Macedonia, while Slovenia's figure (11,305 dinars) exceeded Croatia's by 51 percent and Kosovo's by 436 percent (Cobeljic 1975, table 67, p. 302). Per capita NDP in Delhi was 2.56 times the national average in 1960–61, but 3.22 times as great in 1977–78 (World Bank 1980a, p. 230); this disparity, however, fails to approach the magnitude of the Tanzanian and Yugoslav differentials.

15. The weighted coefficient of variation is defined as:

$$CV_W = \frac{\left[\sum_i (Y_i - \bar{Y})^2 \frac{N_i}{\bar{N}}\right]^{1/2}}{\bar{Y}}$$

where

Y_i = per capita income of the i^{th} state
$\bar{Y}$ = national per capita income
N_i = population of the i^{th} state
$\bar{N}$ = national population.

16. As Koichi Mera pointed out in a review of a draft of this book, because state policymakers are aware of this pattern, virtually all states favor rapid industrialization and have pressured the federal government for funds and legislation to promote industrial growth. As a consequence, industrial location patterns in India are determined largely by political considerations.

17. As John Hansen and Felipe Morris noted in a review of a draft of this book, these efforts may well be contributors to the declining productivity of industrial investment in India. In a detailed study, Sekhar (1983, pp. 97–98) reaches a similar conclusion: "In general, most of the [industrial location] policies have resulted in a wastage of resources and led to distortions in allocative efficiency and growth."

18. See Becker (forthcoming). This observation holds true whether one uses per capita GNP figures converted at official exchange rates or Kravis-adjusted exchange rates. Kravis-adjusted rates offer a superior indicator of per capita purchasing power across countries and thus are preferred as an indicator of development. As it turns out, Kravis-adjusted figures also generate more significant results.

19. Urban per capita incomes, U_i, were not available for much of the sample. Noting that

$$U_i = \frac{\delta_i \, \mathrm{GNP}_i}{(\delta_i - 1)\,\mathrm{PCTURB}_i + 1}$$

where

GNP_i = i^{th} country's per capita GNP
PCTURB_i = i^{th} country's ratio of urban to total population
δ = ratio of urban-rural per capita income.

Becker (forthcoming) assumed that δ_i was a constant equal to 2 for all countries. He thus defined

$$U\text{GDP}khs_i = \text{GDP}khs_i + \ln\left(\frac{2}{1 + \text{PCTURB}_i}\right)$$

20. The curve estimated in equation (8.15) also reaches a minimum bottom quintile value of 5.0 percent, with the Kravis-adjusted urban per capita income estimate taking a value of $871 at that point. The maximum Gini coefficient reached in (8.16) is 0.456; its associated income value is $807. The Kravis-corrected estimate of Indian urban per capita income is ($\ln \delta = 1.35$; PCTURB = 0.192) only $338, although using the standard δ value of 2 raises it to $448.

21. If equation (8.3) is redefined to express urban inequality (the log variance) as a function of modern urban, *MU*, and traditional urban, *TU*, sector inequality, P_{TU} is initially large; as it declines, aggregate inequality increases.

22. Of the major states, Rajasthan has the lowest literacy rate (19.07 percent) and Kerala the highest (60.42 percent) according to World Bank data.

9

Conclusions and Research Needs

THIS BRIEF CONCLUDING CHAPTER summarizes some of the findings of preceding chapters, outlines the prospects for Indian urban growth until the end of this century, and suggests areas for additional research.

India, like nearly all developing countries, urbanized rapidly during the 1960s and 1970s. But the worldwide cross-sectional regressions reported in chapter 2 indicate that India is not more urbanized relative to its stage of development than would be expected on the basis of the experience of other countries. Those regressions show that urbanization is tied closely to easily available measures of stage of development. They also indicate that countries with large or dense populations do not have larger shares of their populations living in urban areas than do less populous countries at similar stages of development. Although India is not overurbanized in comparison with other countries, it is at a stage of development at which urbanization responds rapidly to economic growth. If economic growth continues or accelerates in India during the remainder of the century, the country will certainly urbanize rapidly. Although such urbanization is desirable and almost inevitable, it will create severe problems with respect to the assimilation of migrants and the planning and financing of urban infrastructure, and it will exacerbate social conflict between and within urban and rural groups.

India urbanized slowly but rather steadily during the first half of the twentieth century. Urbanization has been more rapid since about the time of independence. Chapter 3 showed that urbanization has not been closely associated with the overall growth of the manufacturing sector since independence, for the share of the labor force in manufacturing has not increased by much. Instead, urbanization has resulted from the movement of manufacturing workers from household to factory manufacturing. Household manufacturing is predominantly rural; factory manufacturing is predominantly urban. In all likelihood, the 1981 census will show that household manufacturing has become a small part of the total manufacturing sector. If so, future urbanization will depend on the growth of factory manufacturing and of other predominantly urban sectors.

Chapter 4 examined a wealth of data on the size distribution of urban agglomerations. India has one of the most dispersed systems of city sizes—the share of its urban population living in the largest urban agglomeration is smaller than the corresponding figure in just about every other country. India's city size distribution has been remarkably stable during the twentieth century. It became gradually more dispersed during the first half of the century and has become slightly less dispersed since independence. The city size distribution varies among India's five large regions, but in recent decades the regional distributions have become more similar to one another and to the national distribution.

The determinants of growth of individual cities were analyzed at the end of chapter 4. It was found that city growth is only loosely correlated with the city's manufacturing growth; the observation that larger cities tend to grow faster by virtue of their size holds true only for cities with somewhat less than 1 million people. Study of the determinants of city growth was continued in chapter 5 with a detailed analysis of city growth in Madhya Pradesh from 1961 to 1981. We found that city growth there depended strongly on the structure of a city's industrial sector; city growth was promoted by the presence of one of several modern manufacturing or mining industries. There may or may not be a conflict between the findings in chapters 4 and 5. The analysis in chapter 4 was based on a national sample of cities, whereas that in chapter 5 was based on a sample from one state. Furthermore, the findings in chapter 5 relate to a much more detailed specification of industrial structure than do those in chapter 4. More research remains to be done on the effects of industrial structure on urban growth in India. Limited evidence was found that unusually large local government expenditures promote growth.

In chapter 6 we studied the effects of city size on total factor productivity in manufacturing. Using a carefully constructed index of total factor productivity in manufacturing, we found that productivity is high in states where a large percentage of the population resides in urban areas and is high in states whose large cities have large populations. This study confirms for India the findings of related studies in other countries that total factor productivity increases with city size.

Chapter 7 discussed issues related to Indian government attempts to control the size of large cities. After presenting a brief survey of Indian government programs to limit the growth of large cities, we analyzed the reasons for the excessive size or expansion of large cities and concluded that the most important reason in most developing countries is excessive expenditure on local government services financed by taxes levied elsewhere. If this is true, the appropriate remedy is to reform the system of allocating funds for local governments so that a better match between needs and funds can be made. This kind of excessive spending is probably not such a problem in India, since most local government spending is financed by local taxes through a democratic process. The situation may change, however, if the share of local government expenditure financed by Indian state and national governments increases in future years.

In chapter 8 we concluded our studies with a survey of national and urban income inequality. In India, as elsewhere, income inequality is somewhat greater in urban than in rural areas. But the data on Indian income distribution do not reveal a trend of increasing inequality; this is contrary to the experience of other countries that have proceeded through similar stages of development. It is possible that continued growth during the remainder of the century will be accompanied by increasing inequality in India.

What are the prospects for urban growth in India during the remainder of this century? India is at a stage of development at which urbanization responds rapidly to economic growth. If we assume a continuation or mild acceleration of recent Indian growth of real GNP per capita, the percentage of the population living in urban areas will inevitably be higher in 1991 or 2001 than in 1981. In chapter 2 we suggested that percent urban is likely to increase from its 1981 value of 23.7 to a 1990 value of about 28. By 2000 it is almost certain to be well over 30, perhaps 32 to 34.

Attempts to slow or stop urbanization can succeed only at the cost of slowing economic growth. As productivity growth in agriculture enables a smaller share of the labor force to produce what can be consumed and as higher urban wage rates induce rural to urban migration, economic growth will be accompanied by rapid decreases in the share of the labor force in agriculture.

In the normal course of events much of this urban growth will take place outside the largest urban agglomerations. This tendency has already begun. During 1971–81, only one of India's four largest urban agglomerations—Delhi—grew as fast as the total urban population, and its growth was an inevitable consequence of government expansion. Even without specific government programs to encourage it, the most rapid urban growth in the future will probably be in Class I cities smaller than Bombay, Calcutta, Delhi, and Madras. The result will probably be a shift to a somewhat more dispersed city size distribution by the end of the century. This has been a common pattern in other developing countries.

As was emphasized in chapters 3, 4, and 7, governments can influence the growth of particular cities, and perhaps even of particular size classes of cities, by investing in infrastructure and providing public services. Decisions of these kinds are extremely difficult and deserve and require all the intellectual and political resources that can be brought to bear on them. Direct controls on the growth of cities by restrictive licensing of production in large cities lead to distortions in resource allocation.

We conclude with a few suggestions for further research on Indian urbanization. The suggestions are presented in roughly the same order in which subjects were studied in this book.

The statistical analysis of the determinants of urbanization in chapters 2 and 3 vastly oversimplified a complex process. Urbanization occurs during economic development because of the growth in the employment shares of sectors that are predominantly urban. Understanding urbanization mainly requires understanding the growth process that shifts employment from rural to urban sectors. One way to study the process is with the aid of a multisector, multi-input growth model. Such a model has been formulated, estimated, programmed, and simulated for India by Jeffrey Williamson and the authors of this book, and the results will be reported separately. This type of computer analysis is expensive and time-consuming. Additional insights can probably be gained at much less cost by analyzing Indian regional trends and differences. For example, it should be possible to analyze the relationship between agricultural and urban development in a small model that distinguishes among agriculture and agricultural processing and other urban sectors. Indian regional data show a great variety of types and levels of agricultural development. Combined time series and regional analysis should be fruitful.

A second area in need of further research concerns urban infrastructure investment and local government service provision: transportation, education, health, and utilities. Two sets of questions are particularly important. First, what kinds of service provision should have highest priority, when both efficiency and equity considerations are taken into account? What investments have the largest benefits relative to costs, and what investments best meet the needs of low-income groups? Second, what are the effects of infrastructure investments on urban migration? Does elaborate government service provision attract migrants, and, if so, what services have the largest effects? Can growth of medium-size cities be stimulated by careful allocation of infrastructure investments? What political process for raising and spending infrastructure funds will lead to the most efficient and equitable allocation? These issues can be studied only with a cross-sectional sample of infrastructure, city size, and migration data.

Much more research is also needed on the relation between productivity and city size. Verma's study in chapter 6 has confirmed for India the finding of studies of U.S. and other data that total productivity increases with city size. More analysis—using data from individual cities—is needed. Verma's chapter discussed the techniques that can be used. Further study should include not only data on inputs employed by firms but also data on infrastructure, so that it will be possible to estimate the effects of infrastructure on productivity in manufacturing and other sectors. Every issue of government policy toward cities hinges on the outcome of such an analysis. Moreover, such an analysis would shed further light on the determinants of city growth and efficiency and on the effects of government spending on the living standards of residents, especially poor residents.

Indian scholars have devoted much effort to fine studies of income distribution. We reviewed these studies in chapter 8. The next logical question is, What government programs are likely to be most efficient in raising the incomes and income shares of the lowest income groups? An extensive system of direct transfers is probably impractical in the Indian context. If so, the correct approach is to promote programs that increase the productivity, and therefore earnings, of the poor. The programs that are usually suggested are those that increase the demands for skills possessed by low-income people and those that improve the skills of low-income people. Programs in the first category include irrigation and other investments in low-income rural areas and efforts to industrialize backward districts. Programs to improve the health, education, and training of low-income people make up the second category. What combinations of programs are likely to have the greatest effects? In India even more than in most other countries it is important to improve the living standards of the poorest groups with programs that do not slow economic growth. Hence, programs to improve the health and education of low-income groups are preferable to programs to increase the demands for services of low-productivity workers. Programs of the latter kind divert resources from places and uses of potentially high productivity to lower productivity uses.

In connection with the issue of equity, careful research is needed of the effects of various programs to raise the productivity of the poorest groups. The first step should be careful "hedonic" studies of the determinants of earnings. This would reveal the efficacy of various government programs to improve productivity and earnings and would make cost-benefit analysis of programs possible.

Index

Edwin S. Mills is professor of economics at Princeton University. Charles M. Becker is assistant professor of economics at Vanderbilt University. Satyendra Verma is a Ph.D. candidate at the University of Maryland.