

11910

# The Economics of Urbanization and Urban Policies in Developing Countries



edited by  
George S. Tolley  
Vinod Thomas

A World Bank Symposium

077-01 0006  
Zlatichz Morko 1987  
R 3008



---

The Economics of Urbanization  
and Urban Policies  
in Developing Countries

A World Bank Symposium

---





# The Economics of Urbanization and Urban Policies in Developing Countries

*edited by*  
George S. Tolley  
Vinod Thomas

The World Bank  
Washington, D.C., U.S.A.

---

© 1987 by the International Bank  
for Reconstruction and Development / THE WORLD BANK  
1818 H Street, N.W., Washington, D.C. 20433, U.S.A.

All rights reserved  
Manufactured in the United States of America  
First printing April 1987

The findings, interpretations, and conclusions expressed in this study are entirely those of the authors and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent. Maps that accompany the text have been prepared solely for the convenience of readers; the designations and presentation of material in them do not imply the expression of any opinion whatsoever on the part of the World Bank, its affiliates, or its Board or member countries concerning the legal status of any country, territory, city, or area, or of the authorities thereof, or concerning the delimitation of its boundaries or its national affiliation.

**Library of Congress Cataloging-in-Publication Data**

The Economics of urbanization and urban policies in developing countries.

(A World Bank symposium)

Conference held in 1983.

Includes bibliographies.

1. Urbanization—Developing countries—Congresses.
2. Urban policy—Developing countries—Congresses.
3. Urban economics—Congresses. I. Tolley, George S., 1925–  
II. Thomas, Vinod, 1949– . III. Series.

HT169.5.E25 1987 307.7'6'091724 87-2181  
ISBN 0-8213-0786-X

---

## Foreword

Urban problems in developing countries have become more acute in recent decades as people have flocked to cities, and the largest cities have been affected the most. In coming years, as population growth continues throughout the developing world, urban problems promise to become increasingly severe. This volume seeks to promote better understanding and evaluation of policies designed to cope with these issues.

To devise policies intelligently, we need to understand why urbanization has taken the forms we see today and what its future course is likely to be. Considerable progress has been made recently in the economic analysis of urbanization as a whole and its facets, including employment, housing, transport, pollution, and poverty. This book draws together studies of the causes of observed urbanization patterns and builds on them to provide a *better foundation for policy analysis*.

Almost all countries have grappled with policies to control urbanization and manage growing cities more effectively. This book reviews the accumulating experience in both market and centrally planned economies and provides a critique of successes and failures.

Urban policies are of two main types. The first attempts to influence the degree of urbanization and the distribution of population among regions and cities. In addition to surveying country experiences with such policies, this book provides methods for evaluating them.

A second type of urban policy attempts to make the best of urbanization by coping with the urban management problems encountered during economic development, particularly as they are aggravated by urban poverty. Policies of the second type must be based on sound economic analysis of urban functions and must

take account of practical difficulties in implementing projects and the realities of institutional settings. This book places the best in economic analysis alongside real-world problems, contributes to fruitful interaction between the two, and provides analytical tools for improving policies.

A unique contribution of this book is that the findings as a whole provide a unified framework for earlier work on urban problems. The studies build on progress to date to make new contributions that should be of value to policymakers, advisers, project analysts and other practitioners, academicians, and observers of urban problems.

The book consists of four parts, which are introduced by an overview. Part I provides the economic underpinnings for understanding urbanization. Comparative advantage in urban and agricultural production, the relations between urban and rural poverty, and the effects of total population growth are considered, and the ways in which these and other factors have led to observed differences in country urbanization experiences are brought out.

Part II provides a framework for policy analysis. Correction of market failures, redistribution of income, and other goals are considered as bases for policy. Policy approaches in both market and planned economies are reviewed.

The remainder of the book draws on the first two parts to derive tools for the evaluation of policies. Part III surveys national policies concerning urbanization and gives examples of policies that encourage or discourage concentration of a country's population in one or a few large cities. Experiences in Brazil, Colombia, the Republic of Korea, and Venezuela are evaluated. Part IV ex-

amines urban management problems, especially the provision of urban services and the maintenance of a good or at least tolerable environment in large cities. The emphasis is on techniques of quantitative analysis and on project management. Housing, transport, urban finance, and pollution control are considered. A concluding chapter draws policy lessons and makes suggestions.

The book, which contains several contributions that originated from work at the World Bank, distills the

Bank's accumulated knowledge of urban policy. It is based in part on a conference at the University of Chicago in 1983 in which experts from the Bank and from universities throughout the world participated.

*Anthony A. Churchill*

Director

Water Supply and Urban Development Department  
The World Bank

*June 1986*

---

# Contents

Contributors	xi
Abbreviations	xii
1. An Overview of Urban Growth: Problems, Policies, and Evaluation <i>George S. Tolley and Vinod Thomas</i>	1
Patterns of Urbanization	2
Urbanization and Economic Development	4
Sources of Future Urbanization	5
Economic Causes of Urban Problems	5
Urbanization Policy in Market and Mixed Economies	7
Urbanization Policy in a Centralized Economy	7
Concentration and Decentralization Policies	8
Addressing Urban Problems	10
Notes	12
Bibliography	12
PART I. UNDERSTANDING URBANIZATION IN THE DEVELOPING WORLD	13
The Determinants of Urbanization	13
Urbanization in Representative Developing Economies	14
Conclusion	14
Bibliography	14
2. Urbanization and Economic Development <i>George S. Tolley</i>	15
Demand and Supply Factors	16
A More General Explanation	21
Conclusion	29
Bibliography	30
3. What Drives City Growth in the Developing World? <i>Allen C. Kelley and Jeffrey G. Williamson</i>	32
The Limits on City Growth	33
Predicting the Past	35
Understanding City Growth: Some Major Influences	38
The OPEC Watershed and Recent Growth Trends	41
Conclusion	43
Notes	43
Bibliography	45

PART II. A FRAMEWORK FOR URBANIZATION POLICY	47
Urban Problems and Policy Implications	47
The Scope of Urbanization Policy	47
A Case of Central Planning	48
Bibliography	48
4. Market Failures as Bases of Urban Problems	49
<i>George S. Tolley</i>	
Urbanization and Poverty	49
Urbanization as a Source of Growth: An Illusion	51
Has Overurbanization Occurred?	52
Primal Cities	56
Economic Policy and Urbanization	57
Bibliography	58
5. Urban Development Policies in Developing Countries	60
<i>Bertrand Renaud</i>	
Basic Trends in Urban Policy	60
Urbanization Policies	63
Analytical Issues	65
The Direction of Urbanization Policies	66
Conclusion	70
Notes	71
Bibliography	71
6. Urban Policy in Centralized Economies: China	73
<i>William L. Parish</i>	
City Size and Stability	73
Useful and Productive Employment	76
Health Care and Housing	79
Conclusion	81
Notes	82
Bibliography	83
PART III. CONCENTRATION OR DECENTRALIZATION: EVALUATION OF POLICIES	85
Impetus to Urban Concentration	85
Locational Choice and Decentralization	85
Impact of Decentralization Policy	86
Bibliography	86
7. The Analysis of Urban Concentration and Decentralization: The Case of Brazil	87
<i>J. Vernon Henderson</i>	
Urban and Regional Patterns	87
Government Influence on Industrial Location	88
Sources and Magnitudes of Economies of Scale	90
Environmental Considerations in Decentralization	92
Conclusions and Policy Implications	93
Notes	93
Bibliography	93
8. Employment Location and Spatial Policies: Colombia and Korea	94
<i>Kyu Sik Lee</i>	
Observed Patterns of Employment Location	95
Behavioral Underpinnings of Location Choice in Manufacturing	98
A Framework for Evaluating Policy Effects	101
Policy Implications	102
Notes	103
Bibliography	103

9. Industrial Deconcentration Policy: Venezuela	106
<i>Benjamin Reif</i>	
Has Deconcentration of Industry Taken Place?	106
Is Deconcentration a Result of Policy?	109
Other Factors	110
Conclusions	118
Notes	119
Bibliography	119
<b>PART IV. MANAGING THE CITY</b>	<b>121</b>
Toward Greater Fiscal Efficiency	121
Urban Housing	121
Urban Transport	122
Urban Services and the Environment	122
Improving the Efficiency of Shelter Projects	123
Bibliography	123
10. Intergovernmental Fiscal Relations in Developing Countries	124
<i>Roy Bahl and Johannes Linn</i>	
Urban Government Finance as a National Problem	124
The Urban Fiscal Gap	125
Expenditure Pressures	125
Revenue Constraints and Opportunities	129
Policy Options for Urban Fiscal Reform	130
The Politics and Prospects of Urban Fiscal Reform	131
Notes	132
Bibliography	133
11. Housing Demand in the Developing-Country Metropolis	135
<i>Gregory K. Ingram</i>	
The Price Term in Housing Demand Equations	136
Housing Demand and Workplace-Based Price Variation	137
The Setting and the Data	138
The Hedonic Price Equations	139
The Housing Demand Equations	140
Aggregate Estimates of Income Elasticities	143
Conclusion	144
Notes	144
Bibliography	144
12. Urban Transport Policy: Colombia	145
<i>Alvaro Pachón</i>	
Operations	145
New versus Old Buses	148
Buses or Busetas?	150
Summary and Conclusions	153
Notes	154
Bibliography	154
13. Evaluating Pollution Control: The Case of São Paulo	156
<i>Vinod Thomas</i>	
Pollution: Problems and Policies	156
The Costs of Pollution Control	157
The Benefits of Pollution Control: Health Effects	159
Policy Implications	162
Notes	164
Bibliography	164

14. Improving the Effectiveness of Urban Projects	166
<i>Douglas H. Keare</i>	
Project Efficiency and Effectiveness	167
Some Implications for Project Design and Policy Formulation	174
The Contributions of Evaluation	178
Notes	179
Bibliography	180
15. Urban Economic Policy: Directions for the Future	181
<i>George S. Tolley and Vinod Thomas</i>	
Urbanization Problems in Mixed Economies	182
Lessons under Central Planning	182
Urban Concentration and Urban Policy	183
Bibliography	184

---

## *Contributors*

Affiliations are those at the time of the conference.

- Roy Bahl Professor of economics and public administration and director of the Metropolitan Studies Program, Maxwell School, Syracuse University, New York
- J. Vernon Henderson Professor of economics, Brown University, Providence, Rhode Island
- Gregory K. Ingram Economic adviser, Water Supply and Urban Development Department, The World Bank
- Douglas H. Keare Chief, Education Division, Latin America and the Caribbean Regional Office, The World Bank
- Allen C. Kelley James B. Duke professor of economics, Duke University, Durham, North Carolina
- Kyu Sik Lee Senior economist, Water Supply and Urban Development Department, The World Bank
- Johannes Linn Senior economist, Country Programs Department, East Asia and Pacific Regional Office, The World Bank
- Alvaro Pachón General Manager, Corporación Centro Regional de Población, Bogotá, Colombia
- William L. Parish Professor of sociology, University of Chicago, Chicago, Illinois
- Benjamin Reif Professor, Faculty of Architecture and Urban Planning, Central University of Venezuela, Caracas
- Bertrand Renaud Senior economist, Water Supply and Urban Development Department, The World Bank
- Vinod Thomas Senior economist, Latin America and the Caribbean Regional Office, The World Bank
- George S. Tolley Professor of economics, University of Chicago, Chicago, Illinois
- Jeffrey G. Williamson Laird Bell professor of economics, Harvard University, Cambridge, Massachusetts

*Note:* John Crihfield made a major contribution to the shaping of this volume.

---

## *Abbreviations*

CEDE	Centro de Estudios para Desarrollo Económico (Colombia)
CETESB	Companhia Estadual de Tecnologia de Saneamento Ambiental (Brazil)
CFT	Corporación Financiera del Transporte (Financial Transportation Corporation) (Colombia)
CIS	Centro de Informacoes de Saude (Brazil)
DANE	Departamento Administrativo Nacional de Estadística (National Statistics Department) (Colombia)
EMPLASA	Empresa Metropolitana de Planejamento de São Paulo (Brazil)
FEDESARROLLO	Fundación para la Educación Superior y el Desarrollo (Colombia)
FSDVM	Fundación Salvadoreña de Desarrollo y Vivienda Mínima (El Salvador)
GDP	Gross domestic product
GNP	Gross national product
GSP	Greater São Paulo Metropolitan Area
IDRC	International Development Research Centre (Canada)
INTRA	Instituto Nacional de Transporte (Colombia)
ISIC	International Standard Industrial Classification
LCC	Lusaka City Council (Zambia)
MIT	Massachusetts Institute of Technology
OECD	Organisation for Economic Co-operation and Development
OPEC	Organization of Petroleum Exporting Countries
RDE	Representative developing economy
SEADE	Fundação Sistema Estadual de Analises de Dados Estatísticos (Brazil)
SEPLAN	Secretaria de Economia e Planejamento (Brazil)
TFFPG	Total factor productivity growth

---

## *An Overview of Urban Growth: Problems, Policies, and Evaluation*

George S. Tolley  
Vinod Thomas

The world has been urbanizing rapidly for a long time and shows every sign of continuing to do so: more than 40 percent of the world's population today are urban dwellers. Dramatic urbanization over the past two or three decades has been concentrated in the developing countries, where the urban population has been expanding at twice the rate observed in the countryside. In fact, the world's biggest cities have increasingly been in the developing nations, and population concentrations in such urban centers as Mexico City, São Paulo, and Shanghai have reached extraordinary levels.

The rapid growth of cities and their large sizes notwithstanding, the urban share of total population in the developing countries as a whole is less than one-third, compared with more than three-fourths in industrial market economies. Furthermore, the current pace of urbanization in developing countries is not unprecedented: the urban population in some of today's advanced economies grew even faster in the past. The difference lies in the much larger absolute sizes of a few cities and of urban agglomerations in the developing countries today. The urban populations constitute a relatively low share of the national totals, however, as a result of significant population increases from large bases in the rural areas.

Whether or not the current pattern of urbanization in developing countries is alarming when viewed from a historical, spatial, and cross-country perspective, it greatly concerns most governments in those countries. Consequently, some countries follow policies intended to slow down urban concentration and even reverse rural-to-urban migration. Some measures—such as the creation of new cities and the discouragement or pro-

hibition of further industrial concentration—are intended to decentralize economic activity from large, congested urban agglomerations. The stated objectives include the need to reduce the economic and managerial costs of overcrowding and improve regional and urban-rural income distributions. At the same time, other economic policies—for example, credit subsidies, import protection, and urban food subsidies—more often than not indirectly promote urban concentration. Awareness of such effects has prompted some policymakers to reconsider these implicit spatial policies.

This volume is a response to the increasing preoccupation with urbanization in developing countries and to the associated policy issues. In part, the book is concerned with an understanding of why and how urbanization occurs. An examination of the patterns of urbanization among countries and over time puts in perspective the more recent developing-country experiences and brings out similarities in urbanization in different countries during economic development. Explanations of the processes of urbanization are also essential to better appreciate the benefits and costs connected with those processes. These concerns are linked to another goal of the book, an evaluation of urbanization policies. Policy initiatives that attempt superficially to counter deep-rooted urbanization trends can be costly; policies to address the underlying causes of concern and ameliorate urban (as well as rural) problems directly can be beneficial.

Linn (1983) and Renaud (1981) have made important contributions in this area. Renaud compares macroeconomic, urbanization, and interurban policies in developing countries with those in the developed countries and

discusses the need to correct the undesirable spatial effects of national economic policies, make internal management of cities more efficient, and increase economic efficiency by eliminating barriers to mobility of resources and dissemination of innovations. Linn focuses in greater detail on intraurban issues and on how to increase the efficiency and equity of growing cities. He covers a wide range of areas—urban employment, income distribution, transport, housing, and social services—and evaluates the effectiveness of such policy instruments as public investment, pricing, taxation, and regulation.

This study focuses more sharply on the policy issues connected with concentration and decentralization. Evaluation methods and quantitative assessments of policy effects have been drawn from the experiences of different countries. The book complements the more general urban review of Renaud and the city-specific findings of Linn in that it seeks to sharpen the economic analysis and evaluation of urbanization policies at the national and city levels. A unique aspect of this book is its attempt to deepen our understanding of the economic benefits and costs of urbanization and urban policy interventions without denying the contribution of other disciplines to an understanding of urbanization.

### Patterns of Urbanization

In comparisons of urbanization it must be kept in mind that countries do not report their urban populations uniformly. The reasons for nonuniformity range from use of different criteria in defining urban areas to differences in concepts and in the accuracy of statistics among countries. Furthermore, global data for the urban sector usually gloss over important intraurban differences among countries, including variations in the definition and nature of large as against small cities. Statistics on urbanization should therefore be used with caution and perhaps only for broad comparisons.

### Historical Comparisons

Some argue that developing-country urbanization today is qualitatively different from the historical pattern in today's developed countries. The main difference relates to the absolute levels of urbanization today, which in good measure are based on large overall populations. High levels of absolute poverty are also associated with urbanization. During 1970–82 the developing countries experienced a more rapid urbanization than did the industrial countries (table 1-1). The orders of magnitude are even more striking between 1950 and 1980, when the urban areas of the developing countries (excluding China) absorbed nearly 600 million additional people—twice the total number of urban dwellers in industrial countries at the beginning of that period.<sup>1</sup>

If the recent trends continue, an additional 1 billion urban dwellers may be added in developing countries by the end of this century. The magnitudes are historically unique. Unprecedented concentration of people is likely to prevail in several cities in the developing countries which are already among the biggest urban agglomerations in the world. According to U.N. projections, Mexico City and São Paulo each may contain more than 25 million people by 2000, closely followed by Shanghai, Beijing, and Rio de Janeiro; of the twenty-five cities that are likely to have more than 10 million people, twenty are expected to be in what are now considered developing countries. The magnitude of this growth is likely to compound the problems of managing cities and of addressing absolute poverty and unemployment.

Although the sheer magnitude of urbanization and of associated problems in developing countries is phenomenal, some other salient aspects of the present urbanization pattern do not depart markedly from past experience. A comparison of trends over many years reveals no dramatic changes in the growth rate of the world's urban population. The share of the world's population in urban areas has increased steadily, from an estimated 3 percent in 1800 to more than 40 percent today, but the percentage change per decade in the

**Table 1-1. Urban Population: Share and Growth, by Country Group**  
(percent)

Country group	Urban population as share of total population		Average annual growth in urban population		Average annual compound growth in urban share, 1960–82
	1960	1982	1960–70	1970–82	
Low-income	17	21	4.1	4.4	1.06
Middle-income	33	46	4.4	4.2	1.67
Industrial market	68	78	1.9	1.3	0.69
Industrial nonmarket	48	62	2.6	1.8	1.29

Source: World Bank (1984), annex table 22, p. 260; the last column is derived from the first four.

proportion that is urban has remained around 16 percent during this period.<sup>2</sup> As seen in table 1-1 the developing countries today are adding to their urban populations more rapidly than are the developed countries, but these rates are not very unlike those observed during earlier periods of urbanization in what are now developed economies. Hence, the long-term processes under way are not totally surprising.

### *Comparisons among Countries*

Behind the total figures on urban population lies a variety of individual country experiences. Latin America is by far the most urbanized among the developing regions; some two-thirds of its population live in urban centers. In contrast, low-income Asia and Africa are predominantly rural, with average urbanization levels of 25 percent. Intraregional differences are also significant: the urban share is about 83 percent in Argentina and 46 percent in Ecuador; it is 24 percent in India and 12 percent in Bangladesh.

Notwithstanding these differences, some comparisons seem to apply broadly across countries. Descriptively, a clear and well-known association is observed between income level and the percentage of population which is urban. As shown in table 1-1, for countries with low incomes, 21 percent of the total population was urban in 1982. In contrast, 46 percent of the population was urban in middle-income countries, and 78 percent was urban for the higher-income industrial market economies. A broad positive association between income levels and degree of urbanization also seems to hold within regions. (There are, of course, many exceptions to the simple relation between income level and the percentage of population that is urban.)

Furthermore, growth in urban population appears to have a tendency to decline with rising incomes. This tendency is reflected in differences in the rate of urban population growth among countries with different incomes. During 1970-82 urban population grew at 4.4 percent annually in the low-income countries and at a much higher rate if China and India are excluded. In the middle-income countries urban population growth was 4.2 percent annually, whereas for the high-income industrial market economies the urban population growth rate was only 1.3 percent a year.

These regularities in urbanization in developing countries might be expected as part of the development process. Some aspects of present-day concentration may nevertheless be considered especially noteworthy. A tendency often noted is that large portions of the populations of rapidly developing countries tend to be concentrated in one or a few cities. For the low-income countries, on average, 16 percent of the population was in the

largest city in 1980. In the middle-income countries the proportion was substantially higher, 29 percent. In the industrial market economies, however, it was only somewhat higher than in the low-income countries—18 percent. Because of the high representation of rapidly growing countries in the middle-income group, the figures lend some support to the observation that one way in which rapidly developing countries of today differ from countries that underwent development earlier is in the tendency for their urbanization to be more heavily concentrated in large cities. In fact, the overcrowding in one or a few cities, and the visible poverty connected with this phenomenon, are of prime concern; the overall level and rate of urbanization in the country as a whole are less prominent issues.

### *Some Generalizations*

Some relations between urbanization and economic development can be further distinguished with the help of simple descriptive regressions. For sixty-six low- and middle-income economies for which data are available, regressions have been run using the following independent variables: (1) the per capita income rank of the economy, (2) the percentage rate of growth of per capita income, (3) the percentage rate of growth of total population, and (4) zero-one variables that represent the region in which the economy is located. (For example, for an economy in Asia the Asia variable takes on a value of one and the values of all other regional variables are zero for that economy.)

*Degree of Urbanization.* When the independent variables listed above are used in a first-regression equation to explain urbanization, as measured by the percentage of the population that was urban in 1980, a high degree of association is found. The multiple regression coefficient,  $R^2$ , is 0.769. The most significant variable in explaining the percentage of the population that is urban is the income rank. This finding corroborates the positive relation already noted between income and the degree of urbanization. The growth of total population has a significant negative effect in explaining the percentage of the population that is urban. This result may be attributable to slow population growth in higher-income countries, which, in view of the positive income-urban relation, also tend to be the most highly urbanized. The growth of per capita income has a negative, though less significant, coefficient in explaining the urban share of population. The most rapid percentage gains in per capita income occur in countries other than those which already have high levels of income and urbanization. Thus, countries with the most rapidly rising incomes usually have not yet reached the highest

6 *George S. Tolley and Vinod Thomas*

which are discussed in case study detail in later chapters—and points out that urbanization and its accompanying phenomena are products of wider development processes. Care must be taken in the design of measures and policies concerning urbanization to avoid extreme corrective actions that may hamper development.

### *Urban Poverty*

The most widely observed and acutely felt urban problem in developing countries is the large numbers of poor and unemployed people in the cities. The extent of poverty in the economy as a whole depends on the degree of economic development. Urban poverty is viewed as part of overall poverty; the rural poor move to urban areas, which tends to broadly equilibrate real incomes across locations. Poverty is related in great measure to the size of the low-skilled population, which is distributed among urban and rural areas according to economic and other considerations that affect the entire population.

Growth of production leads to growth in demand in urban areas for all factors of production, including unskilled labor. The change in the proportion of unskilled workers in urban areas depends on growth in demand for urban and nonurban commodities, on substitutabilities between unskilled workers and other factors of production, and on the comparative decline of so-called traditional production in urban and in rural areas. Traditional production may be defined as production by firms or households which, although they can modify their output and techniques in response to changes in prices of products and factors, are not viable in the face of competition from modern domestic or foreign manufacturing or from modern farming enterprises. Traditional production usually uses unskilled labor, and its decline may release unskilled labor to other parts of the economy. Although some analysts explain economic development exclusively as a decline in traditional production, development is here viewed as a broader process which involves an accumulation and transfer of knowledge that could and probably would occur even if no units became unviable. The large-scale decline of industries that engage in traditional production may be an important accompaniment of development in some countries—since the decline is accompanied by substitution of capital for labor and by introduction of new techniques—but in other countries this process may be an unimportant detail. Other things being equal, if the decline in traditional production is greater in rural than in urban areas, the proportion of unskilled workers and hence of poor city residents will increase.

Even in the absence of traditional production, development may be concentrated in industries or firms

that initially make up a small part of the economy, most often in the urban sector, in early stages of development. The rise in marginal productivity of all labor, including unskilled labor, may be slow at first and then may accelerate as a greater proportion of the economy develops. As the marginal productivity of labor rises in the course of development, the incomes of those at the low end of the income scale also rise. Furthermore, since development increases the returns to education and other forms of human capital investment, many persons who would otherwise be unskilled transform themselves into skilled laborers and thus directly raise their incomes. The number of unskilled employees falls, and their marginal productivity consequently rises.

Urbanization, which accompanies the strengthening of comparative advantages in urban sectors with development, can be associated with considerable visible poverty and other urban problems, particularly as unskilled workers become displaced from traditional rural activities. Sustained and continued development over a long time would be desirable to ameliorate adjustment problems and increasingly absorb the previously unskilled laborers into the developing economy. The origins of urban problems lie in inadequate and unsustainable development and rural-urban adjustment, but those are by no means the only sources of difficulties.

### *Spatial Concentration and Overurbanization*

Increasing concentration of economic activities and people has been viewed as a reflection both of development and of economic deterioration. As already noted, whereas the generally higher incomes in urban areas are associated with the benefits of urbanization, urban poverty and unemployment and a host of problems associated with pollution and congestion are the most noted indicators of urban failure. Embedded in the urbanization process are elements that represent development and that deserve to be promoted, but urbanization also brings with it externalities which can bring about overcrowding.

Three types of externalities may be associated with overurbanization. First, environmental externalities, such as those connected with pollution and congestion, can mean that cities' sizes are larger than would maximize national (or regional) income and welfare. Second, protected employment that maintains urban wages above market-clearing levels may make cities larger or smaller than they otherwise would be, depending on the elasticity of demand for labor. Third, the attractiveness of urban areas (because of the availability of free or subsidized public services and the advantages of proximity to government activities) can lead to excessive urbanization. Going beyond these urban externalities, a

hypothesis offered here is that the phenomenon of great urban concentration in one or a few cities is connected with the difficulty and relatively high cost of providing intercity transport in developing countries, and that more general infrastructure decisions also play a role.

### Urbanization Policy in Market and Mixed Economies

Some of the differences among today's developing countries that merit attention are emphasized by Renaud in chapter 5. In many of today's developing countries additions to population are larger, income levels are lower, and opportunities to relieve population pressures through migration are more limited than in others. Urbanization in today's advanced economies presents a different set of issues. In some countries there is a process of a slowdown and an eventual end to urbanization and the emergence of yet another industrial revolution based on new technologies which are less tied to concentrated manufacturing centers.

Chapter 5 provides a review of urban-related problems in developing countries—regional inequalities, congestion, pollution, and inadequate provision of urban services—and a threefold classification of related urbanization policies: national economic policies, such as import controls, that have spatial effects; explicit regional development policies, such as investments in infrastructure; and policies concerned with the management of cities. Problems and policy emphasis vary among countries: they are influenced by the age-old structure of cities in Asia, the rapid urbanization which has been going on for some time in Latin America, and the new urbanization in several parts of Africa, which started from a low base and is proceeding at unprecedented rates.

National economic policy sometimes produces significant and unintended effects on urbanization which may outweigh any direct and intended effects of urbanization policies. The unintended spatial biases of national economic policies usually favor some urban centers and are commonly generated through trade policies that protect the manufacturing sector. For example, credit allocation, public investment, and pricing policies may give preferential treatment to economic activities that are concentrated in a few cities and regions. The management practices of the central government and its regulation of economic activities, which make location close to the capital necessary or advantageous for firms, also contribute to the urban vortex.

Explicit regional development policies sometimes attempt to favor decentralization, but they are not always successful or efficient. An unintended by-product

of regulations in India that favored rural over urban development has been the encouragement of the growth of the urban underground economy. The use of water resources for regional development purposes in Malaysia has been dramatic, and policy attention to leakages and regional multipliers has been greater than usual. The Republic of Korea, because of its shortage of land, has an intense interest in land-use guidelines (see chapters 5 and 8), but in spite of prohibitions and incentives, export-oriented growth policies and industrialization ensured that the largest urban centers, especially Seoul and Pusan, would continue to grow. Venezuela and Brazil have pursued industrial deconcentration policies, with limited success (chapters 9 and 13). In addition to policy incentives or disincentives, the balance of labor market considerations, economies of scale within an industry, and the effect of the total size of a city determine location decisions (chapter 7).

The appropriate internal management of cities is important to the success of national spatial policies. Sub-Saharan Africa exhibits acute problems of city management in the face of rapid urbanization, albeit from modest bases. Centralization of decisionmaking in Nigeria, for example, has made it difficult to deal with local problems. In very large cities, policies to limit or stop population growth are not good substitutes for policies that address the urban bias and that directly seek to correct congestion and pollution and provide adequate services. If other cities are not efficiently and effectively managed, their chances of attracting industries and migrants from the largest urban centers will be small. Regional policies can strengthen promising secondary urban centers through well-chosen, cost-efficient actions, better policies for transport investment and management, industrial estates policies, and more important, the systematic development of organized informational networks, such as banking networks, industrial association networks, and better administrative structures, between the secondary cities and the capital region. Good city management and selected regional investments, combined with national economic policies that do not discriminate against rural areas, would do much to alleviate the problems of urbanization.

### Urbanization Policy in a Centralized Economy

Parish (chapter 6) describes how China's socialist leaders, since their coming to power in 1949, have tried to shape the nation's cities to avoid many of the urban problems encountered elsewhere in the developing world. Policymakers planned to restrain the rate of urbanization, since rapid urban growth has often been accompanied by an insufficient number of suitable jobs for new migrants from the countryside. They sought to

avoid the pattern found in other countries of rapid growth of a few large coastal cities at the expense of smaller interior cities. In addition, they tried to narrow the significant gaps in living standards between Chinese cities and rural areas by creating secure jobs and guaranteeing such basic services as health, education, housing, and essential food supplies for both urban and rural populations.

China's history since 1949 brings out both the potential and the disadvantages of urban development in a centrally planned economy. With greater control over economic resources than in the average market society, the government was able to shift investment funds to promote the development of medium over large and interior over coastal cities, which helped reduce regional inequalities. Control over jobs and rationed consumer supplies meant that for a time the government was able to limit severely the growth of all cities, and funds that might have been spent on an elaborate urban infrastructure for new migrants were spent instead on rapid industrial growth.

The government has been able to put the urban population to work. Most able-bodied women, and more than half the urban population, are employed. Few of the jobs are part-time or likely to be lost tomorrow; they are primarily full-time jobs that promise to last for a full career. In these ways Chinese cities have avoided some of the severe problems of unemployment and employment instability that have afflicted other developing-country cities.

Policy mistakes have, however, been made and have led to some of the same problems found in other societies. For instance, it was an error to downplay the role of light industry and consumer services. Some developing-country cities may have too many people in informal service activities, but China illustrates the problems of the opposite extreme. The restricted opportunity for growth of light industry and the bureaucratic restraints on small individual enterprises have fed problems of youth unemployment. Although elimination of urban-rural income gaps was an objective, on balance a pro-urban and proindustrial bias might nevertheless have existed, as in many other countries.

The unemployment problem was heightened by the rapid expansion of employment for women. Jobs were created during the 1960s and 1970s, but since those jobs were taken almost as frequently by women as by men, the need to create additional openings was greater than in other developing societies. Furthermore, the rapid expansion of secondary school education, as in many other developing societies, contributed to growing numbers of unemployed educated persons. China avoided this problem at the university level but not at the secondary level. The unemployed youths were particularly

frustrated because the same level of education had guaranteed good jobs just a decade before, when education was less common. This frustration—shared with youths in many developing countries but perhaps felt more acutely in China because of the socialist promise of secure jobs and rapid development—contributed to the outbreak of petty crime in the 1970s and continues to fuel the social alienation of some youths.

China's experience with the virtual elimination of small informal service activities illustrates the necessity for these types of activities in developing cities. Some argue that because of high population growth rates and the difficulty of centralized provision of essential urban services, the informal service sector in poor societies is essential and should be embraced rather than shunned. China's leaders seem to be moving closer to that view, which entails a greater reliance on market forces, even though they reject its extreme version. Smaller, more makeshift work arrangements, organized ad hoc by neighborhoods and individuals, with lower rates of pay and security, are now approved as a way of providing both employment and essential urban services. Further efforts to improve the relative position of rural activities are being made, and farmers are beginning to narrow the gap between average rural and urban incomes.

### Concentration and Decentralization Policies

At different times and in different countries the growth of large cities has been welcomed and deplored. Governments have used policy instruments to encourage location in major cities or to foster—sometimes even to force—diffusion. Chapters 7, 8, and 9 examine the reasons behind spatial location policies in four countries and the outcomes of those policies.

#### *Encouragement of Concentration: Brazil*

The relative emphasis on concentration and on decentralization in Brazil has changed with the varying fortunes of the economy; concern about concentration has often diminished during times of economic difficulty. Neither concentration nor decentralization may be undesirable in itself; government policies, however, may have intended and unintended effects on industrial location which affect economic welfare. In chapter 7 Henderson examines how government policies have influenced concentration in Brazil, and he attempts to evaluate the desirability of such influences when the existence (or absence) of certain types of economies of scale in urbanization and industrialization is taken into account. The data base for the core of the paper relates to 1970, and much of the discussion therefore concerns the

historical evolution of policies which essentially promoted concentration. Some aspects of the distinct efforts in the 1970s to achieve decentralization are covered in chapter 13.

Localization economies are found to be strong in Brazil, and therefore agglomeration of firms into specialized cities—to take advantage of such benefits as efficiencies in labor markets and in services specific to an industry and greater specialization among firms within an industry—is advantageous. The present results do not show any significant urbanization economies at the scale of activities prevalent in the urban centers of the South and Southeast of Brazil in 1970. The rationale for efforts to encourage industrialization of the largest urban areas rests on the putative net benefits for heavy industries from locating in areas with a large general scale of economic activity. Henderson's findings do not, however, support this rationale. Rather, they indicate that efforts to limit or counter decentralization initiatives may not be desirable.

In addition, negative externalities in the form of environmental degradation could constitute grounds for actually promoting some degree of decentralization. But the size distribution of Brazilian cities is by no means excessively skewed, and efforts to bring about decentralization or a different distribution of city sizes for its own sake may not be warranted. Nevertheless, the provision of more uniform incentives to middle-size cities—which could simply mean the elimination of any special incentives, direct or indirect, for larger cities in the southern region—coupled with environmental restrictions in the highly damaged and built-up areas, could lead to an economically beneficial decentralization of activities.

#### *Industrial Mobility and Decentralization: Colombia and Korea*

Lee (chapter 8) documents the changing patterns of employment location in Bogotá and Cali, the first and third largest cities in Colombia, and summarizes some econometric work on location choices by manufacturing firms. The study also outlines a framework for measuring policy effects and draws policy conclusions, particularly in the context of Korean experience with spatial policy.

The main phenomenon observed is the policymakers' frequent attempts to relocate industries from the traditional industrial districts of large cities to outer areas or to smaller cities. The government's plans may include developing new industrial towns or estates or expanding existing ones to induce new firms or firms that are moving to settle in a desired area. In all cases implementation of such plans and programs requires the selection of particular types of industries to occupy sites

with particular attributes. Hence it is important for policymakers to understand the requirements of firms for attaining equilibrium at new locations and to be able to assess the level and costs of government subsidies and infrastructure investment needed to meet such requirements.

The Bogotá study did not test the effectiveness of explicit policy instruments, partly because such instruments were not implemented in that city. Nevertheless, the behavioral underpinnings established in the study provide clues as to which policy instruments are most appropriate for influencing both the location choices of particular types of firms and the aggregate locational patterns. It is apparent that government policies intended to influence employment locational patterns can be effective if they influence the site attributes which are important to firms.

Lee also examines urban policy in Korea. During the past decade various spatial policies to control the growth of Seoul and to disperse its population have been implemented. For example, in 1971 the greenbelt surrounding Seoul was established. Subsequently the 1977 Industrial Location Act in effect prevented new manufacturing firms from locating within Seoul and enabled the government to issue relocation orders to establishments already set up there. In 1977 the government initiated a ten-year comprehensive plan for redistributing population and industry away from Seoul.

Several other developing countries have tried to decentralize economic activity away from the central city. The economic desirability of decentralization policies has not been established, and not much is known of their effects or their welfare implications. The key policy question is how to guard against spatial policies that are excessive in relation to prevalent decentralization trends, since such measures might result in serious welfare losses. In developing countries the lack of empirical information on decentralization and policy effects does not yet permit the formulation of more efficient spatial policies, but policies to decentralize population and economic activity are probably not good substitutes for better internal management of city growth. For example, the effect on air pollution or on traffic congestion of reducing the population or employment in a large city by a certain amount is likely to be small.

#### *A Decentralization Program: Venezuela*

Reif (chapter 9) uses the case of Venezuela to provide a full-fledged evaluation of decentralization policy. The goals of Venezuelan policy have been to prohibit the location of new manufacturing in Caracas and its surroundings, to induce hazardous industries to relocate

and to encourage others to move to designated development areas, and to attract new manufacturing plants to the designated areas. Policy instruments include direct financial and fiscal incentives, indirect benefits, and negative incentives such as locational control. Reif finds that although little overall deconcentration took place from 1971 to 1978, new firms did tend to leave the country's dominant industrial area.

It remains to be established, however, that the deconcentration was the result of government policy. An analysis of the effects of financial, fiscal, and negative incentives indicates that these had some, though not major, effects on deconcentration. Reif investigates other factors and hypothesizes, among other things, that firms which receive government support benefit from locations near government centers. A series of logit regressions indicates that wages, access to markets, a well-trained work force, union activity, and availability of water significantly affect firm location decisions, in contrast to the relatively less significant performance of government financial incentives. The work shows that strong economic tendencies must be overcome if any appreciable decentralization is to be achieved.

### Addressing Urban Problems

Part IV of this book contains discussions of some of the significant problems that accompany urbanization. Among the most pressing are management of urban fiscal resources, housing, transport, and environmental protection.

### *Management of Urban Finance*

Growing fiscal problems in cities of the developing countries have been caused in part by unprecedented urban growth, for which many countries are ill-prepared, and by the growing demands on local services. The demand for local services is sensitive to increasing population, especially of the poor. Positive income elasticities for public services and demonstration effects from the developed world also influence the demand for local services. Factors that affect the cost of local services include wages, labor unions, rising land and energy prices, inflation, and the increasing costs of borrowing funds for large urban infrastructure investments.

The major revenue sources—taxes, user charges, and external funding—tend to rise less than expenditures during urbanization, and the result is fiscal and service deficits. One solution to fiscal problems would be to

reassign responsibility for urban services from local to central authorities, but Bahl and Linn (chapter 10) do not recommend this approach. An alternative solution is to increase local tax authority and make more use of property taxes and motor vehicle taxes. A third solution, also favored, is to encourage reliance on user charges (which have the advantage of being directly linked to services provided). Equity considerations and social costs (for instance, congestion) may, however, keep user charges low. A fourth solution, to increase fiscal transfers from central to local authorities, has many pros and cons but might be recommended if proper encouragement to cities can be combined with preservation of local autonomy. Reforms in urban financial arrangements in developing countries are proceeding slowly; political factors contribute to the inertia.

### *Housing Policy*

In developing countries, housing is a major consumption category; it constitutes, on average, 15–25 percent of total urban household expenditures. Housing is also one of the most problematic areas in urban development. The most glaring aspect of the problem concerns sprawling squatter areas, delapidated shelter, and appalling lack of basic public services. Other dimensions of the problem are shortages of land, infrastructure services, off-site services, amenities, and employment opportunities and the inadequate supply and rising cost of housing.

For the overwhelming majority, housing is privately provided. Public authorities, however, implement licensing, building codes, zoning, and recordkeeping to ensure clarity in property rights. Governments almost universally have some active interest in housing finance. Overt programs of public housing for low-income groups are common. Housing policy must account for people's desires and should maximize incentives for individuals to expand the housing supply. Public interventions should be limited to areas where the public sector is best suited to perform, such as direct investment, pricing policy, and regulation.

To provide housing requires expertise in many areas. Ingram (chapter 11) stresses the need for understanding the underlying demand for housing. He analyzes housing demand by renters and owners in Bogotá and Cali, Colombia, for 1972 and 1978. Hedonic price coefficients for housing attributes are used to estimate the cost of a standardized unit of housing for each workplace, which is then used as the price variable to estimate a demand function for housing that includes other variables as well. The effect of income on the demand for housing is highly significant, and its elasticity is in the upper end of

the range 0.2–0.8. Price performs less well: its elasticity appears to be less than one. Other variables include age and sex of the household head, family size, and distance to work.

### *Transport Issues*

Transport influences the rate and pattern of urban development and presents multidimensional issues. Infrastructure in the form of roads and streets is publicly provided but entails private use. Urban transport planning requires road layout and in some cases provision for commuter rail transit. These large-scale projects are examples of investments that need rigorous benefit-cost analysis.

The planning of transport and other large urban infrastructure investments is complicated not only by income distribution considerations but also by the difficulty of taking account of the feedback effects of investments on urban development. If provision of services merely follows demand, infrastructure may be a bottleneck to development, and development opportunities may be lost. But if infrastructure leads development and, among other things, fosters new spatial patterns, several problems arise. Infrastructure services that do not closely follow existing patterns carry the risk that demand will not materialize as had been projected. As a minimum, benefit-cost analysis of urban projects might introduce spatial modeling to show probable future residential and business locations and to estimate urban development with and without the proposed project.

Another aspect of the urban transport problem is the provision of public transport, which again involves interactions between public and private decisions. In his analysis of public transport Pachón (chapter 12) points out that private organizations are responsible for most mass transit in Colombia, although local governments grant licenses and route authorizations. Disadvantages can arise from the political allocation of routes; examples are parallel routes, duplication of service, and information problems created by the multiplicity of routes. The system is, however, flexible and appears to be capable of handling the information problem.

Pachón gives an economic rationale for preferring small, less capital-intensive vehicles (such as school buses, minibuses, and collective taxis) over large metropolitan buses. The growth in the number of small buses in Bogotá is a result of lower costs, shorter waiting times, higher trip frequencies, and a greater income elasticity of demand for small-bus service. The desirability of old and of new buses is also relevant; an analysis of the age structure, operating costs, and profitability of Colombian buses is provided. Subsidies that encourage

investments in buses, fare structures that favor new vehicles, and licensing requirements that restrict vehicle stocks are features of the Colombian system.

The dependence on public transport of a large part of the urban population, particularly the poor, heavily influences fares and services and increases government involvement. The Colombian analysis indicates the scope and problems of providing efficient services within this framework.

### *Urban Environment*

Environmental problems are intimately related to urbanization because growing pollution in the developing countries has been caused largely by growth of activity in urban areas. Concern has been heightened as consciousness about pollution has spread worldwide. Thomas (chapter 13) analyzes policies for dealing with the key tradeoff between reducing environmental damage, on the one hand, and, on the other, paying the cost of pollution control and maintaining industrial competitiveness and growth.

In Brazil the capital and operating costs of air pollution control equipment and spare parts can be used to show how costs of pollution abatement vary among producers. Where effluents per unit of output decline with the size of a firm, it would be advantageous to require less than proportionate abatement for larger producers. An offsetting effect, however, is economies of scale in pollution control. The variation in pollution control costs by type and size of firm should be considered in devising policies that will achieve a given amount of environmental control at the lowest possible cost.

The damages to health from the high degree of air pollution in São Paulo are large by any standards and in comparison with the relatively lower risks in outlying areas and in Rio de Janeiro. A regression of the mortality rate on the pollution level, population density, per capita income, hospital beds per person, and percentage of people sixty-five years of age and older yields a positive and significant coefficient for particulates. For example, an annual increase of one ton of particulates per square kilometer in the Greater São Paulo metropolitan area is associated with an increase in mortality of twelve persons per million. The analysis emphasizes that benefits from a given amount of pollution reduction in an area will depend on the size of the population.

The most advantageous tradeoff between reducing environmental damage and maintaining growth can be achieved by a policy that allows industrial and spatial variation in pollution control rather than mandating uniform controls. A promising approach is to use emission taxes or pollution abatement standards that can be

adjusted to conditions in different industries and districts.

### *Evaluating Urban Projects*

The findings of this book point to net benefits from selected policy interventions that directly address selected urban problems. In this connection careful evaluations of intended and actual outcomes of urban programs would be helpful. Part of the work would be better evaluations of urban projects, many of which involve shelter. Keare (chapter 14) shows that rigorous evaluation of projects can assist in improving the efficiency and effectiveness of future endeavors in urban policymaking and in project formulation and implementation. Although much of the paper deals with specific shelter projects financed by the World Bank, the lessons drawn are likely to apply to policymaking as a whole.

The projects are evaluated on the basis of eight criteria: project design, the selection of project beneficiaries, construction methods, materials loan programs, housing completion, occupancy, maintenance of housing infrastructure, and community participation. Given a target group of beneficiaries and a policy objective (in this case, shelter), desirable projects are those that strive for efficient resource use through decentralized decisionmaking. Keare stresses market solutions whenever possible and presumes that project participants are the best judges of their own self-interests. In general, projects should provide participants with suitable locations, secure tenure, and adequate credit but beyond these should leave most decisions to the participants. Advantages and disadvantages are associated with construction projects, self-help requirements, housing standards, rentals, and restricted credit policies. The costs of delayed occupancy and inadequate maintenance and the importance of project cost recovery are stressed in this context.

### Notes

1. Much of the information in this section is from World Bank (1984).

2. These estimates are based on Davis (1972), vol. 2, p. 51, and on more recent data from the World Bank.

### Bibliography

- Andors, Stephen. 1978. "Urbanization and Urban Government in China's Development: Towards a Political Economy of Urban Community." *Economic Development and Cultural Change*, vol. 26 (B), pp. 525-45.
- Berry, Brian J. L., Peter Goheen, and Harold Goldstein. 1971. "Metropolitan Area Definition: A Reevaluation of Concept and Statistical Practice." In L. S. Bourne, ed., *Internal Structure of the City: Readings on Space and Environment*. New York: Oxford University Press.
- Davis, Kingsley. 1965. "The Urbanization of the Human Population." *Scientific American*, vol. 213 (September), pp. 28, 40.
- . 1972. *World Urbanization 1950-1970*. 2 vols. Population Monograph Series 4 and 9. Berkeley, Calif.: University of California Press.
- Hauser, Philip M., and Robert W. Gardner. 1982. "Urban Future: Trends and Prospects." In Philip M. Hauser and others, eds., *Population and the Urban Future*. Albany, N.Y.: State University of New York.
- Linn, Johannes F. 1983. *Cities in the Developing World: Policies for Their Equitable and Efficient Growth*. New York: Oxford University Press.
- Mills, Edwin S. 1972. "Welfare Aspects of National Policy toward City Sizes," *Urban Studies*, vol. 9, no. 1, pp. 117-24.
- Mohan, Rakesh. 1979. *Urban Economics and Planning Models: Assessing the Potential for Cities in Developing Countries*. Baltimore, Md.: Johns Hopkins University Press.
- Renaud, Bertrand. 1981. *National Urbanization Policy in Developing Countries*. New York: Oxford University Press.
- Richardson, Harry W. 1977. *City Size and National Spatial Strategies in Developing Countries*. World Bank Staff Working Paper 252. Washington, D.C.
- United Nations. Population Division. 1977. *Manual 8. Methods for Projections of Urban and Rural Population*. Population Studies 55. New York.
- Wheaton, W. C., and H. Shishido. 1979. "Urban Concentration, Agglomeration Economies and the Level of Economic Development." Working Paper. Cambridge, Mass.: MIT.
- World Bank. 1984. *World Development Report 1984*. New York: Oxford University Press.

# Part I

---

## *Understanding Urbanization in the Developing World*

Part I of this volume contains contributions by Tolley and by Kelley and Williamson that provide a framework for understanding urbanization processes. In chapter 2 Tolley views urbanization as a result of the interplay between supply of and demand for urban goods and services. Kelley and Williamson in chapter 3 adopt a more elaborate general equilibrium model to analyze city growth in the developing world. The two studies focus on fundamental economic reasons for urbanization and offer complementary insights into the process.

### The Determinants of Urbanization

Chapter 2 begins with a simple model that introduces the principal determinants of urbanization and considers increasingly complex and more realistic situations. Income elasticities of demand are higher for commodities produced in urban areas than for those produced in rural areas, and this factor encourages urbanization. Changes in income elasticities are hypothesized to occur as an economy develops and to make for first a rising and then a falling rate of urbanization. Capital accumulation and other reasons for changes in the use of labor affect the number of workers needed to fulfill the demand for urban and nonurban commodities and thus further affect urbanization.

The progression from a closed-economy model to a more general open-economy model of urbanization based on comparative advantage underscores the importance of a neglected reason for urbanization, specialization connected with international and domestic trade. In countries that achieve development by specializing in manufactured commodities, the urbaniza-

tion that occurs during development is intensified. A key factor behind this phenomenon may be that the transfer of production techniques between nations appears to be less costly for urban than for agricultural products. In addition to these factors which affect the proportion of a country's population which is urban, population growth in the country as a whole increases the number of people living in urban areas.

Following the above line of reasoning, a simple model of urban population growth would be based, among other things, on (1) total population growth, (2) urban productivity derived from nonlabor sources, and (3) rural productivity derived from nonlabor sources. An application to sixty-six developing countries reveals the possible effect of those variables. As expected, total population growth contributes to the rate of urbanization, but it accounts for only a 2 percent annual rate, and actual urbanization growth rates range up to 7 percent. Low rates of growth of nonlabor rural income are found to be associated with high rates of urbanization; the reason appears to be sluggish growth in agricultural productivity rather than land scarcity. More vigorous increases in nonlabor inputs in urban areas provide further impetus to urbanization. Relatively stronger productivity growth in urban nonlabor income sources, which arises from the relatively easier transfer of techniques among nations in urban as against rural production, fosters urbanization at a rate faster than overall population growth. Finally, in countries in which the level of urbanization is low—mostly the poorer nations—a given absolute addition to the urban population translates into a higher urban growth rate. The existence of many variations in the general patterns indicates, however, that conditions in individual countries play a role in urbanization.

## Urbanization in Representative Developing Economies

A general equilibrium model is used in chapter 3 to analyze city growth in developing economies. Prices of outputs and inputs are flexible, and capital and labor are essentially mobile. The eight sectors of the model are manufacturing, modern services, informal urban services, low-quality urban housing services, high-quality urban housing services, agriculture, informal rural services, and rural housing services. The nontradable services are important in determining spatial cost-of-living differences that affect migration behavior.

Savings by businesses (from after-tax profits), government, and households are allocated endogenously to investment in physical capital, human capital, and housing. Exogenous variables are foreign-supplied capital, the unskilled labor force, sectoral rates of total factor productivity change, prices of imported raw materials and fuels, and the terms of trade between primary exportables and manufactured importables as influenced by commercial and trade policies.

The results lead the authors to suggest that rapid rates of population growth are not the most important factor in rapid urban growth in developing economies. (This finding is not inconsistent with Tolley's view that overall population growth as such contributes to both urban and rural expansion.) Capital transfers to developing countries do not seem to have been significant. Nor is rural land scarcity necessarily an important influence on urbanization. The findings suggest that fuel prices have been less influential in promoting urbanization than the relative price of manufactured goods. Kelley and Williamson concur with Tolley that productivity advances that favor modern urban activities have been a potent influence on urbanization. They see trade policy in industrial countries and price policies in developing coun-

tries as strong influences on urbanization in the coming decades.

## Conclusion

The studies in part I draw attention to the importance of the macroeconomic development of an economy, trade policy, and relative prices and productivities in urban and rural areas as major influences on urbanization. Although these factors are not always considered in urban policy analysis, they may be highly significant. Macroeconomic policy biases, particularly those that affect trade, industry, and agriculture directly or indirectly, may also have significant impacts on urbanization. Some of these relations are further examined in part II.

## Bibliography

- Beier, George, Anthony Churchill, Michael Cohen, and Bertrand M. Renaud. 1975. *The Task Ahead for the Cities of Developing Countries*. World Bank Staff Working Paper 209. Washington, D.C. Also in *World Development*, vol. 4, no. 5 (May 1976), pp. 363-409.
- Keesing, Donald B. 1979. *Trade Policy for Developing Countries*. World Bank Staff Working Paper 353. Washington, D.C.
- Kelley, Allen C., and Jeffrey Williamson. 1984. *What Drives Third World City Growth? A Dynamic General Equilibrium Approach*. Princeton, N.J.: Princeton University Press.
- Tolley, G. S., and G. K. Kripalani. 1974. "Stages of Development and Urbanization." In A. A. Brown, J. A. Licori, and E. Neuberger, eds., *Urban and Social Economics in Market and Planned Economies*. New York: Praeger.
- World Bank. 1984. *World Development Report 1984*. New York: Oxford University Press.

## *Urbanization and Economic Development*

George S. Tolley

Concerns about urbanization and economic development go hand in hand, for a number of reasons. Beyond the question of whether urbanization is a cause or a consequence of development, it is the backdrop for development in almost all nations. Rapid urbanization poses complex management problems in housing, transport, land use, the environment, energy, public services, and finance. These problems are responsible in part for interest in whether overurbanization has occurred and whether more active policies are needed to control and direct urbanization. To deal adequately with urbanization issues, an understanding is needed of why the process takes place and what determines its rate, its geographic pattern, and the characteristics of people in cities.

That there is a close relation between urbanization and economic development is evident from experience in most countries that have undergone development. It also seems clear that urbanization is not a primary cause of economic development.

In a very long view of history the role of cities in promoting economic development—through being centers of the learning and interactions that lead to innovation—would be a concern. In recent history, as communications have been revolutionized, economic development has come more to consist in applying ideas, no matter where they originate.

It seems reasonable that some, and perhaps most, urbanization is the result of economic development processes. But accepting a general line of causality from economic development to urbanization does not answer

the question of how the causality works. Economic development makes it possible for people to produce more. Why should this possibility lead to urbanization? Can our understanding of the links between development and urbanization be carried to the point of quantification? What are the main implications of a better understanding of urbanization?

In seeking answers to these questions in this chapter it is recognized that the explanation of urbanization is largely an explanation of economic activity. Cities would not grow unless there were productive things for people to do in them.

In the past hundred years enormous growth of cities occurred in the Western world as people took the fruits of development in the form of products, other than food and fiber, that could be produced more economically in cities than in the countryside. Rapid urbanization in developing countries today is in some part a result of a similar wish to take the fruits of development in the form of goods and services suited to urban production.

But more than in the past, urbanization appears to be related to international specialization that is freed from great dependence on domestic demand. People are producing goods in the cities of the developing world and trading them in international markets for all manner of other goods, including food. The links between the composition of domestic consumption and urbanization have been greatly weakened. Actually, the links were never exclusive, and they have become progressively less so throughout the world with the increasing importance of international trade to all economies.

Domestic and international demands for urban and nonurban goods, as just discussed, are the basis for demand influences on urbanization. The ability to pro-

*Note:* Kenneth Miranda provided valuable help in the preparation of this chapter.

duce more urban and nonurban goods acts to increase abundance and is a supply consideration. Population growth results in growth of human inputs and is likewise a supply consideration. Many of the puzzles of urbanization can be understood by analyzing how these demand and supply influences interact. The demand and supply influences provide headings under which urbanization is analyzed in this chapter.

A traditional model of urbanization, appropriate to closed economies, is considered first. Relations between urban and nonurban productivity, income elasticities of demand, and population growth as determinants of urbanization are analyzed. Then urbanization is considered in a more general setting, with international trade. Attempts are made to distinguish between first-order and second-order effects for countries that are price takers in world markets. One of the conclusions from the resulting model, which is applicable to most countries, is that greater urban productivity leads to more urbanization rather than (as in the traditional model) less. Symmetrically, and again contrary to the traditional model, greater nonurban productivity leads to less urbanization.

The model is applied to sixty-six developing economies, leading to explanations of the degree of urbanization in each economy in terms of growth in urban productivity, nonurban productivity, and population. The use of the framework developed in this chapter to further deepen understanding of urbanization in individual countries is discussed, and some important policy implications are noted.

## Demand and Supply Factors

A first step toward an understanding of urbanization is to consider the underlying demand and supply conditions that determine how the resources of an economy are divided between urban and nonurban activities. A first approximation is to associate nonurban activities with agriculture and urban activities with industry and other nonagricultural pursuits. The distinction ignores nonagricultural production by farm households and industry in rural areas, and it begs the question of exactly how *urban* is defined—that is, whether small towns are urban or rural and where the dividing line between urban and nonurban phenomena lies. Although the distinction is not perfect, the discussion here proceeds on the assumption that some degree of approximation is inevitable and that nonurban situations in most parts of the world are overwhelmingly agricultural.

Demand conditions determine how people split their income between urban and nonurban products. Important factors in the allocation of income are the level of

income and the relative prices of the two types of goods. These factors are in turn influenced by supply conditions. The total resources of an economy are devoted to either urban or nonurban activities. The amounts of output of the two types of activities depend on how the available resources are split between them and on production considerations that determine how much of each type of output can be obtained from a given amount of resources.

Over time the total number of persons working in the economy may grow, and for both urban and nonurban activities the amount of output that can be obtained for a given amount of inputs may increase. Given the increases in the number of persons, in the total stock of capital in the economy, and in productivities, the responses to the demand and supply changes that occur during economic growth will determine changes in the allocation of resources between urban and nonurban activities.

The view that urbanization accompanies economic development implies that the demand and supply changes are such that increasing urbanization in fact results. Why is this, and is it necessarily to be expected? The explanation of urbanization has been best explored for the high-income Western countries which have experienced dramatic declines in the proportion of people working in agriculture and in this sense have become the most fully urbanized. The received explanation of this result centers on one element on the demand side and one on the supply side.

### *Income Elasticities of Demand*

The element on the demand side is income elasticity of demand. Because of the low income elasticity of demand for agricultural products—say, 0.10 to 0.20—people are viewed as choosing to spend an increasing proportion of income on urban products as per capita income rises. The low income elasticity of demand for agricultural products has thus been cited as one of the principal reasons why a society devotes more and more of its resources to nonagricultural pursuits and thus becomes more highly urbanized. Since people must spend all their income on either urban or nonurban commodities (where commodities are comprehensively defined to include both consumption goods and the capital goods on which savings are spent), the income elasticity of demand for agricultural (nonurban) commodities can be translated into an income elasticity for urban commodities. If all income elasticities are unity, a 1 percent increase in income leads to a 1 percent increase in expenditures on every commodity. If the income elasticity for some commodities is less than unity, expenditures on them will rise by less than 1 percent and

the income left over will enable expenditures on other commodities to rise by more than 1 percent. Thus, if the income elasticity for nonurban products is less than unity, the income elasticity for urban products will be greater than unity.

The balance of income elasticities above and below unity depends on the relative amounts of expenditure on each kind of commodity. A well-known condition is that the expenditure-weighted sum of income elasticities over all commodities is equal to unity. For example, if 20 percent of expenditures are for nonurban commodities and 80 percent for urban commodities, and if the income elasticity of demand for agricultural products is 0.10, then the income elasticity for urban products is 1.225, implying a 1.225 percent increase in expenditures on urban products for every 1 percent increase in income. To obtain this result, note that the weighted income elasticity condition is  $w_u\eta_u + w_a\eta_a = 1$ , where the  $w$ s are the urban and nonurban expenditure weights and  $\eta_u$  and  $\eta_a$  are the income elasticities. The solution of the income elasticity for urban products,  $\eta_u$ , is  $(1 - w_a\eta_a)/w_u$ , or, with the numbers in the example,  $(1 - 0.2 \cdot 0.1)/0.8$ , or 1.225.

The example is for a high-income country; a different result may be obtained for a low-income country. It is generally thought that the income elasticity demand for agricultural products goes down as income rises. Thus, for lower-income countries a higher agricultural income elasticity implies a lower income elasticity for urban commodities and hence, other things being equal, a lower rate of urbanization. A country could still experience urbanization, but the lower income elasticity for urban commodities might contribute to a lower rate of urbanization in the early stages of development.

In addition to differences in income elasticities (the  $\eta$ s), another difference between high-income and low-income countries—and one that operates in the opposite direction—is the fraction of expenditure devoted to urban and nonurban commodities (the  $w$ s). If the fraction spent on nonurban commodities is high, the weight on the income elasticity for nonurban commodities will also be high. At an early stage of development, when per capita income is low, a large fraction of income is likely to be spent on agricultural commodities. The income elasticity of demand for agricultural commodities will not be as far below unity as in a high-income country and, because its weight is greater, it will tend to make the income elasticity for urban commodities only slightly above unity.

As development proceeds, the gradual decline in the income elasticity for agricultural commodities will act to increase the income elasticity for urban commodities, while the decline in the agricultural weight will act to decrease it. Whether the income elasticity for urban

commodities rises or falls during development depends on which effect predominates.

It is conceivable that the income elasticity for urban commodities is low during the early stages of development, rises to a peak during a middle period, and falls at the later stages. If so, there will be a tendency for the rate of urbanization to start out low, rise to a peak, and fall to a low rate. The first column of table 2-1 shows various values of the income elasticity of nonurban commodities and the second column shows accompanying percentages of expenditures on those commodities that might be observed. The third column shows the income elasticity of demand for urban commodities that results from these combinations, as derived from the formula for  $\eta_u$ , above. The entries in the table progress from the top row, which corresponds to conditions that might be encountered in the initial stages of development, to intermediate situations (the middle of the table), to advanced development (the last row).

In the first row the income elasticity of demand for agricultural products is at its high value, unity. The income elasticity of demand for urban products is then also unity, no matter what the expenditure fractions are. The income elasticity effect would not lead to any relative urbanization, for it would leave the urban proportion of the population unchanged. In the second set of entries the income elasticity of demand for agricultural commodities has fallen to 0.8. Some possible values of the expenditure share accompanying this elasticity are shown. If the share of expenditure on agricultural commodities remains high (80 percent), the income elasticity for urban commodities rises to 1.8; it does not rise as much if the expenditure shares are less. In the next set of entries the income elasticity for agricultural products is 0.5. If the expenditure share has fallen by this time to 50 percent, the income elasticity of demand for urban commodities will be 1.5, with higher and lower values for different expenditure shares. The last row corresponds

Table 2-1. *Determinants of Income Elasticity of Demand for Urban Commodities*

<i>Income elasticity of demand, nonurban commodities</i>	<i>Income spent on nonurban commodities (percent)</i>	<i>Income elasticity of demand for urban commodities</i>
1.0	0–100	1.0
0.8	80	1.8
0.8	50	1.2
0.8	20	1.05
0.5	80	3.0
0.5	50	1.5
0.5	20	1.125
0.1	20	1.225

to the example originally given for a high-income country, with an income elasticity for urban commodities of 1.225.

Table 2-1 suggests that the income elasticity for urban commodities will first rise and then fall during development. If the share of expenditure on agricultural commodities falls only slowly, the income elasticity for urban commodities rises longer and further than if the share of expenditures on agriculture falls rapidly.

In addition to demand conditions within the country, the proportion of the population which is urbanized will be affected by changes in the degree of international specialization that occur during development. A country may develop primarily by increasing its output of manufactured goods for export to the rest of the world. If agricultural output does not increase enough to meet domestic demand, the exports of manufactured goods will be used in part to pay for increasing imports of agricultural commodities. As a result, the increase in production of urban commodities will be larger than if there were no changes in the degree of international specialization, and a smaller fraction of the increased demand for agricultural commodities will be met from production inside the economy. The income elasticity for agricultural commodities must be redefined to pertain only to the percentage increase in domestic agricultural production associated with a 1 percent rise in income, and increases in output of urban commodities will be larger than implied by domestic elasticities. Changes in the degree of international specialization may thus accelerate the rate of urbanization during development.

Changes in international specialization appear to be the rule and not the exception during development. They have been particularly pronounced in such dramatic growth experiences as the "Pacific growth miracles" in which Hong Kong, Japan, Korea, and Taiwan all experienced rapid development simultaneously with increases in output of nonagricultural goods for export. In these economies and in others with similar, although less extreme, growth patterns, concern has arisen over the loss of agricultural self-sufficiency as imports of agricultural products rise along with the rising output of the economy. A more positive way of looking at the decline in agricultural self-sufficiency is to view it as a manifestation of gains from international specialization. It is more economical to produce manufactured goods and exchange them for agricultural products than to forgo the additional manufactured output and try to produce agricultural products that can be produced more efficiently elsewhere.

Although generalizations are hazardous, historically economic development has been almost universally

accompanied by increasing specialization of production. Indeed, the very possibility of economic development has often been based on such specialization. Increasing international specialization is part of a wider process of specialization and provides a reason for increasing urbanization that is related to economic development.

Increasing specialization can occur in agricultural as well as in urban production. It is conceivable that some countries could become less rather than more urbanized during development, although no prominent example comes to mind. A large part of the internationally traded portion of agricultural production comes from high-income countries, which are already highly urbanized. Growth of domestic demand and comparative advantage in nonagricultural production are more important in explaining urbanization than is specialization in international agricultural trade.

### *Supply Factors*

When attention turns from the demand to the supply side, concern shifts from changes in output to changes in the numbers of people used to produce the output. If human inputs were equally substitutable for physical capital in urban and in nonurban production activities, and if the rates of growth of factor productivity in the two types of activities were the same, changes in output and changes in numbers of people would go together exactly, and the demand considerations already discussed would suffice to explain the rate of urbanization during development. But this is not the case. Conditions vary widely, and output per person in production of urban commodities may rise more or less rapidly than for nonurban commodities.

Traditional explanations have emphasized rapid technological change in agriculture, which leads to more rapid growth of output per agricultural worker as the major supply consideration that acts in conjunction with low income elasticity of demand to eject resources from agriculture. Several considerations other than technological change, however, affect output per worker, as will now be discussed.

Land scarcity is overrated as a factor that could decrease output per worker in agriculture. If output did go down, the rate of urbanization would be retarded as more workers were kept in agriculture to produce a given agricultural output. This outcome could occur if growing population pressed seriously against scarce land supplies, but the evidence even for the prime example of a country that might be in these straits—China—is unclear. New land has been coming into production in some parts of the world, notably in Latin America. Improvement of the quality of land through investment

retards declines in output per worker. International specialization has worked against agriculture in many countries and has eased pressure on land supply.

Agricultural research has effects opposite to those of scarce land supply. An increase in output per worker in agriculture permits a given demand to be met with fewer agricultural workers, releases people for urban pursuits, and increases the rate of urbanization. Even if the increase in productivity as a result of agricultural research leads to income growth, workers will be released to urban areas if the income elasticity of demand for the products whose output is increased is less than unity, as seems likely. It is also true that the Green Revolution and other innovations resulting from agricultural research have sometimes been more saving of land and related physical inputs than of workers. The effect of these innovations has in any case been extremely varied among and within countries, so that it is more than usually difficult to generalize about it.

As development proceeds there is a tendency for more of the inputs used in agriculture to be produced in urban areas. This phenomenon is part of the increasing specialization that has already been mentioned and appears to be speeded up by innovations resulting from agricultural research. Gasoline, feed concentrates, and fertilizers and other nutrients are among the inputs produced in urban areas that replace inputs formerly produced on farms. In addition, home garden plots and home food processing decline.

Successful economic development is characterized by rising returns to human inputs, resulting in part from substitution of capital for labor. These changes may affect urban and nonurban production differently. Urbanization will be favored if the substitution of capital for labor is higher in nonurban than in urban production—another of the circumstances that may differ from one country to another. A large source of the substitutions is replacement of labor-intensive human or animal sources of power by machine power, and possibilities for substitutions abound in nonurban as well as urban production. (The mechanization of cotton growing in the United States had wholesale effects on rural-to-urban migration and profoundly affected American cities.) Urban production, which uses readily transferable techniques from other countries, may offer greater possibilities for substitutions, at least in manufacturing industries. But urban production is more intensive than nonurban production in services, where capital appears to be less substitutable for labor.

One of the effects of economic development is the gradual elimination of so-called traditional production in both urban and nonurban areas. Urbanization will be speeded if there is more of this traditional production to

be eliminated in nonurban than in urban areas. Here, again, countries are not alike. Family and home businesses abound in many cities, as do small unviable farming operations in many parts of the world. (Much campesino agriculture is of this type.) The important characteristic of traditional production for the purposes of this discussion is that, in spite of all of its adjustments to relative prices and other changes during development, it is outcompeted by other forms of production. During economic development people are released from the declining traditional forms of production to other activities. If they are released from traditional employment in urban areas and go into other employment in urban areas, there is no effect on urbanization, but if they are released from traditional employment in agricultural areas where there are no expanding employment opportunities, they will seek employment in urban areas, and urbanization will increase.

Because of the low amount produced per person in traditional production, a reduction in traditional employment leads to increased output per person. Whether output per person rises more in urban or in nonurban areas depends on the rate of reduction in traditional production in the two types of areas and on its relative importance in the areas.

Changes in the quality of human inputs through education and other forms of human capital formation deserve mention in relation to urbanization. Urbanization and human capital formation both appear to be responses to factors that lead to economic development, rather than urbanization being a primary cause of human capital formation or vice versa. Like physical capital formation, human capital formation results in increases in output per person and allows a given demand to be fulfilled by fewer people. Given the demands for urban and for nonurban commodities, a higher rate of increase in human capital per person in nonurban than in urban production will be conducive to a high rate of urbanization, since the demand for nonurban commodities can be fulfilled by fewer people. The increase in educational levels and the growth of skilled and professional categories of employment in cities are sometimes thought to be necessary conditions for development. Yet farmers also become better educated. Nonurban production may use a lower amount of human capital per person at the early stages of development, and the level may remain lower through a great part of the development process even though human capital per person in nonurban production is continually rising and may indeed be rising at a faster rate than in urban production. The faster rise in human capital per person in nonurban than in urban production leads to the hypothesis that increases in nonurban human capital could contribute to a high

rate of urbanization. This effect, however, may be secondary to the consideration that both urbanization and human capital formation are basically results rather than causes of economic development.

### Population

The demand and supply considerations discussed so far pertain to the relative rate of urbanization, or urbanization per capita, and would be sufficient to explain urban growth in a country with stable population. Growing total population provides a kicker to both urban and nonurban population growth. The rapid population growth rates in some parts of the world mean that much urbanization comes simply from the fact that there are more people in the country.

### The Traditional Model of Urbanization

The discussion has considered the role of the income elasticities of demand in determining urbanization, the factors that affect output per person in each sector, and the role of total population growth. The following formula brings those factors together into an expression for understanding their combined influence on the rate of urbanization.

$$(2-1) \quad \dot{N}_u = \frac{(\eta_u/\eta_a)\dot{T}_a - \dot{T}_u}{(\eta_u/\eta_a)(N_u/N_a) + 1} + \dot{N}.$$

Since the urban population  $N_u$  equals the proportion urban  $n_u$  times total population  $N$  (or,  $N_u = n_u N$ ), the percentage change in the urban population  $\dot{N}_u$  is the percentage change in proportion urban  $\dot{n}_u$  plus the percentage change in total population  $\dot{N}$  (or,  $\dot{N}_u = \dot{n}_u + \dot{N}$ ), which is the basis of equation 2-1. The variable on the left-hand side,  $\dot{N}_u$ , is the annual percentage rate of growth of the urban population. The first term on the right-hand side is the percentage rate of change in the proportion of the population which is urban. It includes all the considerations discussed earlier except the effect of total population growth of the country, which is given by the second term on the right-hand side,  $\dot{N}$ .

Note that the ratio of the percentage increases in output per capita of urban and nonurban commodities will be in the ratio of their income elasticities. (Recall that the income elasticities are redefined to take account of international trade effects.) This first condition can be written  $\dot{y}_u/\dot{y}_a = \eta_u/\eta_a$ , where  $\dot{y}_u$  is the percentage rate of growth of urban output per capita ( $\dot{y}_u = Y_u/N$ , or urban output divided by the country's total population) and  $\eta_u$  is the trade-adjusted income elasticity of demand for urban commodities. Similar definitions apply to  $\dot{y}_a$  and  $\eta_a$ , where the subscript  $a$  stands for nonurban commodities.

Considerations that affect output per person are brought in by noting that the percentage increase in output per capita for any commodity is equal to the percentage change in output per person engaged in production of the commodity plus the percentage change in the proportion of people in the nation engaged in that production. The second and third conditions are thus  $\dot{y}_u = \dot{T}_u + \dot{n}_u$  for urban commodities and  $\dot{y}_a = \dot{T}_a + \dot{n}_a$  for nonurban commodities, where  $\dot{T}$  is percentage change in output per person engaged in production and  $\dot{n}$  is percentage change in the proportion of people engaged in production. The second and third conditions follow from differentiation of the identity that output equals number of workers times output per worker. The fourth condition is that any change in the urban proportion of the population must be accompanied by an equal and opposite change in the nonurban proportion. In percentage terms the fourth condition is  $\dot{n}_u = - (N_a/N_u)\dot{n}_a$ , where  $N_a$  and  $N_u$  are the numbers of persons in nonurban and urban areas. The four conditions that have been given determine the four variables  $\dot{y}_u$ ,  $\dot{y}_a$ ,  $\dot{n}_u$ , and  $\dot{n}_a$ . The first term on the right-hand side of the above expression for the rate of growth of urban population is the solution of the set of four conditions for  $\dot{n}_u$ , the percentage change in the proportion of the population which is urban.

As can be seen from equation 2-1, the numerator of the term giving the percentage change in the proportion urban is equal to the negative of the rate of growth of output per person in production of urban commodities (that is, an increase in urban productivity, acting alone, leads to a decrease in urban population because the demand for urban commodities can be met by fewer people) plus the ratio of urban to nonurban income elasticities multiplied by the rate of growth of nonurban output per person (growth in nonurban output per person releases persons for urban activities, and the magnitude of the effect depends on the income elasticities). The denominator—one plus the product of the ratio of income elasticities and the ratio of the numbers of people in urban and in nonurban areas—arises because of the interactions of the various effects.

The second term in equation 2-1 is  $\dot{N}$ , the percentage rate of growth of total population of the country. Its presence indicates that when changes in the proportion of the population that is urban are taken into account, the urban population varies in direct proportion to the number of people in the country.

From the formula it is clear that rates of urbanization may differ with conditions in particular countries. As an example consider a country, still at an early stage of development, in which per capita income is rising somewhat rapidly. If the growth is oriented toward growth of manufacturing products that are exported to world mar-

kets, the ratio of the increase in urban and in nonurban output,  $\eta_u/\eta_a$ , will be rather high—say, 4. Suppose that one-third of the population is in urban areas and two-thirds is in rural areas, so that the value for  $N_u/N_a$  is 0.5. Output per worker in urban areas is assumed to be rising at 2 percent a year ( $\dot{T}_u = 0.02$ ). As a result of progress in a commercialized part of agriculture and the decline of subsistence agriculture, output per worker in agriculture is also assumed to rise at the rate of 2 percent a year ( $\dot{T}_a = 0.02$ ). It is assumed that the total population of the country is growing at 1.5 percent a year ( $\dot{N} = 0.015$ ). The values are substituted into the formula for the yearly percentage growth in the urban population.

$$(2-2) \quad \dot{N}_u = \frac{4 \cdot 0.02 - 0.02}{4 \cdot 0.5 + 1} + 0.015 = 0.035.$$

That is, the annual growth rate for the urban population is 3.5 percent.

Examples of much more rapid or of slower urbanization can be constructed by varying the values. The purpose of this section, however, has not been to explain the rate of urbanization for any one country but to begin to suggest the nature of the underlying changes in demand and supply that determine urbanization.

### A More General Explanation

The discussion points to generalizations about why worldwide urbanization is occurring. One reason is that incomes are rising in many places, albeit in varying degrees. The higher income elasticity of demand for urban than for nonurban commodities is a fundamental reason why urbanization can be expected to accompany economic development throughout the world.

Another reason for world urbanization is more closely related to the nature of the economic development that is occurring. For much of the world economic development entails a transfer of techniques of production, and by and large the techniques are more transferable for urban than for agricultural commodities. Techniques of agricultural production are much more tied to local land and climate conditions. Efforts are being made to develop and adapt agricultural techniques to suit the unique conditions of each area of each country, but the pace, at best, is slow. Moreover, agriculture, as an inherent user of location-specific resources, is subject to intrinsic comparative advantages. Increasing specialization of agricultural production among major regions of the world may be in store.

Urban production, on the other hand, is not nearly as location-bound. Factories, machines, and electronics parts tend to be physically similar the world over in their requirements for productive inputs. It is not surprising

that developing countries are increasingly getting out of the agricultural business and into urban-based production.

Although the model of the preceding section allows for the effects of trade in the specification of income elasticities, it is not fully consistent with the general reasons for rising urban population connected with specialization in production noted above. In equation 2-1 a rise in agricultural productivity will increase urbanization because a given demand for agricultural products can be satisfied by fewer people and the excess people move to the cities. In like fashion, a rise in urban productivity acts to decrease urbanization, since fewer people are needed to satisfy the demand for urban products. The amounts of labor required for production in both sectors change according to a life of their own, and amounts of output are not affected by productivity changes.

Yet one might think that if a sector became more efficient, it would increase its production, not decrease it. An increase in efficiency lowers the price at which a given quantity can be produced. If the price elasticity of demand is greater than zero, a lowering of the price will increase the amount that can be sold. A production response is to be expected.

Indeed, if the responses to changes in production take the form of adjustments in foreign exports and imports, prices are effectively set in world markets where demands are highly elastic and where in many cases prices can be taken as given. Even within an economy, a more adequate view of the differences between urban and rural production will recognize that there is not a strict dichotomy in the types of goods produced. Aside from local or nontraded goods in each sector, there are goods that may be produced in either sector and that can be traded between the two sectors. There may be little farm output in urban areas (although chickens and garden plots can be significant in cities), but factory production and home production of urban-type goods abound in rural areas, and trade between sectors in processed foods and other commodities is commonplace. For most of the world it may be that most consumption in either rural or urban areas is of goods produced within each area and that trade in marketed surpluses of rural and urban goods is confined to a small part of output, much of which can actually be produced in either sector.

These considerations of the way trade occurs suggest that it may be more reasonable to take prices facing rural and urban areas as given rather than to take relative quantity growth as given, as in the model presented above. If, as a first approximation, prices are taken as given, productivity changes will have the opposite effect from that implied by the model. A rise in productivity in urban areas enhances cities' ability to compete in inter-

national markets and to compete with production in rural areas and thus leads to an increase in urban production, not a decrease. An analogous result is obtained in nonurban areas, with the implication that rising nonurban productivity will retard urbanization. The results are consistent with the hypothesis that urbanization results from changing advantages in production.

Given that a rise in productivity leads to an expansion rather than a contraction of sector output, the question becomes how great the expansion will be. Since relative quantities of output change are endogenous rather than exogenous, a careful distinction is needed between changes in average labor productivity and changes in marginal labor productivity as affected by diminishing returns to employment in a sector. If there were no diminishing returns as employment increased in a sector, the assumption of given output prices (suggested here as a first approximation) would imply that, except by accident, one sector would have a higher marginal productivity of labor than the other. Migration would then lead to all the employment being in one sector or the other—one sector would go to zero and the other would have all the employment.

To avoid this unrealistic solution the assumption of given prices might be dropped, leading to a solution governed by changing prices of urban goods in relation to those of nonurban goods. But the idea that urbanization is primarily explained by massive changes in prices of urban goods in relation to those of nonurban goods—in particular, by increases in relative prices of urban goods—is not appealing. This type of change does not appear to be a pervasive world phenomenon. Endogenous relative prices can be introduced, but it would be best to do so only after a situation of given relative prices has been more fully explored.

Another way to avoid the unrealistic solution of having all employment end up in one sector or the other would be to base a solution on impediments to labor migration. Then a rise in, say, urban compared with nonurban labor productivity would increase the attractiveness of urban employment. Imperfect labor mobility would lead to a rising gap in real incomes between urban and rural sectors that might be reduced only gradually over time. One problem with this solution is that it does not avoid the outcome that all employment ends up in one sector; it only delays that outcome. Eventually, over generations, if not earlier, people all migrate to the cities in response to relative differences in earnings opportunities. The only reason for not doing so would be nonpecuniary attachments to rural life. Basing an explanation of urbanization on permanent nonpecuniary attachments to a way of life is at least as unappealing as basing it on changing relative output prices of the sectors.

If one has to choose between no labor mobility and a

degree of mobility that leads to rough equalization of opportunities for real earnings between sectors, the latter appears far more defensible. One is likely to find that, historically, real earnings of comparable labor in urban and rural areas have changed by about the same amount, and the degree of net and gross migration observed between urban and rural areas is so great as to belie the notion that responses do not occur. The single most important distinguishing feature in the relations between the urban and rural parts of an economy may be that they are connected by migration. If one is not to base an explanation of urbanization on a growing disparity between real earnings of comparable labor in urban and in rural areas, it is reasonable to suppose that there is sufficient labor mobility to roughly equalize the urban and rural changes in marginal productivity that occur during economic development. The implications of lesser degrees of mobility could be considered, but this can be done more easily and with more insight after the first approximation is explored.

In distinction to the variable  $T$  that was used above to refer to average labor productivity, let  $A$  denote the influence of nonlabor inputs on the amount of output (taken to be synonymous with income produced). For nonurban production, the availability of land may partly determine  $A$ . Port facilities and other natural limitations could play a similar role in urban production. For both urban and nonurban production, private capital and the amounts of various types of infrastructure influence  $A$ . With labor  $N$  measured in terms of numbers of people or hours,  $A$  is influenced by the productive quality per person, which in turn is determined in part by human capital investment. For both urban and nonurban production, perhaps the greatest influence on  $A$  is technical knowledge used in production, which depends on encouragement of or impediments to international transfer of technology, entrepreneurship, incentive structures, and other policy and institutional considerations.

To obtain a formulation consistent with growth source analysis, take time derivatives of the production relation  $Y = AN^\epsilon$  and express as yearly percentage changes. Then growth in output will equal the percentage growth in nonlabor influences,  $\dot{A}$ , plus the product of elasticity of output with respect to labor,  $\epsilon$ , and the percentage change in labor. For urban production,

$$(2-3) \quad \dot{Y}_u = \dot{A}_u + \epsilon \dot{N}_u$$

and for nonurban production,

$$(2-4) \quad \dot{Y}_a = \dot{A}_a + \epsilon \dot{N}_a.$$

The elasticity of output,  $\epsilon$ , is almost certainly between zero and one and is closer to one than to zero. If it were one or greater, the implication would be that—other inputs being held constant—output could be increased

in equal or greater proportion to increases in labor. Under the assumption of fixed proportions the elasticity would be zero. That assumption, however, ignores the myriad of possibilities for drawing on a given amount of nonlabor inputs to produce varying physical quantities and qualities of outputs in modern factory production, office work, retailing, personal services, and traditional production as the amount of labor is varied. The relation between average labor productivity  $T$  in the preceding section and the influence of nonlabor inputs  $A$  in this section can be seen by combining the conditions  $T = Y/L$  and  $Y = AL^\epsilon$  to obtain  $T = AL^{\epsilon-1}$ . If  $\epsilon$  is close to one and if nonlabor influences  $A$  are more predominant as causes of change in average labor productivity than the elasticity effect of changes in labor inputs  $L^{\epsilon-1}$ , then  $T$  and  $A$  will move in a similar fashion.

To proceed to the rest of the model in this section, suppose that labor is paid a wage  $w$  that is equal to its marginal product  $m$  multiplied by price output  $p$  (or,  $m = w/p$ ). The elasticity of output (the marginal product  $m$  multiplied by the amount of labor  $n$  divided by output  $y$ ) is seen to be equal to wage payments  $wn$  divided by the value of output  $py$ . This implies that  $\epsilon$  in equations 2-3 and 2-4 equals the share of product paid to labor,  $wn/py$ . The order of magnitude of  $\epsilon$  as labor's share can in principle be observed rather readily. Although the elasticity of output may be different in urban and in nonurban production, in this example it is taken to be the same. Not to do so would imply that major reliance is placed on differences in factor shares in explaining urbanization.

The rate of urbanization may indeed be affected by urban-rural differences in elasticity of output as well as by changes in elasticities in particular countries. There are no compelling studies that indicate systematic differences and changes in factor shares, and the shares do seem in fact to be fairly stable. The formula for elasticity of output followed here therefore seems appropriate for an overview of major factors that explain rate of urbanization.

Given observed values of  $\epsilon$ , equations 2-3 and 2-4 can be used in connection with measurements of income growth  $\dot{Y}$  and population growth  $\dot{N}$  to obtain estimates of the growth of nonlabor influences,  $\dot{A}$ . When either equation 2-3 or equation 2-4 is solved for  $\dot{A}$ ,  $\dot{A}$  is found to be  $\dot{Y}$  minus  $\epsilon\dot{N}$ . The estimate of growth of nonlabor sources of income in a sector is the observed income growth less that part of income growth attributable to labor, which in turn is equal to labor's elasticity of output multiplied by the population growth of the sector.

An essential feature of the model is that labor will adjust to keep the relative earnings of urban and rural employment in line. To incorporate this feature, the condition pertaining to the elasticity of output derived

above,  $\epsilon = wn/py$ , can be rearranged as an expression for the wage in a sector,  $w = \epsilon py/n$ . With the use of rates of change for the two sectors,

$$(2-5) \quad \dot{W}_u = \dot{P}_u + \dot{Y}_u - \dot{N}_u$$

and

$$(2-6) \quad \dot{W}_a = \dot{P}_a + \dot{Y}_a - \dot{N}_a$$

assuming that  $\epsilon$  is a stable parameter ( $\dot{\epsilon} = 0$ ). The  $P$ s in these conditions refer to prices of output produced; the relevant deflator that determines labor supply response is an index of prices of goods consumed. The level of prices may differ greatly between urban and rural areas. If prices move proportionately over time and relative real earnings in urban and rural areas remain roughly the same, the rate of growth of wages in the two sectors will be similar, which yields the condition  $\dot{W}_u = \dot{W}_a$ . The right-hand sides of equations 2-5 and 2-6 can then be set equal to each other.

If the approximation is retained that changes in prices of urban-produced goods relative to those of rural goods will not in the first instance be used as an explanation of the rate of urbanization,  $\dot{P}_u = \dot{P}_a$ . Then the condition obtained from setting the right-hand sides of equations 2-5 and 2-6 equal to each other is

$$(2-7) \quad \dot{Y}_u - \dot{N}_u = \dot{Y}_a - \dot{N}_a$$

This condition has the strong empirical implication that the excess of the rate of income growth over the rate of employment growth will be the same in urban and in nonurban areas. Put another way, average labor productivity will change by the same proportion in both sectors.

Income produced in the two sectors may change by greatly different amounts—that is,  $\dot{Y}_u$  and  $\dot{Y}_a$  may be very different. The differential changes in income produced are a result of differences in growth of nonlabor sources of income, which lead to corresponding differences in growth of demand for labor. For example, if nonlabor sources of income growth increase more rapidly in urban than in rural areas, the demand for labor will increase more in urban than in rural areas. The increase in nonlabor sources of income and the associated increase in labor inputs will raise income by  $\dot{Y}_u$  and employment by some lesser amount,  $\dot{N}_u$ . Meanwhile, because of the slower growth of nonlabor income sources in rural production,  $\dot{Y}_a$  will be smaller than  $\dot{Y}_u$  and  $\dot{N}_a$  will be correspondingly smaller than  $\dot{N}_u$  (owing to a lesser growth in demand for labor in rural areas). If factor shares and relative wages and prices are not to change, the rise in ratio of income to employment must be the same in the two sectors.

How well does the condition stated in equation 2-7, that average labor productivity will change by the same

proportion in urban and rural employment, fit the facts? For sixty-six low- and middle-income economies for which the requisite income and population figures are available to study growth rates for 1960–80, the mean of the excess of the proportionate change in average urban labor productivity over average rural productivity is  $-0.0001$ . The standard error of this mean is  $0.0029$ . In these comparisons real nonagricultural and agricultural incomes are used as surrogates for urban and rural incomes. It is assumed that employment is roughly proportional to population, so that urban and rural populations can be used as surrogates for employment.

Alternatively, still with nonagricultural and agricultural incomes as surrogates for urban and rural incomes, the labor force in agriculture can be used as a surrogate for rural employment and the nonagricultural labor force as a surrogate for urban employment. Or, more simply, the comparison may be viewed as an application of the model to agricultural versus nonagricultural employment rather than to urban versus rural employment. If agricultural and nonagricultural labor forces are used, rather than rural and urban populations, the mean of the excess of the proportionate change in average nonagricultural labor productivity over that of average agricultural productivity is  $0.0024$  and the standard error is  $0.0029$ .

To further test the extent to which equation 2-7 is satisfied, if the error attributable to the fact that observed changes in average productivity do not exactly fulfill the equation is small, most of the observed changes in output and employment will be explained by the model. The small size of the error in relation to total income or employment change is an indication of the extent to which the hypothesis of equal proportionate changes in average labor productivity is fulfilled.

An error-of-fit measure, calculated for each economy, is the proportionate annual change in average labor productivity as a fraction of the mean of annual proportionate urban and rural population growth. The median of the error-of-fit measures over all sixty-six economies is  $0.030$ . Two-thirds of the measures fall between  $-0.735$  and  $0.826$ . When nonagricultural and agricultural labor forces are used instead of urban and rural populations, the median error-of-fit measure for the sixty-six economies is  $0.067$ . Two-thirds of these measures fall between  $-0.981$  and  $0.828$ .

The fact that most of the error-of-fit measures are substantially less than 1 indicates that most of the observed variation in output and employment is associated with growth of nonlabor income sources and the hypothesized proportionate changes in labor and that relatively little change is left to be explained by errors or departures from the hypotheses of the model. The existence of nonzero error of fit may be partly a result of

errors or differences in the way variables are calculated from available data. The uncertainty of output and employment figures is well known, and the reasons for differing changes in measured employment and output connected with changing market orientation of production during development have often been discussed. In general, more production may pass through the marketplace as economic development proceeds over a twenty-year period, but because of lags in data collection, statistics may capture these changes only imperfectly.

Nonzero error of fit may also occur because assumptions of the model that are justified as first approximation may not be exactly fulfilled. But the small size of the errors suggests that the principal factors are indeed captured by the model. In view of all the possible reasons for existence of error, it is fair to say that the data fit the model well.

It was specified that growth in output can be represented as the sum of labor and nonlabor effects in each sector in equations 2-3 and 2-4. Condition 2-7 (for changes in employment relative to output in each sector as migration occurs in response to changes in demand) was then derived. The model is completed by using the condition discussed above—that the total amount of labor used in the economy is the sum of urban and nonurban labor, or  $N = N_u + N_a$ . The equation is differentiated with respect to time and divided by the original values of the variables. The annual proportionate rate of growth of the total population is expressed as the product of the urban share of population and annual proportionate rate of growth of the urban population, plus the product of the nonurban share and the annual proportionate rate of growth of the nonurban population; that is,

$$(2-8) \quad \dot{N} = F_u \dot{N}_u + F_a \dot{N}_a.$$

The four-equation model given by formulas 2-3, 2-4, 2-7, and 2-8 determines the four rates of growth that pertain to output and employment in the two sectors, that is,  $\dot{Y}_u$ ,  $\dot{Y}_a$ ,  $\dot{N}_u$ , and  $\dot{N}_a$ . To solve for the variable of interest—urban population growth,  $\dot{N}_u$ —equations 2-3 and 2-4 are substituted into equation 2-7 to eliminate the  $Y$ s. The result is a relation between  $\dot{N}_u$  and  $\dot{N}_a$  that depends on the rates of growth of nonlabor sources of income in the two sectors. The relation can be written  $\dot{N}_a = \dot{N}_u + (\dot{A}_u - \dot{A}_a) / (1 - \epsilon)$ , indicating that urban and rural populations move proportionately except as they are influenced by differential changes in growth of nonlabor sources of income, which have a multiplier effect of  $1 / (1 - \epsilon)$  because under the model enough labor must be hired to drive the return to labor down to the same level in each sector. The result for  $\dot{N}_a$  is inserted into condition 2-8 for total employment and the equation is solved for  $\dot{N}_u$ :

$$(2-9) \quad \dot{N}_u = \dot{N} + \left( \frac{F_a}{1-\epsilon} \right) \dot{A}_u - \left( \frac{F_a}{1-\epsilon} \right) \dot{A}_a.$$

The solution of the model for  $\dot{N}_u$  given by equation 2-9 provides the basis for decomposing the rate of growth of the urban population into three factors. The first is total population growth of the nation,  $\dot{N}$ , which has a proportionate effect on urban population growth. The second is an urban productivity factor that stems from growth of nonlabor sources of urban production. The positive effect of increases in nonlabor sources of urban production is enhanced by the multiplier  $1 / (1 - \epsilon)$  that is connected with keeping earnings growth the same in the two sectors, and the magnitude of the effect is also proportional to the fraction of the population in agriculture. A given change in the absolute level of employment as labor is reallocated will have a greater proportionate effect on urban population the smaller is the existing fraction of the population that is urban (which is to say, the larger the fraction that is nonurban). The third factor into which urban population growth is decomposed is a rural productivity factor which stems from growth of nonlabor sources of rural income. It is exactly analogous to the urban productivity factor, but it acts on urban population growth with a negative sign, since in this model a rise in rural productivity increases rural population.

#### *Application to the Sample*

Measures of the rate of population growth  $\dot{N}$  to be applied to equation 2-9 are readily available. The fraction of the population that is not urban,  $F_a$ , is available for many countries but may be subject to more error than total population. Systematic measures of labor's elasticity of output  $\epsilon$  that are comparable among sectors and nations are apparently not available, but impressions about general orders of magnitude can be obtained. Although the growth in nonlabor sources of income (the  $\dot{A}$ s) cannot be directly observed, first-cut measurements can be obtained with the use of equations 2-3 and 2-4. That is, when either equation is solved for  $\dot{A}$ , the contribution of growth in nonlabor sources of income to observed income growth is equal to the observed income growth  $\dot{Y}$  minus the product of elasticity of output of labor  $\epsilon$  and the growth rate of labor  $\dot{N}$ . (That is, the nonlabor source of income growth is total income growth minus the contribution of labor to income growth.)

The calculations just described give only first-cut or, at most, suggestive measurements of the  $\dot{A}$ s because they ignore the stipulation of the model that average labor productivity rises at the same rate of growth in both sectors. The measures are dependent on the esti-

mates of growth of urban and rural income, which may be subject to great measurement errors and possible inconsistencies in definition. To minimize the effects of measurement problems, the procedure used here is to average the changes in income in the two sectors and thus obtain income measures that conform to the productivity condition. Specifically, the average excess of the rate of growth of income in each sector over the rate of population growth in each sector yields an estimate of the common rise in average labor productivity in the two sectors. This estimate made on the basis of the observed average for the two sectors can be added to population growth in each sector to obtain the measures of income growth in the sectors that would prevail if the model fitted the real situation precisely. After this attempt to purge the income measurements of error and to eliminate the effects those errors would spuriously introduce into estimates of the factors that affect urbanization, estimates of the  $\dot{A}$ s may be obtained as before by subtracting the product of labor elasticity of output and the amount of labor from the income values that conform to the average labor productivity condition.

Table 2-2 presents the results obtained when the model is applied to sixty-six low- and middle-income economies, using urban population statistics from the World Bank and an estimate of labor's share  $\epsilon$  of 0.7. Column 1, the annual proportionate growth of urban population in each economy from 1960 to 1980, is the variable to be explained. Column 2 is the average fraction of the population that is urban for 1960-80. This fraction was calculated to conform to equation 2-8, which relates the rates of urban and rural population growth to the rate of growth of the total population. The fraction of the population that is rural,  $F_a$ , is  $1 - F_u$ . Arrangement of equation 2-8 as an expression for  $F_u$  gives as the fraction of the population that is urban  $F_u = (\dot{N} - \dot{N}_a) / (\dot{N}_u - \dot{N}_a)$ , the ratio between the excess of total population growth over nonurban population growth and the excess of urban over nonurban population growth. This measure is quite close to the simple average of the 1960 and 1980 values for the fraction of the population that is urban, and in all cases it lies between those values. The method just described for calculating  $F_u$  in conformance with equation 2-8, along with the method described above for calculating the growth of nonlabor income sources  $\dot{A}$  to conform to equations 2-3 and 2-4, ensures an internally consistent system of accounting in which the factors that explain urban population growth exactly add up to urban population growth. Although other procedures for defining variables might have given about the same results, the procedures used here have the advantage of eliminating idiosyncracies in definition as a perceived reason for differences in rate of urbanization. The analytic factors that explain urban population

Table 2-2. *Urban Population Change, Sixty-six Economies, 1960-80*

Economy	Annual proportionate change in urban population, $\dot{N}_u$ (1)	Urban share of population, 1980, $F_u$ (2)	Annual proportionate growth in urban nonlabor sources of income, $\dot{A}_u$ (3)	Annual proportionate growth in nonurban nonlabor sources of income, $\dot{A}_a$ (4)	Components of change in urban population			Error of fit, $(\dot{A}_u - \dot{A}_a) / \frac{1}{2}(\dot{N}_u + \dot{N}_a)$ (8)
					Annual proportionate change in total population, $\dot{N}$ (5)	Urban productivity factor, $F_u \dot{A}_u / (1 - \epsilon)$ (6)	Nonurban productivity factor, $-F_a \dot{A}_a / (1 - \epsilon)$ (7)	
Bangladesh	0.066	0.076	0.006	-0.007	0.027	0.017	0.022	-0.413
Ethiopia	0.070	0.099	0.016	0.001	0.024	0.048	-0.002	-0.171
Mali	0.054	0.151	0.016	0.005	0.024	0.045	-0.015	0.196
Somalia	0.052	0.231	-0.025	-0.036	0.024	-0.065	0.093	1.527
Chad	0.067	0.118	-0.010	-0.026	0.020	-0.030	0.077	-1.101
Burma	0.039	0.228	0.019	0.012	0.022	0.048	-0.030	-1.193
Burkina Faso	0.045	0.068	-0.001	-0.010	0.016	-0.002	0.032	0.828
India	0.032	0.199	0.019	0.015	0.022	0.051	-0.041	0.567
Malawi	0.070	0.062	0.028	0.016	0.029	0.089	-0.049	-0.249
Rwanda	0.062	0.029	0.018	0.007	0.027	0.059	-0.024	-0.096
Sri Lanka	0.041	0.223	0.019	0.012	0.021	0.050	-0.030	-0.315
Guinea	0.061	0.141	0.013	0.002	0.029	0.038	-0.006	-0.190
Zaire	0.060	0.242	-0.006	-0.021	0.023	-0.016	0.053	-0.176
Niger	0.070	0.091	-0.001	-0.014	0.031	-0.003	0.042	-0.554
Benin	0.044	0.119	0.007	0.001	0.027	0.021	-0.004	0.633
Pakistan	0.041	0.249	0.030	0.025	0.029	0.074	-0.062	0.442
Tanzania	0.072	0.080	0.023	0.009	0.028	0.070	-0.026	0.004
Central African Republic	0.051	0.315	0.007	-0.005	0.022	0.017	0.012	-0.747
Haiti	0.055	0.247	0.004	-0.012	0.016	0.009	0.030	-1.060
Mauritania	0.128	0.102	0.009	-0.025	0.026	0.026	0.076	-0.313
Lesotho	0.067	0.033	0.039	0.025	0.021	0.126	-0.080	0.969
Uganda	0.077	0.080	0.006	-0.008	0.034	0.018	0.026	-0.724
Sudan	0.070	0.166	0.003	-0.013	0.024	0.009	0.037	-0.027
Togo	0.053	0.132	0.034	0.025	0.026	0.098	-0.071	0.671
Kenya	0.070	0.101	0.017	0.005	0.035	0.051	-0.016	-0.246
Senegal	0.029	0.240	0.010	0.008	0.025	0.025	-0.021	-0.538
Indonesia	0.035	0.174	0.037	0.031	0.020	0.101	-0.087	0.826
Egypt	0.031	0.415	0.037	0.033	0.022	0.073	-0.065	1.046
Ghana	0.049	0.292	-0.001	-0.011	0.026	-0.002	0.025	-0.735
Honduras	0.054	0.292	0.017	0.007	0.032	0.040	-0.017	-0.473
Zambia	0.054	0.301	0.019	0.008	0.029	0.045	-0.020	-1.449
Thailand	0.033	0.135	0.046	0.045	0.029	0.133	-0.129	1.101
Bolivia	0.042	0.284	0.025	0.018	0.026	0.059	-0.043	-0.030
Philippines	0.038	0.330	0.032	0.028	0.029	0.071	-0.062	0.009
Congo	0.028	0.350	0.004	0.001	0.023	0.008	-0.002	1.138
Nigeria	0.047	0.163	0.024	0.016	0.025	0.066	-0.045	1.500
El Salvador	0.036	0.395	0.022	0.020	0.033	0.043	-0.040	0.495
Morocco	0.044	0.348	0.021	0.013	0.027	0.045	-0.028	0.260
Peru	0.047	0.567	0.027	0.014	0.028	0.039	-0.020	-0.149
Côte d'Ivoire	0.080	0.278	0.032	0.018	0.045	0.077	-0.042	0.009
Nicaragua	0.043	0.470	0.041	0.034	0.031	0.073	-0.060	-0.494
Colombia	0.045	0.593	0.040	0.026	0.026	0.054	-0.035	-1.148
Paraguay	0.031	0.375	0.034	0.032	0.027	0.071	-0.067	0.077
Dominican Republic	0.055	0.402	0.032	0.019	0.029	0.064	-0.037	0.044
Guatemala	0.037	0.360	0.035	0.031	0.028	0.075	-0.066	-0.057
Syrian Arab Republic	0.047	0.434	0.048	0.040	0.032	0.091	-0.076	-0.384
Malaysia	0.036	0.270	0.042	0.039	0.028	0.103	-0.095	0.652
Jamaica	0.034	0.419	0.015	0.005	0.015	0.029	-0.010	-1.283
Korea, Republic of	0.056	0.410	0.064	0.047	0.022	0.125	-0.092	0.334

Table 2-2 (continued)

Economy	Annual proportionate change in urban population, $\dot{N}_u$ (1)	Urban share of population, 1980, $F_u$ (2)	Annual proportionate growth in urban nonlabor sources of income, $\dot{A}_u$ (3)	Annual proportionate growth in nonurban nonlabor sources of income, $\dot{A}_a$ (4)	Components of change in urban population			Error of fit, $(\dot{A}_u - \dot{A}_a) / \frac{1}{2}(\dot{N}_u + \dot{N}_a)$ (8)
					Annual proportionate change in total population, $\dot{N}$ (5)	Urban productivity factor, $F_u \dot{A}_u / (1 - \epsilon)$ (6)	Nonurban productivity factor, $-F_a \dot{A}_a / (1 - \epsilon)$ (7)	
Turkey	0.047	0.383	0.036	0.026	0.025	0.075	-0.053	0.301
Algeria	0.065	0.452	0.002	-0.017	0.030	0.004	0.031	-0.195
Mexico	0.046	0.592	0.029	0.018	0.033	0.039	-0.025	0.122
Panama	0.044	0.475	0.034	0.026	0.030	0.060	-0.046	0.012
Taiwan	0.038	0.680	0.057	0.044	0.024	0.061	-0.046	1.408
Chile	0.028	0.749	0.027	0.016	0.019	0.022	-0.014	-2.719
Costa Rica	0.038	0.400	0.033	0.039	0.030	0.067	-0.059	0.276
Brazil	0.046	0.556	0.042	0.030	0.028	0.062	-0.045	0.260
Uruguay	0.010	0.821	0.016	0.011	0.008	0.009	-0.007	-5.082
Argentina	0.019	0.782	0.029	0.022	0.014	0.021	-0.016	-0.914
Portugal	0.018	0.269	0.029	0.023	0.003	0.070	-0.055	5.940
Yugoslavia	0.030	0.348	0.042	0.033	0.010	0.091	-0.071	0.430
Venezuela	0.044	0.756	0.042	0.029	0.034	0.035	-0.024	-1.436
Hong Kong	0.023	0.895	0.044	0.042	0.023	0.015	-0.015	2.877
Greece	0.024	0.526	0.054	0.042	0.006	0.085	-0.066	-0.601
Spain	0.024	0.658	0.048	0.037	0.011	0.055	-0.042	0.894
Israel	0.039	0.836	0.050	0.037	0.031	0.027	-0.020	-0.642

Note: Economies are listed in increasing order of 1978 per capita income.

growth in fact add up exactly to measured urban population growth.

Columns 3 and 4 present measures of the average annual rate of growth of the sources of nonlabor urban and nonurban income  $\dot{A}$ , with the use of the purged income values, as described above.

The key columns are 5, 6, and 7. They show the decomposition of the annual proportionate urban population growth given in column 1 into the three terms in equation 2-9. Thus, column 5 is the annual proportionate overall population growth of the economy, which makes a direct and equal contribution to urban population growth. Column 6 is the urban productivity factor, which gives the effect of nonlabor urban income source growth on the rate of urban population growth. Column 6 is column 3 multiplied by the nonurban fraction of the population (calculable from the table as 1 minus the urban fraction in column 2) and divided by the share of income that accrues to factors of production other than labor,  $1 - \epsilon$  (in this case,  $1 - 0.7$ , or 0.3). Column 7, the nonurban productivity factor, is calculated from column 4 in the same way as the urban productivity factor is calculated from column 3, except that the sign is changed, since nonurban productivity acts in the opposite direction from urban productivity as an influence on the rate of urbanization.

Column 8 presents the error-of-fit measure for each economy. As discussed earlier, that measure is the difference between urban and nonurban average productivity change (based on observed income) divided by the average employment change in the two sectors. A positive sign indicates a greater urban than nonurban average increase in labor productivity, using observed measures of income; a negative sign indicates the reverse. A value of less than one in absolute terms indicates that relatively more of the employment changes are associated with the factors explained by the model than with failure to satisfy exactly the proportionate productivity condition. The effect of the error is smaller the closer is the value to zero.

Column 5, proportionate total population growth, shows, as expected, that total population growth accounts for a rate of urbanization that is generally on the order of 2 percent a year (0.020, as a decimal fraction, as expressed in the table). It ranges up to 3 percent in some cases and for certain economies, such as Yugoslavia, is very low. The direct effects of population growth thus cannot account for the high rates of urbanization—up to 7 percent—which are found in the table. Population growth could also cause pressure on limited agricultural land supplies and force growing numbers of people to seek work in urban areas. This effect would be

manifested in a resource scarcity situation in which supplies of nonlabor resources were more expansible in urban than in rural areas. If nonlabor resources were as scarce in urban as in rural areas, in the sense of having an elasticity of supply that is zero or equally small, nonlabor inputs would tend to increase by similar relative amounts in both rural and urban areas and there would be little impetus to urbanization. In a situation in which limited supplies of land provided the major impetus to urbanization, one would expect to see limited increases in nonlabor sources of income growth in rural areas, coupled with larger increases in labor inputs in urban areas as the supply of urban sites with associated infrastructure and other urban facilities expanded to accommodate people displaced from agriculture. These endogenous increases in urban nonlabor inputs would be greater than the increases in agricultural nonlabor inputs, but if the only important impetus to urbanization were displacement of agricultural labor, the increases in urban nonlabor inputs still might not be great, since in this depiction of the world more vigorous positive impetuses to economic development are lacking.

One might suppose that economies with less than 1 percent a year growth in agricultural nonlabor inputs may have a land scarcity situation. In table 2-2 twenty-five of the sixty-six economies have growth of nonlabor sources of nonurban income of less than 1 percent a year (or, as a decimal fraction, 0.010) as indicated in column 4. Of these, ten experienced negative rates of change in nonlabor sources of income, which indicates retrogression and other special factors not explained by the usual land scarcity situation. Note that land scarcity does not imply negative growth of nonlabor sources of income but rather no growth or a very small increase. The economies with low and negative growth for nonlabor sources of nonurban income tend to be among those with the lowest incomes, as indicated by their early appearance in the table. (Economies are ranked by 1978 per capita income.) It is also true that these economies generally have high urbanization rates. In the five lowest-income economies (Bangladesh, Ethiopia, Mali, Somalia, and Chad) rates of change for nonlabor sources of rural income are negative or less than 1 percent a year. Each had a rate of growth of urban population of over 5 percent a year (0.050, as a decimal). Of the twenty-five economies with low or negative growth of nonlabor sources of rural income, only three—Senegal, Congo and Jamaica—had an urban population growth of less than 4 percent a year. By contrast, in only twenty of the forty-one economies where growth of nonlabor sources of rural income was greater than 1 percent were annual rates of urban population growth more than 4 percent.

The apparent tendency for high rates of urbanization

to be associated with low rates of growth of nonlabor rural income sources does not prove that land scarcity has been a highly important reason for rapid urbanization. A more likely explanation of the finding may be the general sluggishness of agriculture in the twenty-five economies and a general failure to advance knowledge or introduce new techniques—in short, a failure of non-land conventional and nonconventional inputs to increase, even though land supply may have been elastic. Lack of inputs other than land may have acted to depress the marginal productivity of labor and hence the demand for labor in rural areas. If so, it is a lack of effective economic development in agriculture and not a lack of land that helped to foster high urbanization rates among the countries that experienced low rates of growth of nonlabor rural income sources.

The interpretation of events in terms of a more general economic development perspective that goes beyond preoccupation with the role of agricultural land is strengthened by considering additional results in table 2-2. One of the striking relations in the table is the positive association between growth rates of urban and rural nonlabor sources of income, as revealed by looking down columns 3 and 4. The two series have a simple correlation coefficient ( $R^2$ ) of 0.964. Two reasons may be noted for the positive association. First, the nonlabor income sources include capital, which is mobile between sectors. Increases in the capital stock of an economy will tend to be allocated among sectors so as to equalize the marginal returns to capital and impart changes in the same direction to nonlabor income sources in the different sectors. Second, and perhaps more important, the nonlabor income sources reflect more general economic development changes, in addition to conventional inputs, that act to increase the productivity of the conventional inputs. There is some tendency for the effects of economic development changes to be pervasive; for example, the influence of research, infrastructure, incentive structures, and institutions tends to be felt throughout an economy. Just as there is a tendency to equalize returns to conventional inputs among sectors, there is some tendency to equalize returns to deeper economic development efforts, even though the latter tendency may be less perfect. The fact that the economies in table 2-2 which have low rates of growth of nonlabor rural income sources also tend to have low rates of growth of nonlabor urban income sources suggests that general economic development differences, and not primarily agricultural land, are at work in explaining those economies' high rates of urbanization.

An ancillary finding in table 2-2 is that the growth in nonlabor sources of income tends to rise with the per capita income of the economy (which is to say, it rises as one looks down the table). The simple  $R^2$  between the income rank of an economy (as indicated by its order in

the table) and the urban nonlabor income source growth rate in column 3 is 0.658. The simple  $R^2$  between income rank and the rural nonlabor income source growth rate in column 4 is 0.642. The finding reflects the tendency among these economies for the more affluent among them to gain in relation to the less affluent.

To return to factors that affect the rate of urbanization, it is noteworthy that when columns 3 and 4 of table 2-2 are compared, for every one of the sixty-six economies the growth rate of urban nonlabor income sources is greater than that for rural nonlabor income sources. This finding is consistent with the hypothesis noted earlier, that the techniques of urban production are easier to transfer among economies than those of agricultural production. Since an excess of urban over rural nonlabor income source growth acts to foster urbanization in this model, the rate of urban population growth is greater than the rate of overall population growth for every one of the sixty-six economies.

A final consideration in interpreting table 2-2 is the role of the fraction of the population which is urban. If this fraction is small, any differential change between urban and rural nonlabor income source growth has a magnified effect on urban population. The reason is that a given absolute change has a greater percentage effect on a small number than on a large number. This effect is reflected in equation 2-9, in which the coefficient  $F_a / (1 - \epsilon)$  multiplies the growth rate of nonlabor income source to arrive at the urban and rural productivity factors. Thus, economies with large nonurban fractions in their populations tend to have greater urbanization rates. Since these economies generally have low per capita incomes, they are the ones that have predominantly experienced the highest urbanization rates.

Progressing down the table to increasingly higher-income economies, one finds that the difference between urban and rural nonlabor income source growth stays about the same or becomes larger, but the contributions of the urban and rural productivity factors decrease because of the declining value of the multiplier  $F_a / (1 - \epsilon)$ . Since  $F_a = 1 - F_w$ , the multiplier will be greater than one if the fraction of the population that is urban is greater than the share of income that accrues to labor. The condition would be fulfilled for all but a handful of the economies, regardless of the value of  $\epsilon$ , but the multiplier falls more and more toward one toward the bottom of the table. The rates of growth of urban population decline until they are not much greater than the rate of overall population growth. This is shown in equation 2-9, which gives an analytic explanation for the descriptive finding of the association between the fraction that is urban and the growth rate of urban population.

The analysis presented here of factors that affect urbanization was carried out for three other cases: a

population analysis in which labor share,  $\epsilon$ , is 0.5 instead of 0.7; and two labor force analyses that use nonagricultural labor force as a surrogate for urban labor force, with  $\epsilon$  at 0.7 in one case and 0.5 in the other. The results were essentially similar to those presented here.

The existence of many variations from the general patterns that have been discussed indicates that conditions in individual economies play a part in urbanization. India and Burma are among the economies which stand out particularly as exceptions to some of the patterns in table 2-2. Both rank low in per capita income and have high nonurban proportions of population, yet unlike most other economies with these characteristics they have relatively low urbanization rates. Growth rates of urban and rural nonlabor income sources are largely offsetting in these two countries. The size of the Indian economy might influence the result, but this consideration would not apply to Burma.

India and Burma are only two examples that indicate the usefulness of more analytic work for understanding patterns of urbanization. For any one of the economies, more detailed analysis could be undertaken, with more refined identification of the factors that affect urbanization as an object.

Further comparative analytic work to contrast and understand differences among regions and different types of economies is recommended to strengthen the knowledge base for prediction and policy concerned with economic development. Further work should give attention to differences in industrial structure and to considerations that affect trade. Greater detail on types of goods could be included, as could allowance for price effects, age composition, education, dependency, and other considerations bearing on work participation. The assumption that the number of workers in urban areas goes hand in hand with the population of these areas would be modified.

## Conclusion

This chapter suggests that the degree to which a country succeeds in fostering growth of urban productivity is likely to be the major determinant of its urbanization. Population growth will continue to be a source of both urban and nonurban growth. Countries that lag in growth of agricultural productivity will experience added urbanization pressures.

As important as these general findings is the usefulness of the framework of this chapter as a basis for quantifying the experience of individual countries. Prediction and planning are enhanced, and the economic factors that influence the effectiveness of policies concerned with urbanization can be better understood.

## Bibliography

- Armstrong, Harvey, and Jim Taylor. 1978. *Regional Economic Policy and Its Analysis*. Deddington, Oxford, U.K.: Philip Allan.
- Arriaga, Edward. 1970. "A New Approach to the Measurement of Urbanization." *Economic Development and Cultural Change*, vol. 18, no. 2 (January), pp. 206–18.
- Berry, Brian J. L. 1961. "City Size Distributions and Economic Development." *Economic Development and Cultural Change*, vol. 9, no. 4, pt. 1 (July), pp. 573–88.
- Bourne, L. S., and J. W. Simmons, eds. 1978. *Systems of Cities: Readings on Structure, Growth, and Policy*. New York: Oxford University Press.
- Breese, Gerald. 1966. *Urbanization in Newly Developing Countries*. Englewood Cliffs, N.J.: Prentice-Hall.
- Chapin, Stuart F., Jr. 1971. "Selected Theories of Urban Growth and Structure." In L. S. Bourne, ed., *Internal Structure of the City: Readings on Space and Environment*. New York: Oxford University Press.
- Colosio, Donaldo. 1979. *Urbanization and Economic Development in Mexico*. Working Paper 79-19. Laxenburg, Austria: International Institute for Applied Systems Analysis.
- Currie, L. 1975. "The Interrelation of Urban and National Economic Planning." *Urban Studies*, vol. 12, no. 1 (February), pp. 37–46.
- Davis, Kingsley. 1965. "The Urbanization of the Human Population." *Scientific American*, vol. 213 (September), pp. 28, 40.
- . 1971. "The Role of Urbanization in the Development Process." In Raanan Weitz, ed., *Urbanization and the Developing Countries. Report on the Sixth Rehovot Conference*. Praeger Special Studies in International Economics and Development. New York: Praeger.
- Felstehausen, H. 1971. "Planning Problems in Improving Colombian Roads and Highways." *Land Economics*, vol. 47, no. 1 (February), pp. 1–13.
- Friedmann, J., and C. Weaver. 1979. *Territory and Function*. University of California Press.
- Hansen, Niles M. 1969. "Regional Development and the Rural Poor." *Journal of Human Resources*, vol. 4, no. 2 (Spring), pp. 205–14.
- Hansen, Niles M., ed. 1976. *Public Policy and Regional Economic Development: The Experience of Nine Western Countries*. Cambridge, Mass.: Ballinger.
- Harberger, C. 1970. "Economic Policy Problems in Latin America: A Review." *Journal of Political Economy*, vol. 78, no. 4, pt. 2 (January–August), pp. 1007–16.
- Henderson, J. V. 1980. "A Framework for International Comparisons of Systems of Cities." Urban and Regional Report 80-3, World Bank, Washington, D.C. Processed.
- Holleb, D. B. 1976. "Mao's Legacy for Revolutionary Urban Development." Working Paper. University of Chicago, Center for Urban Studies. December.
- Hoselitz, Bert. 1954. "Introduction; The Role of Cities in Economic Development and Cultural Change." *Economic Development and Cultural Change*, vol. 3, no. 1 (October), pp. 4–5.
- . 1957. "Urbanization and Economic Growth in Asia." *Economic Development and Cultural Change*, vol. 6, no. 1 (October), pp. 42–54.
- Kelley, Allen C., and Jeffrey G. Williamson. 1979. *Modeling Urbanization and Economic Growth*. Working Paper 79-81, Laxenburg, Austria: International Institute for Applied Systems Analysis. September.
- Lasvwn, J. R. 1973. "Urbanization and Development—The Temporal Interaction Between Geographical and Sectoral Cluster." *Urban Studies*, vol. 10, no. 2 (June), pp. 163–88.
- Linn, J. F. 1979. "Automotive Taxation in the Cities of Developing Countries." *Nagarlok* (January).
- Lipton, Michael. 1977. *Why Poor People Stay Poor: Urban Bias in World Development*. Cambridge, Mass.: Harvard University Press.
- Miller, J. C. 1979. "Regional Development: A Review of the State of the Art." AID 20523. Agency for International Development, Washington, D.C. Processed.
- Mills, Edwin. 1971. "City Sizes in Developing Economies." In Raanan Weitz, ed., *Urbanization and the Developing Countries. Report on the Sixth Rehovot Conference*. Praeger Special Studies in International Economics and Development. New York: Praeger.
- . 1972. "Welfare Aspects of National Policy toward City Sizes." *Urban Studies*, vol. 9, no. 1 (February), pp. 117–24.
- Pascal, Anthony H. 1971. "Urban Growth and Development: Discussion." *American Economic Review*, vol. 61, no. 2 (May), pp. 357–59.
- Renaud, Bertrand. 1981. *National Urbanization Policies in Developing Countries*. New York: Oxford University Press.
- Reynolds, L. G. 1975. "China as a Less Developed Economy." *American Economic Review*, vol. 65, no. 3 (June), pp. 418–28.
- Richardson, Harry W. 1977. *City Size and National Spatial Strategies in Developing Countries*. World Bank Staff Working Paper 252. Washington, D.C.
- Rondinely, Dennis A., and Kenneth Ruddle. 1978. *Urbanization and Rural Development: A Spatial Policy for Equitable Growth*. New York: Praeger.
- Tolley, George S. 1974. "The Welfare Economics of City Bigness." *Journal of Urban Economics*, vol. 1, no. 3, pp. 324–45.
- Tolley, G. S., and G. D. Gwyer. 1967. "International Trade in Agricultural Products in Relation to Economic Development." In Herman M. Southworth and Bruce F. Johnston, eds., *Agricultural Development and Economic Growth*. New York: Cornell University Press. Pp. 403–47.
- Tolley, G. S., and G. K. Kripalani. 1974. "Stages of Development and Urbanization." In A. A. Brown, J. A. Licori, and E. Neuberger, eds., *Urban and Social Economics in Market and Planned Economies*. New York: Praeger.
- Tolley, G. S., and S. Smidt. 1964. "Agriculture and the Secular Position of the U.S. Economy." *Econometrica*, vol. 32, no. 4 (October), pp. 554–75.

- Tolley, G. S., and Barton Smith. 1975. "Social Costs and the Rural-Urban Balance." In Earl O. Heady and Larry R. Whiting, eds., *Externalities in the Transformation of Agriculture: Distribution of Benefits and Costs from Development*. Ames, Iowa: Iowa State University Press.
- Tolley, George S., Philip E. Graves, and John L. Gardner, eds. 1979. *Urban Growth Policy in a Market Economy*. New York: Academic Press.
- Tolley, G. S., R. J. Krumm, and P. Graves. "Design of Incentives for Spatial Intervention for Developing Countries." Working Paper. Chicago: University of Chicago.
- Townroe, Peter M. 1979. *Employment Decentralization: Policy Instruments for Large Cities in Less Developed Countries*. Vol. 10, pt. 2, in D. R. Diamond and J. McLoughlin, eds., *Progress in Planning*. Elmsford, N.Y.: Pergamon.
- Wheaton, W. C., and H. Shishido. 1979. "Urban Concentration, Agglomeration Economies, and the Level of Economic Development." Working Paper. Cambridge, Mass.: MIT.

## *What Drives City Growth in the Developing World?*

Allen C. Kelley and Jeffrey G. Williamson

Cities are capital-intensive, a stylized fact which has led many pessimists to assert that current rates of city growth in the developing world cannot be sustained. To make matters worse, cities are stocked with public capital at low or zero user charges, and in that sense they are too capital-intensive. Since the social costs of immigration thus exceed private costs, the number and sizes of cities turn out to be excessive. Furthermore, the increasing scarcity of urban land encourages the substitution of capital for land and increases capital intensity. When the economy can no longer finance such urban costs—when it fails, for instance, to cope with social overhead and housing requirements—the process of urbanization is retarded.

No doubt the qualitative analysis is correct: cities are capital-intensive and have a voracious appetite for savings and accumulation. But is the quantitative emphasis—the idea that cities are “too big”—warranted?

It is also often alleged that rapid city growth in the developing countries represents a disequilibrium, that these countries have “overurbanized,” and that a painful structural adjustment will eventually be required. Analysts of urban problems in the developing world point out that migrants are attracted to the cities in the hope that they (or their children) will be selected for training and employment in protected, high-wage sectors. Apparently, new in-migrants are willing to accept underemployment in low-wage traditional service sectors while waiting. Eventually, however, social discontent is likely to erupt. Are labor markets in these cities really in serious disequilibrium, and is the developing world overurbanized as a consequence?

It is often alleged that the unusually high rates of

developing-country urbanization in the recent past can be traced to the availability of cheap energy, technological diffusion that favors modern urban sectors, heavy capital inflows, world trade liberalization, a drift toward domestic price distortions that favor city output, and unusually rapid population growth. That these conditions have begun to show signs of changing suggests that the developing world may have overurbanized in the recent past. Urban growth may be expected to decelerate over the next two decades under certain conditions: a recurrence of fuel scarcity, technological regression in modern sectors in the face of a productivity slowdown in industrial countries, diminished capital transfers because of economic austerity in industrial countries, a retreat toward protectionism in industrial countries, and a decrease in population growth.

In short, the pessimists offer three sources for an incipient urban crisis in the developing world: a savings constraint which will bridle the growth of capital-intensive cities; a labor market disequilibrium which has made overurbanization a temporary but serious problem of overshoot; and the disappearance of unusual external conditions which were favorable to urban growth in the past. Although the pessimists have established a plausible case, no one to our knowledge has offered a quantitative assessment of the importance of these forces over the past two decades. Without such an assessment, debate over future trends in developing countries will be dominated by allegation and anecdotal evidence. It is our view that the debate can be better informed by the application of a general equilibrium model of development that includes some of the costs of urbanization, so that “natural limits” to urban growth can be evalu-

ated and the impact of changing economic and demographic conditions assessed (see Kelley and Williamson 1984a, b).

### The Limits on City Growth

The limits on urban growth are set, on the one hand, by urban costs that affect migration decisions and, on the other, by rising urban investment requirements that compete with productive capital accumulation. Urban land constraints raise rents, increase living costs in urban areas relative to rural areas, and thus inhibit migration into cities. To the extent that rising rents and urban disamenities are caused by high densities, crowding, and other manifestations of inelastic urban land supply, city rents reflect the quality of urban life as well as living costs. To evaluate the impact of urban land constraints on city rents, a general equilibrium model is needed which, at the minimum, admits housing service activities and confronts equilibrium land use issues. Furthermore, a variety of urban land requirements must be included—residential squatter settlements, factory sites, land use for public social overhead, and luxury housing sites.

The housing and social overhead investment requirements of city growth must also be analyzed. Such “unproductive” urban investment requirements (first analyzed by Coale and Hoover 1958) compete directly with “productive” capital accumulation and may check urban growth. If such overhead investment is forgone, housing costs rise, the quality of urban services falls, and migration to cities is discouraged. Thus, in addition to the effect of the rise in the relative cost of living in the city, the rise in unproductive investment requirements in cities may lower the rate of productive capital accumulation and job creation and set a limit to urban growth.

Modern urban sectors also tend to be relatively intensive in skills, in intermediate inputs, and in such imported inputs as energy. To the extent that cities are energy-intensive, fuel scarcity can limit urban growth. If capital and skills are complements and different labor skills are poor substitutes for each other, rapid rates of urban capital accumulation imply increasing demands for skilled labor, which can constrain capacity expansion, retard growth of employment, and limit urban growth. An effort to relax the labor constraint through skill accumulation is likely to compete with productive urban capital accumulation and itself constitute a limit on urban growth.

This chapter uses a computable general equilibrium model to analyze past, present, and future growth of cities in the developing world. Since the model has appeared elsewhere (Kelley and Williamson 1980, 1982,

1984b), only a suggestive outline is presented here. The model is in the neoclassical general equilibrium tradition. Prices of outputs and inputs are completely flexible, and most are endogenously determined; firms are driven by profit maximization; households are driven by utility maximization; and even government demand decisions obey well-defined rules from consumer demand theory. Mobility of capital and labor is constrained to reflect the institutional realities of factor markets in developing countries, but economic agents are motivated to search for the optimal sectoral and spatial use of resources.

The model has eight sectors (table 3-1). Tradables and nontradables—the latter include various location-specific services—are distinguished. This is not the first multisectoral model to recognize nontradables, but it is the first spatial development model to stress the importance of nontradables as an influence on spatial cost-of-living differentials, on migration behavior, and thus on the rate of urban growth.

The model is savings-driven, and the aggregate savings pool is generated endogenously from three sources: retained after-tax corporate and enterprise profits, government savings, and household savings. This savings pool is allocated competitively and endogenously to three uses: investment in physical capital (productive investment), investment in human capital (training), and investment in housing (unproductive investment). It should be emphasized that these three modes of accumulation are competitive and are determined endogenously; that is, investment in skills (training) takes place up to the point at which rates of return are equated to the economywide rate on physical capital accumulation. Physical capital goods are allocated across the three capital-using sectors so as to minimize differentials in after-tax rates of return. Dwelling investment utilizes household savings only up to the point at which rates of return equal the economywide rate on physical capital accumulation. Of course, there are institutional and technological features which seriously restrict the ability of the economy to equate rates of return at the margin. Any of the three dwelling markets (rural, urban squatter settlements, and formal urban housing) may be starved for funds, since the absence of an intersectoral mortgage market may leave housing investment requirements in excess demand. The immobility of sector-specific capital stocks makes it likely that current investment allocations are insufficient to equalize sectoral rates of return. Indeed, the larger are housing requirements, the smaller is the residual pool available for physical capital accumulation and the more likely it is that current investment allocations are insufficient to equalize sectoral rates of return. Furthermore, firms' demands for skills may remain unsatisfied if the stock of

**Table 3-1. Sector Characteristics of the Kelley-Williamson Model**

Sector	Location	U.N. ISIC counterpart	Determination of market price	Tradability characteristics	Production inputs		Production function form
					Primary	Intermediate	
Manufacturing and mining (M)	Urban	Manufacturing, mining	Exogenous	Traded internationally and interregionally	Capital, skills, labor	Imported raw materials and fuels, A,KS	Nested CES, constant returns
Modern capital-intensive services (KS)	Urban	Electricity, water, and gas, banking, public administration, trade, commerce, construction	Endogenous	Traded interregionally	Capital, skills, labor	Imported raw materials and fuels, A,M	Nested CES, constant returns
Informal urban services (US)	Urban	Personal services, some trade and commerce	Endogenous	Not traded	Labor	None	Cobb-Douglas diminishing returns
Low-quality (squatter) urban housing services (H, US)	Urban	Dwellings (rent and imputed ownership)	Endogenous, owner-occupier shadow price	Not traded	Dwellings, land	None	Cobb-Douglas constant returns
High-quality urban housing services (H, KS)	Urban	Dwellings (rent and imputed ownership)	Endogenous, owner-occupier shadow price	Not traded	Dwellings, land	None	Cobb-Douglas constant returns
Agriculture (A)	Rural	Agriculture, livestock, forestry, fishing, hunting	Exogenous	Traded internationally and interregionally	Capital, land, labor	Imported raw materials and fuels, M, KS	Cobb-Douglas constant returns
Informal rural services (RS)	Rural	Personal services, some trade and commerce	Endogenous	Not traded	Labor	None	Cobb-Douglas diminishing returns
Rural housing services (H, RS)	Rural	Dwellings (rent and imputed ownership)	Endogenous, owner-occupier shadow price	Not traded	Dwellings	None	Leontieff

Sources: Kelley and Williamson (1980, 1982, 1984b).

potential trainables is insufficient to meet desired levels of investment in training. In short, capital market disequilibrium may well be a chronic attribute of the economy.

Finally, some exogenous variables that help drive the economy over time are alleged to have influenced city growth. These variables are the nominal value of foreign capital and aid ( $F$ ) available each year to help finance the development effort and forestall balance of payments problems; the total unskilled labor force ( $L$ ), which is determined by earlier demographic events; the sectoral rates of change in total factor productivity, which favor modern sectors and are labor saving; prices of imported raw materials and fuels ( $P_Z$ ), which are influenced by actions of OPEC and by other world market conditions; and the terms of trade between primary exportables ( $P_A$ ) and manufactured importables ( $P_M$ ), which are twisted by domestic price policy and the political economy of protectionist or liberalization winds in the industrial nations.

To permit generalization about urbanization in the developing world, the representative developing economy (RDE) has been employed as the data base on which the model is estimated and validated. The RDE is a fictional economy that embodies the experience of developing economies on three continents since 1960. Three criteria were used to select the RDE group: availability of data beginning in the early 1960s, low per capita income but some successful growth, and close conformity to the model's most important assumptions. (The last-named criterion implied the rejection of economies which have been heavily dependent on foreign capital and on mineral exports such as fuel.) The forty economies which met the criteria<sup>1</sup> seem to exhibit the same average attributes as the far larger group analyzed by Chenery and Syrquin (1975). The RDE group contains eleven of the world's twelve largest developing economies and covers more than 80 percent of the population of the developing world excluding China.

### Predicting the Past

Of the about one hundred endogenous variables generated by the model, three—urbanization, city growth, and rural-urban migration—form the core of this chapter. The illustrations below show how the model replicates past trends for these and other selected variables, both to throw light on past trends and to validate the use of the model for making predictions. Much of the initial discussion focuses on 1960–73, but the late 1970s are also discussed.

### Macroeconomic Indicators

*Aggregate Growth.* As set out in table 3-2, the model generates an annual average growth rate of gross domestic product (GDP) in constant prices of 6.26 percent for 1960–73, and the growth rate of GDP per worker averages 3.58 percent. A significant trend acceleration is indicated; GDP growth rises from an annual rate of 5.92 percent in 1960–65 to 6.60 percent in 1968–73. That these predictions conform fairly closely to the actual pre-OPEC history of the RDE group helps to establish the model's plausibility at the most aggregate level.

*Unbalanced Growth and Industrialization.* The model also replicates the RDE history of industrialization quite closely, as can be seen in table 3-3. The M-goods sector underwent relatively rapid growth, from 15.9 percent of GDP at factor cost (current prices) in 1960 to 20.8 in 1973; the model generates almost identical trends. The service sector underwent a more modest rise, from 47.5 to 50.6 percent of GDP at factor cost, and the model captures these trends too. The rapid decline in the relative importance of agriculture—from 36.6 to 28.6 percent—is also faithfully replicated in the simulation. It is satisfying that the model so closely replicates the quantitative record of the unbalanced growth of sectoral output.

*Investment Allocation and Sources of Finance.* Gross domestic investment as a share in GDP increased markedly between 1960 and 1973, from 15.59 to 19.46 percent (Kelley and Williamson 1984b, table 3.10, p. 89). Total investment shares, including investment in training, also drift upward in the model, a prediction which conventional national accounts cannot document. Although the aggregate investment share increased over the period, the model suggests that the distribution of investment was remarkably stable. Investment allocation by sector and type is poorly documented for the pre-OPEC period, but the model predicts that the distribution of investment by sector and type must have been quite stable over the thirteen years. The sources of

Table 3-2. Annual Average Growth Rate, 1960–73 (percent)

Period	GDP (constant prices)		GDP per worker (constant prices)	
	Model	Actual	Model	Actual
1960–65	5.92	5.78	3.19	3.24
1968–73	6.60	6.12	3.98	3.58
1960–73	6.26	5.80	3.58	3.26

Source: Kelley and Williamson (1984b), table 3.5, p. 84.

**Table 3-3. Output Mix, 1960–73**  
(percent of GDP)

Sector	1960		1970		1973	
	Model	Actual	Model	Actual	Model	Actual
Agriculture (A)	36.6	36.6	30.4	30.9	28.2	28.6
Manufacturing and mining (M)	15.9	15.9	19.3	19.2	20.9	20.8
Modern capital-intensive services (KS)	29.6	29.6	30.4	n.a.	30.8	n.a.
Informal urban services (US)	6.2	6.2	6.9	n.a.	7.2	n.a.
Informal rural services (RS)	3.2	3.2	3.1	n.a.	2.8	n.a.
Housing, all sectors (H)	8.5	8.5	9.9	n.a.	10.1	n.a.
All services combined (KS + US + RS + H)	47.5	47.5	50.3	49.9	50.9	50.6

n.a. Not available.

Source: Kelley and Williamson (1984b), table 3.7, p. 86.

saving, however, did not show the same stability (Kelley and Williamson 1984b, table 3.10, p. 89).<sup>2</sup> The share of gross domestic investment financed by public saving rose to 34 percent. The model predicts a comparable rise to 33.7 percent, with both shares rising by about 5 percentage points. The fall in the private share of finance was attributable not so much to lagging household saving as to a decline in the share of gross corporate saving in gross domestic investment, from 31.7 to 26.6 percent.

*Income Inequality and Wage Patterns.* Size distributions for the model and for an average of eighteen of the RDE for which data exist are given in Kelley and Williamson (1984b, table 3.12, p. 91). In spite of well-known limitations, the top 5 percent and the bottom 20 percent shares of national income exhibit some striking trends. Contrary to conventional wisdom, an unambiguous rise in inequality is not confirmed. The bottom 20 percent suffered an erosion in income shares during the period, but the top 5 percent underwent an even more dramatic erosion. Clearly, the middle classes flourished at the expense of both the very rich and the very poor. Those are exactly the conditions under which Lorenz curves intersect. Inequality inferences are then impossible unless value weights are imposed on social classes. The model exhibits the same behavior: a 3.1 percent fall in the share in national income of the top 5 percent is predicted, compared with the RDE history of a fall of 2.6 percent. Similarly, while the bottom 20 percent found its share declining by 0.3 percent, the model predicts a fall of 0.2 percent. The group in the 60–90 percent range increased its share by 3.7 percent, which implies that the source of the inequality lies with increased inequality of earnings. Those in the 60–90 percent range in the model are urban skilled and unskilled in the formal sectors. In short, the model generates wage inequality and skill scarcity.

### Urban Indicators

*Migration, Urbanization, and City Growth.* Table 3-4 documents four key aspects of urban development: the share of the population that is urban, city growth rates, net rural out-migration rates, and net urban in-migration rates. The rise in urban shares provides some evidence of accelerating rates of urbanization and conforms to the pre-inflection-point phase along logistic urbanization curves that is commonly found in developing-country time series (Preston 1979; Ledent 1980). In the first five years of the simulation (1960–65) urbanization levels rise by 3.18 percentage points, while in the last five years (1968–73) they rise by 5.38 percentage points. Similarly, city growth rates rise over time, con-

**Table 3-4. Migration, Urbanization, and City Growth, 1960–73, Model Predictions**

Year	Percent urban	City growth rate	Net rural out-migration rate	Net urban in-migration rate
1960	32.60	—	—	—
1961	33.55	5.56	1.41	2.91
1962	34.30	4.87	1.13	2.24
1963	34.73	3.85	0.65	1.25
1964	35.25	4.13	0.81	1.51
1965	35.78	4.11	0.82	1.50
1966	36.49	4.60	1.10	1.98
1967	37.23	4.61	1.16	2.01
1968	38.07	4.90	1.35	2.27
1969	38.95	4.91	1.42	2.31
1970	39.93	5.15	1.60	2.51
1971	40.96	5.22	1.71	2.58
1972	42.15	5.56	2.02	2.91
1973	43.45	5.72	2.21	3.07

—Not applicable.

Source: Kelley and Williamson (1984b), table 3.13, p. 93.

forming to the trends reported in the *World Tables 1976* (World Bank 1976, table 2, "Social Indicators") between 1960 and 1970. Rural-urban migration rates also rise.

Table 3-5 presents model predictions and actual experience. City growth rates in the RDE group were 4.60 percent a year during the 1960s; the model predicts a rate of 4.67 percent. The model predicts a rural out-migration rate of about 1.1 percent a year, and the historical rate ranged between 1 and 1.2 percent; the predicted urban in-migration rate is 2.1 percent a year, and the historical rate ranged between 1.8 and 2.3 percent. Finally, 45 percent of the increase in city population is accounted for by in-migration in the model; this figure falls between Preston's (1979) estimate of 39 percent (based on twenty-nine developing countries) and Keyfitz's regional estimate of 49 percent (1980, p. 151). It is also close to the 42 percent figure for developing countries reported by Linn for 1970-75 (1979, p. 73).

*Urban Land Use, Density, and Land Scarcity.* Table 3-6 reports indicators of urban land use and scarcity. The share of urban land devoted to squatter settlements rises sharply (Mohan 1979; Beier and others 1975). Competition for land use generates sharply rising land scarcity; the shadow site rent on urban land almost doubles between 1960 and 1970 and in 1973 reaches a level about 2.3 times that of 1960. Urban land values (deflated by the general price level) surge; the index rises from a base of 100 in 1960 to 195.2 in 1970 and 239.7 in 1973. These trends in urban land values imply a rapid annual growth of 7 percent—exactly the sharp rises which were singled out at the U.N. Habitat Conference in 1976 (United Nations 1976). Urban densities rise everywhere in the model, but they rise most dramatically in "luxury" housing, where the relative scarcity of land compared with structures rises most sharply, encouraging land saving and greater density.

**Table 3-5. Predicted and Actual Migration, Urbanization, and City Growth, Averages, 1960-70 (percent)**

Variable	Predicted	Actual
Annual city growth (compounded)	4.67	4.60
Total increase in urban share of population	7.33	5.30
Annual increase in urban share of population	0.73	0.53
Net in-migrant share of increase in urban population	45.0	39.3-49.0
Net in-migration rate	2.09	1.81-2.26
Net out-migration rate	1.10	0.97-1.21

Source: Kelley and Williamson (1984b), table 3.13, p. 93.

**Table 3-6. Urban Land Use, Density, and Land Scarcity, Model Predictions, 1960, 1970, and 1973**

Variable	1960	1970	1973
Share of urban land in squatter settlements (percent)	43.0	52.2	53.9
Urban land density (persons ÷ area) (1960 = 100)			
High-quality housing areas	100.0	152.0	172.7
Squatter settlements	100.0	121.4	141.7
Shadow site rent on urban land (1960 = 100)	100.0	183.3	233.3
Shadow price on urban land (1960 = 100)	100.0	195.2	239.7

Source: Kelley and Williamson (1984b), table 3.14, p. 95.

*Housing Scarcity and Cost-of-Living Differentials.* In addition to land scarcity, excess demand for housing units in the short run and rising costs of housing construction in the long run can inflate the cost of city life, as summarized in table 3-7. Excess housing demand arises because capital market segmentation excludes intersectoral (mortgage) loans for housing investment, and each socioeconomic class must rely on its own internally generated saving to meet housing investment. Table 3-7 reports that the saving constraint is never binding for rural households or for skilled workers and higher-income households, but for squatter settlements—the faster growing sector—the saving constraint is binding very early in the period and thus excess short-run demand for new low-quality dwelling units is generated. Excess demand and increasing urban land scarcity ensure a rapid upward drift in (shadow) rents in

**Table 3-7. Housing Scarcity and Cost-of-Living Differentials, Model Predictions, 1960, 1970, and 1973**

Variable	1960	1970	1973
Excess housing investment demand <sup>a</sup>			
Urban high-quality housing	-3.77	-2.55	-2.67
Urban squatter settlements	-0.02	0.37	0.28
Rural housing	-0.29	-0.58	-0.88
Housing rent index			
Urban high-quality housing	1.00	1.35	1.46
Urban squatter settlements	1.00	1.72	1.90
Rural housing	1.00	1.38	1.45
Urban squatter ÷ rural	1.00	1.25	1.31
Cost-of-living index			
Urban ÷ rural	1.00	1.09	1.12

a. [Desired housing investment demand - household saving] ÷ desired housing investment demand.

Source: Kelley and Williamson (1984b), table 3.15, p. 96.

urban squatter settlements, and the cost-of-living differential rises as a result.

### Understanding City Growth: Some Major Influences

Which exogenous variables have had the greatest impact on city growth in developing countries? Which are least likely to account for future growth? Answers to these questions can be divided into three parts: the size of the past and future changes in the exogenous variable thought to influence endogenous rates of city growth; the short-run comparative static impact of that exogenous variable; and the long-run forces set in motion by the short-run comparative static impact. This section focuses on short-run comparative static elasticities and explores the impact of some key macroeconomic events on urban growth: unbalanced productivity advance, world market conditions and price policy, accumulation, demographic change, and land scarcity.

Each of these short-run elasticities reflects the full general equilibrium impact of the exogenous variable in question, based on the initial conditions in 1970. Labor markets adjust through migration, and urban land markets seek an optimal land use solution, but capital markets are severely constrained in the short-run analysis: old capital cannot migrate, and new capital goods and newly trained skilled workers are not added to capacity. Investment responses are also ignored in the short-run analysis: recent historical experience with sectoral investment allocation is assumed to guide entrepreneurs who are slow to adjust to the new, unexpected, and shock-distorted rates of return.

#### *Unbalanced Productivity Advance*

If output demand is relatively price inelastic, sectoral total factor productivity growth (TFPG) for a sector tends to generate a relative price decline rather than an elastic supply response. Thus, the rise in the marginal physical productivity of factors used in a technologically dynamic sector will be partially offset by the decline in price, so that marginal value products rise by less, and shifts of resources, including labor, to the technologically dynamic sector are minimized. If, on average, urban sectors tend to have relatively high rates of TFPG, and if the demand for urban output is price elastic, final demand shifts toward the dynamic urban sectors, the derived demand for urban employment is augmented, and the city grows. The higher are price elasticities of demand for urban output, the greater is the city growth attributable to unbalanced productivity advance that favors the modern sectors.

In table 3-8 disembodied TFPG (the  $A_{jS}$ , in rates of change) in agriculture and manufacturing exert a much greater impact on urbanization than do productivity improvements in the service sectors. The price variables  $P_A$  and  $P_M$  are exogenously determined and are fixed by invoking the small-country case of infinite price elasticity. Services are nontradables with price elasticities of demand sufficiently low that the productivity-induced declines in service sector prices ( $P_{KS}$ ,  $P_{US}$ , and  $P_{RS}$ , in response to  $A_{KS}$ ,  $A_{US}$ , and  $A_{RS}$ , respectively) imply stable marginal value products and trivial employment and city growth effects. Although productivity advance in manufacturing is an important potential determinant of urbanization, rapid productivity advance in agriculture tends to forestall out-migration to the city. This result is in contrast to the closed dual economy model in which productivity advance in agriculture meets with demand absorption problems, declining farm terms of trade, and thus a labor surplus which out-migrates to glut urban labor markets.

Unbalanced rates of TFPG that favor manufacturing are likely to have been a key determinant of rapid immigration and city growth since the late 1950s. Not only are the comparative static elasticities in table 3-8 consistent with that position, but limited evidence suggests that annual rates of TFPG in manufacturing have been relatively high in most successful developing economies.

Although technological advance tends to be lower in the service sectors, especially in the informal service sectors, table 3-8 suggests that rapid TFPG in those sectors has had little impact on urbanization experience for the demand elasticity reasons already offered. If rapid TFPG matters little to urbanization even in the modern capital-intensive and skill-intensive KS sector, lagging productivity advance in the service sectors would also matter little. It is believed that urban social overhead is crucial to the profitability and viability of urban firms. In our model, KS activities such as transport, communications, and electricity have that role. Table 3-8 confirms that productivity advance augments KS services supplied in short-run equilibrium ( $Q_{KS}$ ), but because almost all of these productivity gains are passed on to users elsewhere in the economy ( $P_{KS}$  declines), employment in KS itself changes little. Final-demand customers of KS services, who tend to be the urban rich, gain. A major user of KS services as intermediate inputs is manufacturing, and employment rises there. The net effect on urbanization is, however, indirect and small.

#### *Short-Run Constraints on City Growth*

Productivity advances that favor modern sectors foster urbanization, but in the short run the city growth

Table 3-8. Comparative Static Impacts of Unbalanced Productivity Advance on City Growth in Developing Countries (1970 Elasticities)

Endogenous variable	Tradable commodities		Nontradable services		
	$A_M$	$A_A$	$A_{KS}$	$A_{US}$	$A_{RS}$
<i>City growth attributes</i>					
Percent urban	0.50	-0.26	0.03	-0.01	-0.03
City growth rate	10.29	-5.33	0.68	-0.26	-0.67
In-migration rate	20.57	-10.65	1.36	-0.52	-1.33
Squatter house rents	3.57	-0.73	0.40	0	-0.14
Cost-of-living differential (urban ÷ rural)	1.49	-0.80	0.12	-0.11	0.01
<i>Selected economywide attributes</i>					
$P_{KS}$	0.82	0.22	-0.91	0.03	0
$P_{US}$	1.03	0.24	0.15	-1.09	0
$P_{RS}$	0.02	1.17	0.09	-0.01	-1.03
$Q_M$	2.34	-0.57	0.24	0.08	-0.09
$Q_{KS}$	0.02	0.17	0.99	0.01	0
$Q_{US}$	0.23	0	0.02	0.75	-0.01
$Q_A$	-0.34	1.31	0.07	0	0.07
$Q_{RS}$	-0.11	0.32	-0.01	0	0.60
$L_M$	1.36	-0.84	0.10	0.18	-0.09
$L_{KS}$	-0.18	0.15	-0.07	0.11	0
$L_{US}$	0.24	0	0.02	-0.26	-0.01
$L_A$	-0.37	0.15	-0.02	0.01	0.08
$L_{RS}$	-0.11	0.34	-0.01	0	-0.41

Source: Kelley and Williamson (1984b), table 4.1, p. 103.

Note: Variables are defined as follows:  $A_M$ , total factor productivity growth in the manufacturing sector;  $A_A$ , total factor productivity growth in the agricultural sector;  $A_{KS}$ , total factor productivity growth in the modern services sector;  $A_{US}$ , total factor productivity growth in the informal urban services sector;  $A_{RS}$ , total factor productivity growth in the informal rural services sector;  $P_{KS}$ , price per unit of modern services;  $P_{US}$ , price per unit of informal urban services;  $P_{RS}$ , price per unit of informal rural services;  $Q_M$ , output in the manufacturing sector;  $Q_{KS}$ , output in the modern services sector;  $Q_{US}$ , output in the informal urban services sector;  $Q_A$ , output in the agricultural sector;  $Q_{RS}$ , output in the informal rural services sector;  $L_M$ , unskilled labor in the manufacturing sector;  $L_{KS}$ , unskilled labor in the modern services sector;  $L_{US}$ , unskilled labor in the informal urban services sector;  $L_A$ , unskilled labor in the agricultural sector;  $L_{RS}$ , unskilled labor in the informal rural services sector.

response is constrained, partly by problems with absorption of output, partly by short-run capacity constraints and skill bottlenecks, and partly by a rising supply price of unskilled labor in the cities. Table 3-9 focuses on the rising supply price of urban labor. Urban job creation fosters in-migration, but with some limitation. Although land use shifts to squatter settlements to accommodate rising density, urban rents nevertheless rise steeply in the face of the migrant influx. Since most new in-migrants are unskilled and poor, rents in squatter settlements rise more sharply than do rents for high-quality housing. Squatter-settlement rents reflect excess demand for housing and sites; the rents for high-quality housing mostly reflect increased land scarcity as the needs of the poor are partially accommodated by shifting land use. The cost-of-living advantage of the rural area rises sharply as a result of increased urban rents. All these factors raise the average unskilled wage in the cities and tend to restrict city growth in the short run. Furthermore, job creation is constrained by skill

bottlenecks (the skilled wage rises) and capital scarcity (the return to capital in manufacturing rises far above its return in other uses), which suggests that the long-run impact of unbalanced productivity advance on city growth is far greater than these short-run elasticities imply.

#### Scarcities of Fuel and Imported Raw Materials

Since urban-based manufacturing uses intensively fuels and raw materials (importables which we call Z-goods), any increase in the price of imported Z-goods penalizes manufacturing directly and other urban activities indirectly and thus inhibits urban job creation and city growth. Even though the model admits the possibility of direct and indirect substitution away from the more expensive imported fuels and raw materials, table 3-10 reports that in-migration and city growth are still constrained because urban activities tend to be Z-intensive. The elasticity of  $P_Z$ , however, is low compared

Table 3-9. *Short-Run Constraints on the City Growth Response to Unbalanced Productivity Advance (1970 Elasticities)*

<i>Endogenous variable</i>	$A_M$	$A_{KS}$
<i>City growth attributes</i>		
City growth rate	10.29	0.68
In-migration rate	20.57	1.36
<i>Congestion indicators</i>		
Percent urban land in squatter settlements	1.25	0.13
Squatter house rents	3.57	0.40
High-quality house rents	1.16	0.15
Cost-of-living differential (rural ÷ urban)	1.49	0.12
<i>Factor market disequilibrium indicators</i>		
Average urban unskilled wage	1.03	0.15
Rural unskilled wage	0.02	0.09
Urban skilled wage	1.28	0.13
Return to capital in M	2.34	0.27
Return to capital in KS	0.80	0.07
Return to capital in A	-0.30	0.06

Source: Kelley and Williamson (1984b), table 4.2, p. 105.

with the unbalanced productivity advance indicators (the  $A_j$ s) or with the other two exogenous prices reported in table 3-10. At first glance this may suggest that the sensitivity of urbanization to scarcity of fuels and raw materials has been overstated in the literature. In view of the historical record since 1960, however, one must be cautious. After all,  $P_Z$  soared in the 1970s after its recorded stability during the 1960s. Thus,  $P_Z$  may have been a dominant source of developing-country urbanization over the past two decades in spite of the modest elasticities reported in table 3-10. (The issue is discussed again below.)

#### Price Policy and World Market Conditions

Table 3-10 presents the short-run impact multipliers for prices of agricultural and manufactured goods, both of which compete in world markets. These prices are usually heavily distorted by external and internal price policy. The table shows that city growth was far more sensitive to  $P_A$  and  $P_M$  than to  $P_Z$ . Any effort to understand the sources of past and future city growth must therefore carefully sort out these relative price conditions, including the impact of past liberalization and future protectionist trends in the industrial countries and the influence of price-distorting policy regimes in developing countries.

#### Accumulation, Capacity, and Job Creation

Table 3-10 also summarizes the impacts of "productive" capital accumulation and population-sensitive

"unproductive" capital accumulation on employment demand, urban job creation, and city growth. Accumulation in the urban modern sectors fosters job creation, and an investment policy which favors manufacturing at the expense of agriculture fosters urbanization. It is surprising, however, that accumulation of urban squatter housing ( $H_{US}$ ) has the most potent short-run impact on urban job creation and city growth. This result occurs even though the impact multipliers exclude the employment effects of the formation of the housing stocks through investment (and investment activities in squatter dwelling construction are highly labor-intensive). The relatively large urban employment effects associated with the accumulation of squatter housing are therefore all indirect: housing rents de-

Table 3-10. *Factors That Influence City Growth in Developing Countries (1970 Elasticities)*

<i>Exogenous variable</i>	<i>Endogenous urbanization variable</i>		
	<i>Percent urban</i>	<i>City growth rate</i>	<i>In-migration rate</i>
<i>Land and labor</i>			
$R_A$	-0.03	-0.69	-1.38
$R_U$	0.04	0.88	1.77
$L$	-0.57	6.38	12.75
<i>Accumulation</i>			
$K_M$	0.09	1.82	3.63
$K_{KS}$	0.02	0.40	0.81
$K_A$	-0.03	-0.68	-1.35
$H_{KS}$	0	0.07	0.15
$H_{US}$	0.39	7.92	15.83
$H_{RS}$	-0.13	-2.75	-5.50
<i>Prices</i>			
$P_Z$	-0.04	-0.89	-1.77
$P_M$	0.54	11.13	22.23
$P_A$	-0.32	-6.51	-13.01
<i>Productivity advance</i>			
$A_M$	0.50	10.29	20.57
$A_A$	-0.26	-5.33	-10.65
$A_{KS}$	0.03	0.68	1.36
$A_{US}$	-0.01	-0.26	-0.52
$A_{RS}$	-0.03	-0.67	-1.33
<i>Other</i>			
Skilled labor force ( $S$ )	0.23	6.96	13.90
Foreign capital ( $F$ )	0	-0.03	-0.06

Note: Variables, in addition to those defined in table 3-8, are as follows:  $R_A$ , agricultural land stock;  $R_U$ , urban land stock;  $L$ , total unskilled labor force;  $K_M$ , physical (productive) capital in the manufacturing sector;  $K_{KS}$ , physical capital in the modern services sector;  $K_A$ , physical capital in the agricultural sector;  $H_{KS}$ , high-quality urban housing;  $H_{US}$ , low-quality (squatter) urban housing;  $H_{RS}$ , rural housing;  $P_Z$ , price per unit of imported raw materials;  $P_M$ , price per unit of manufactured goods;  $P_A$ , price per unit of agricultural goods.

Source: Kelley and Williamson (1984b), table 4.6, p. 111.

crease with the augmented supply of dwelling space, the relative cost of living in the cities declines, in-migration is fostered, nominal wages of the unskilled are lowered by the temporary labor glut, and employment expands everywhere in the city as a result, especially in manufacturing. In contrast to Coale and Hoover's (1958) emphasis on the tension between unproductive and productive capital accumulation in city growth, table 3-10 suggests no conflict: of the six alternative modes of accumulation listed, accumulation in  $H_{US}$  has the highest urban job creation and city growth effects. An issue for future work is to assess whether this conclusion holds for the longer run.

**Land and Labor**

Popular accounts of rapid urbanization in developing countries emphasize the role of high population growth. Table 3-10 contradicts this view, although how wrong the population account is depends on which aspect of urbanization and city growth is of interest.

Lewis (1977) has stressed the capital intensity of cities, and the present model conforms with that reality, since urban activities are, on average, far less labor-intensive than are rural activities. The Rybczynski theorem in trade theory holds that an increased endowment of any given factor of production should favor the expansion of those sectors which use that factor most intensively. Thus, whatever its source, population-induced labor force growth should foster the expansion of rural activities and suppress urbanization. According to this analysis, population growth does not offer an explanation for urbanization, and the negative impact multiplier in table 3-10 (percent urban,  $-0.57$ ) proves the point. For in-migration and city growth, however, table 3-10 reports a more conventional result, since the impact multipliers are positive and quite large. (The next section elaborates on this issue.)

What about land endowments? Conventional wisdom has argued that scarcity of agricultural land has tended to push labor into the cities. Although an extension of the arable land stock would certainly increase the retention of labor in rural areas and thus retard urbanization, the size of the impact reported in table 3-10 is small. Changes in agricultural land endowment are simply not an important part of the city growth tale.

**The OPEC Watershed and Recent Growth Trends**

The changing growth environment in the aftermath of the OPEC price shock has been significant. It has been manifested mainly in relative prices, which are captured

**Table 3-11. The Growth Environment, Pre-OPEC and Post-OPEC: Dynamic Parameters Assumed in the Simulations**  
(percent)

Exogenous variable (dynamic parameter)	Average annual growth	
	Pre-OPEC (1960-73)	Post-OPEC (1973-79)
Price of imported raw materials ( $\dot{P}_Z^*$ )	0	5.2
Price of M-goods ( $\dot{P}_M^*$ )	-0.7	-1.6
Agricultural land stock ( $\dot{R}_A^*$ )	1.0	0.5
Urban land stock ( $\dot{R}_U^*$ )	1.0	1.0
Total labor force ( $\frac{\dot{L}^*}{L+S}$ )	2.54	2.68

Source: Kelley and Williamson (1984b), table 5.1, p. 126.

in the model by exogenous trends in  $P_A$ ,  $P_M$ , and  $P_Z$  (table 3-11). In relation to prices of primary exports from developing countries, the quality-adjusted price of manufactured goods declined at an annual rate of 0.7 percent up to 1973. The rate of decline accelerated after 1973 and averaged 1.6 percent a year to 1979. In contrast, the price of imported raw materials (including fuels) rose by 5.2 percent a year after 1973 compared with  $P_A$ ; the same relative price exhibited long-run stability before 1973. Although these price trends are affected by the base periods selected, the averaging devices applied, and the underlying price series, there can be no doubt about the epochal character of the post-OPEC price trends which developing countries faced in the 1970s.

As table 3-11 shows, the growth rate of the labor force was higher after 1973, although the acceleration was quite modest. Table 3-11 also documents the assumed rates of farmland growth, showing a decline which captures the apparent rapid exhaustion of possibilities for augmenting extramarginal land. Clearly, there were other (perhaps less important) nonprice changes in the economic and demographic environment that surrounded the RDE group after 1973. For that matter, not all of the epochal price trends were related to OPEC. Nonetheless, we have labeled these two epochs pre-OPEC and post-OPEC.

**Evaluating the OPEC Watershed**

The world economy has been undergoing painful adjustments to the price shocks associated with short-run OPEC policy and long-run scarcities of raw materials and fuels. Since as late as 1979 the world economy was still digesting the impact of these disequilibrating shocks, it seems clear that a long-run general equilibrium model such as ours cannot be expected to account

adequately for the short-run trends that developing countries have undergone since 1979. Nonetheless we can use the post-OPEC conditions to illustrate the model's sensitivity to such epochal shocks.

Tables 3-12 and 3-13 summarize the city growth predictions. Table 3-13 offers a 1960-73 prediction, using the actual pre-OPEC economic and demographic environment, that has already been compared with actual experience in previous sections; a 1973-79 prediction using the actual post-OPEC economic and demographic environment documented in table 3-11; and a counterfactual post-OPEC prediction, which simply allows the pre-OPEC environmental conditions to continue beyond 1973. The counterfactual experiment makes it possible to assess what urbanization would have been like without the post-OPEC epochal shocks.

Comparison of the pre-1973 and post-1973 predic-

tions by the model in table 3-13 might suggest that the OPEC price shocks mattered little to subsequent urban performance in the late 1970s. The annual rate of city growth declines only modestly over the period as a whole (from 4.86 to 4.65 percent); the percent urban continues its rapid climb; and in-migration rates, although somewhat lower, remain high. Note, however, the unambiguous evidence of retardation in the Actual columns of table 3-12, in which the rates of in-migration and city growth lose their steam quite dramatically. More important, in table 3-13 compare the predicted actual post-OPEC experience with the counterfactual 1973-79 experience that assumes pre-OPEC environmental conditions: while the actual city growth rate for 1973-79 is 4.65 percent a year, the counterfactual rate would have been 6.04 percent. Furthermore, without the shocks of the late 1970s, the counterfactual rate of in-migration

Table 3-12. *Post-OPEC Urban Adjustments, Model Predictions, 1973-79*

Year	Actual post-OPEC environment			Counterfactual environment		
	Percent urban	City growth	Net urban in-migration rate	Percent urban	City growth	Net urban in-migration rate
1973	43.45	5.72	3.07	43.45	5.72	3.07
1974	44.45	5.10	2.32	44.78	5.75	3.10
1975	45.22	4.48	1.72	46.25	5.92	3.27
1976	46.26	5.03	2.26	47.80	6.03	3.38
1977	47.05	4.52	1.76	49.46	6.14	3.47
1978	47.85	4.47	1.71	51.22	6.23	3.56
1979	48.58	4.29	1.53	53.03	6.16	3.49

Note: The counterfactual results assume that the 1960-73 dynamic parameters persist after 1973. The net urban in-migration rate is the ratio of annual in-migration to the average urban population in all previous years.

Source: Kelley and Williamson (1984b), table 5.3, pp. 130-31.

Table 3-13. *Pre-OPEC and Post-OPEC Urban Adjustments, Model Predictions, Period Averages*  
(percent)

Variable	Basis for prediction		
	Actual pre-OPEC environment, 1960-73	Actual post-OPEC environment, 1973-79	Counterfactual environment, 1973-79
Annual city growth	4.86	4.65	6.04
Total increase in share urban for period	10.85	5.13	9.58
Annual increase in share urban	0.83	0.86	1.60
Net in-migration share of urban population increase	47.72	41.34	57.50
Net in-migration rate	2.35	1.83	3.30
Net out-migration rate	1.37	1.60	3.13

Note: The 1973-79 counterfactual results assume that the 1960-73 dynamic parameters persist after 1973. The net in-migration rate is the ratio of annual in-migration to the average urban population in all previous years. The net out-migration rate is the ratio of annual rural out-migration to the average rural population in all previous years.

Source: Kelley and Williamson (1984b), table 5.3, pp. 130-31.

would have risen to 3.30 percent rather than fall to 1.83 percent, as in fact happened.

### *Isolating the Sources of Slowdown in City Growth*

It seems clear that exogenous economic and demographic conditions had a powerful impact on developing-country urbanization during the 1970s. Table 3-14 reports an effort to isolate the most important factors. Nine counterfactual stimulations are used. Each generates a history in the 1970s (1973–79), but the table presents only one of the model's predictions—the annual rate of city growth. Each counterfactual case should be compared with the actual 1973–79 performance reproduced in column 1 (which repeats table 3-12). The fuel abundance counterfactual in column 3 maintains all of the exogenous conditions that underlie the actual performance in column 1 except fuel price behavior: while the OPEC-augmented actual  $\dot{P}_Z$  was 5.2 percent a year between 1973 and 1979, the counterfactual assumes that  $\dot{P}_Z = 0$ , as was indeed the case up to 1973.

The urban slowdown had little to do with agricultural land expansion or labor force growth. Rather, it appears that prices were doing most of the work. Furthermore, fuel scarcity, in spite of the attention it has received, was not nearly as important a source of slowdown in city growth as was the accelerated decline in the relative price of manufactures (lower  $\dot{P}_M$ ). This finding obviously supports the view that any future trend toward protectionism in industrial countries will play an important role in shaping city growth in developing countries in the next two decades. The same might be said, of course, of the mix of internal policies which may twist the relative price of manufactures in the future. This position is reinforced by the counterfactual experiment under “stable world markets,” in which the relative price of manufactures is held fixed during the 1970s. Under these more favorable price conditions for manufacturing the model predicts annual city growth rates of 6.49 percent, in contrast to the actual rate of only 4.65 percent.

Although a sharp decline in the rate of population growth would certainly have diminished the rate of city growth, column 8 in table 3-14 suggests that this influence has been grossly overdrawn in the popular literature. The counterfactual explores the impact of a spectacular reduction in population pressure from the actual RDE rate of 2.68 percent a year to the 0.9 percent which prevailed among industrial countries. Even under this enormous diminution in population growth, the rate of city growth would still have reached almost 4 percent a year for the period 1973–79. Rapid population growth

certainly has contributed to spectacular city growth rates, but the experiments suggest that it is not the central force that drives urbanization. Table 3-14 also suggests that a shift to foreign capital austerity would not matter much to urbanization.

Urbanization in developing countries is, however, sensitive to a productivity slowdown. Given a plausible retardation in the economywide rate of TFPG from 1.8 to 1 percent a year, city growth rates decline significantly, and the impact seems to show signs of cumulating over time.

### Conclusion

This chapter has shown the ability of the Kelley-Williamson model to adequately replicate growth, accumulation, distribution, and city growth in developing countries up to the 1973 watershed. On that basis it is concluded that the model can be used to analyze the sources of urban growth during the 1960s, the early 1970s, and the difficult period of structural adjustment since then and to analyze future trends in urban experience under varying world market conditions and domestic policy regimes.

To summarize, rapid rates of population growth are not—as a reading of the popular literature would suggest—the central influence that drives rapid urban growth in developing countries. Capital transfers to developing countries have not played a significant part, and rural land scarcity has had only a modest role. The most potent influences on city growth appear to have been the rate and imbalance of sectoral productivity advances—technological events which have favored the urban modern sectors—and prices (although OPEC-induced fuel scarcity has been less important than the relative price of manufactures). Thus, trade policy in the industrial countries and price policy in developing countries are likely to have the most important impacts on city growth in the next two decades.

### Notes

1. Algeria, Bangladesh, Brazil, Cameroon, Chile, Colombia, Costa Rica, Côte d'Ivoire, Dominican Republic, Ecuador, Arab Republic of Egypt, El Salvador, Ethiopia, The Gambia, Guatemala, Honduras, India, Indonesia, Kenya, Korea, Malaysia, Mexico, Morocco, Nicaragua, Nigeria, Pakistan, Panama, Paraguay, Peru, Philippines, Portugal, Sri Lanka, Swaziland, Syria, Taiwan, Thailand, Togo, Turkey, Uganda, and Yugoslavia.

2. The World Bank's *World Tables* normally report only total private savings and do not separate household savings from firm (corporate) reinvestment. Although the distinction

Table 3-14. Sources of a Slowdown in City Growth, 1973-79  
(percent)

Item	OPEC watershed counterfactuals, 1973-79						Other counterfactuals, 1973-79			
	Actual 1973-79 (1)	Total pre-OPEC environment (2)	Fuel abundance, pre-OPEC, $\dot{P}_Z = 0$ only (3)	World markets, pre-OPEC, $\dot{P}_M$ only (4)	Land expansion, pre-OPEC, $\dot{R}_A$ only (5)	Population pressure, pre-OPEC, $\frac{*}{L+S}$ only (6)	Stable world markets, $\dot{P}_M = 0$ (7)	No population pressure (developed country rate) (8)	Technological slowdown (decrease in TFPG from 1.8 to 1 percent) (9)	Foreign capital austerity, $F = 0$ (10)
Year										
1973	5.72	5.72	5.72	5.72	5.72	5.72	5.72	5.72	5.72	5.72
1974	5.10	5.75	5.35	5.59	5.09	5.06	5.95	4.46	5.10	5.21
1975	4.48	5.92	4.91	5.67	4.51	4.50	6.46	4.03	4.27	5.66
1976	5.03	6.03	5.28	5.90	4.95	4.96	6.51	4.37	4.60	4.37
1977	4.52	6.14	5.13	5.91	4.47	4.48	6.64	3.72	4.22	4.33
1978	4.47	6.23	5.05	5.96	4.36	4.36	6.63	3.68	3.93	4.21
1979	4.29	6.16	4.83	5.79	4.27	4.28	5.76	3.57	3.82	4.24
Average	4.65	6.04	5.09	5.80	4.61	4.60	6.49	3.97	4.32	4.67
<i>Exogenous variables taken to be:</i>										
$\dot{P}_Z^*$	5.2	0	0	5.2	5.2	5.2	5.2	5.2	5.2	5.2
$\dot{P}_M^*$	-1.6	-0.7	-1.6	-0.7	-1.6	-1.6	0	-1.6	-1.6	-1.6
$\dot{R}_A^*$	0.5	1.0	0.5	0.5	1.0	0.5	0.5	0.5	0.5	0.5
$\frac{*}{L+S}$	2.48	2.54	2.68	2.68	2.68	2.54	2.68	0.9	2.68	2.68
$\dot{F}^*$	← $\dot{F}^*$ such that $F/\text{GDP} = 3$ percent →									
TFPG	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.0	1.8

Note: Variables are defined as follows:  $\dot{P}_Z^*$ , fuel and raw materials prices;  $\dot{P}_M^*$ , domestic price of manufactures;  $\dot{R}_A^*$ , agricultural land stock;  $\frac{*}{L+S}$ , population (or labor force);  $\dot{F}^*$ , foreign capital inflow; TFPG, economywide productivity growth.

Sources: Kelley and Williamson (1984a), table 5, p. 435; (1984b), table 5.4, p. 133.

may be fuzzy in the economies being analyzed, the two sources of savings are quite distinct in the model and are thus reported separately. Furthermore, what the model identifies as training investment is embedded in private and government consumption in the RDE national accounts.

## Bibliography

Becker, Charles M., Edwin S. Mills, and Jeffrey G. Williamson. 1986. "Modeling Indian Migration and City Growth, 1960–2000." *Economic Development and Cultural Change* vol. 35, no. 1 (October), pp. 1–33.

Beier, George, Anthony Churchill, Michael Cohen, and Bertrand M. Renaud. 1975. *The Task Ahead for the Cities of Developing Countries*. World Bank Staff Working Paper 209. Washington, D.C. Also in *World Development*, vol. 4, no. 5 (May 1976), pp. 363–409.

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

Coale, A. J., and E. M. Hoover. 1958. *Population Growth and Economic Development in Low-Income Countries*. Princeton, N.J.: Princeton University Press.

Kelley, Allen C., and Jeffrey G. Williamson. 1980. *Modeling Urbanization and Economic Growth*. RR-80-22. Laxenburg, Austria: International Institute for Applied Systems Analysis. May.

———. 1982. "The Limits to Urban Growth: Suggestions for Macromodelling Third World Economies." *Economic Development and Cultural Change*, vol. 30, no. 3 (April), pp. 595–623.

———. 1984a. "Population Growth, Industrial Revolutions, and the Urban Transition." *Population and Development Review*, vol. 10, no. 3 (September), pp. 419–44.

———. 1984b. *What Drives Third World City Growth? A General Equilibrium Approach*. Princeton, N.J.: Princeton University Press.

Keyfitz, Nathan. 1980. "Do Cities Grow by Natural Increase or by Migration?" *Geographical Analysis*, vol. 12, no. 2 (April), pp. 142–56.

Ledent, Jacques. 1980. *Comparative Dynamics of Three Demographic Models of Urbanization*. RR-80-1. Laxenburg, Austria: International Institute for Applied Systems Analysis. February.

Lewis, W. Arthur. 1977. *The Evolution of the International Economic Order*. Princeton University, Woodrow Wilson School, Research Program in Development Studies, Discussion Paper 74. Princeton, N.J.

Linn, Johannes F. 1979. *Policies for Efficient and Equitable Growth of Cities in Developing Countries*. World Bank Staff Working Paper 342. Washington, D.C.

Mohan, Rakesh. 1979. *Urban Economic and Planning Models: Assessing the Potential for Cities in Developing Countries*. Baltimore, Md.: Johns Hopkins University Press.

Preston, Samuel. 1979. *Patterns of Urban and Rural Population Growth*. New York: United Nations, Population Division.

United Nations. 1976. *Global Review of Human Settlements: A Support for Habitat*. Vols. 1 and 2. Oxford: Pergamon Press.

World Bank. 1976. *World Tables 1976*. Baltimore, Md.: Johns Hopkins University Press.



# Part II

---

## *A Framework for Urbanization Policy*

**T**he explanations of urbanization suggested in part I provide a framework of economic development within which to discuss problems connected with urbanization. The policy issues addressed in part II are presented in the context of overall development and the macroeconomic settings of countries. Tolley (chapter 4) goes to the economic sources of widely noted urban problems, Renaud (chapter 5) discusses policy experiences from broad countrywide and cross-country perspectives. This broad approach is intended not to minimize the urban-specific sources of various urban problems, but to bring out some important and sometimes neglected overall considerations which have pervasive effects on the urban economy. Parish (chapter 6) examines urban problem-solving in China's centralized policy framework.

### **Urban Problems and Policy Implications**

The most basic urban problem, as noted in chapter 4, is poverty. The extent of urban poverty, in the first instance, can be understood in the context of the rate and nature of economic development and the generation of employment. Given the level of development, rural-urban migration helps to reduce differences in real incomes between urban and rural areas and bring about an increase in urban poverty. Sustained and rapid development is in most instances a necessary, although not always a sufficient, condition for alleviating urban poverty. Attention to the distribution of the benefits of development is also essential.

This is not to deny that urbanization can itself contribute to problems, including poverty. Many countries face excessive urbanization and the associated problems of overcrowding and deplorable living conditions. In economic terms overurbanization occurs when negative

externalities such as pollution and congestion cause cities to be seen as "too big." Measurement of these effects, however, reveals that the net economic costs of overurbanization (compared with the alternatives) are surprisingly small. Furthermore, these externalities are best dealt with directly rather than by trying to control urbanization or city size. A more satisfactory size distribution might also come about in the absence of urban biases and such incentives as protected employment and special subsidies in urban areas.

In developing countries the disproportionate concentration of populations in big cities makes the scale of overurbanization more dramatic than it has ever been. One reason for urban concentration is that transport costs within cities are lower today than when the developed countries were undergoing rapid urbanization. Because economies of association can now be realized more cheaply within a city, it is easier to locate ancillary local production and the needed housing in the city. Infrastructure planning and construction may also contribute to disproportionate expansion in cities that are already large, as growth feeds on growth.

### **The Scope of Urbanization Policy**

Chapter 5 provides an account of the range of analytical and policy problems that arise in the context of developing-country urbanization, with examples drawn from different continents. The author also examines the differences between the urban policy problems that confront the developing countries and those that the developed countries face. Two structural changes in the developed world help explain the differences: the slowing and eventual end of urbanization, and the emergence of yet another industrial revolution, this one

based on new technologies and less tied to concentrated manufacturing centers.

Problems in developing countries that are associated with urbanization include spatial inequalities, such as disamenities as congestion and pollution, and the need to create employment and provide services. Urbanization policies actually followed by developing countries can be classified as national economic policies, explicit spatial and regional policies, and city management. Their relative importance varies widely from country to country.

Problems of congestion, pollution, and internal city inefficiency are best addressed directly through good city management. The attempt to control city population directly is an inefficient and costly way of dealing with those problems. At the same time it should be recognized that national economic and sectoral policies often have spatial biases. Reversal of such biases, particularly those that hurt rural production, will minimize pressures to control city population directly. Explicit spatial policies may seem desirable, particularly in situations in which the nation's economic policies have, intentionally or unintentionally, produced urban biases. The efficacy of direct spatial initiatives must be reviewed carefully, however, particularly since experience with them is not encouraging.

### A Case of Central Planning

Chapter 6 uses data from a variety of sources to piece together a view of modern cities in China and sheds light on some major goals of Chinese urban policy. One goal has been balanced city size distribution and stability. In the 1950s investment funds and activities were channeled from coastal cities and reinvested in the interior. Tough migration laws and near-total control of social institutions by local and central authorities furnished the necessary muscle. The study offers a detailed review of recent urbanization in China, which contrasts with that in most developing countries.

Another goal has been secure, productive urban employment, guaranteed by the state and devoted to centrally defined output objectives. Several reasons are cited for the rise in the employment rate of the nonagricultural population: the eviction from cities of unauthorized people, the increasing employment of women, the sharp drop in urban birth rates, and efforts to create new urban employment opportunities. The largest employers of the urban population are the various levels of government, and state employees in China seem to enjoy greater benefits than do workers in other developing countries. In the smaller employment sectors, collective enterprises and the private work force, wages are lower, benefits are fewer, and employment is less secure than in the state sector.

The distinction between urban work in China and elsewhere has been highlighted by differences in the distribution of nonagricultural labor: the Chinese have emphasized work in manufacturing, mining, and construction and have downplayed consumption-related activities. The consequences have been less varied and less plentiful consumer goods and rationing and queuing. An additional problem has been the unemployment, especially among young people, that has accompanied investment in capital-intensive heavy industry. Many educated young workers were sent to the countryside to relieve urban unemployment, and unemployment rose to 5–11 percent of the nonagricultural labor force. Reforms after 1978 loosened investment policies and restrictions on work opportunities and made public education more responsive to the needs of industry. By the end of 1981, outright unemployment had fallen to 2.4 percent of the nonagricultural labor force.

Another aim has been to provide adequate urban services. The authorities have displayed ambivalence toward urban services and have agreed only that there should be an equitable distribution of basic needs. Medical services have grown steadily and are of higher quality than in other developing countries; infant mortality is lower and life expectancy higher. Urban housing has received less attention, partly because it has been considered a consumption good, but Chinese urban housing compares favorably with that in other developing countries. The study discusses the unintended side effects of rent control and the recent sharp increases in state funds to remedy shortcomings in housing.

### Bibliography

- Birdsall, Nancy. 1980. *Population and Poverty in the Developing World*. World Bank Staff Working Paper 404. Washington, D.C.
- Henderson, J. Vernon. 1977. *Economic Theory and the Cities*. New York: Academic Press.
- Lipton, Michael. 1977. *Why Poor People Stay Poor: Urban Bias and World Development*. Cambridge, Mass.: Harvard University Press.
- Mills, Edwin, and Charles Becker. 1982. "Urbanization and Economic Development." Princeton, N.J.: Princeton University. Processed.
- Thomas, Vinod. 1982. *Differences in Income, Nutrition and Poverty within Brazil*. World Bank Staff Working Paper 505. Washington, D.C.
- World Bank. 1979. *World Development Report 1979*. New York: Oxford University Press.
- . 1981. *World Development Report 1981*. New York: Oxford University Press.
- . 1982. *World Development Report 1982*. New York: Oxford University Press.

## *Market Failures as Bases of Urban Policies*

George S. Tolley

Urbanization in developing countries remains controversial. At the heart of many economic development efforts has been promotion of industrialization, which fosters the growth of urban centers. But urbanization has also been viewed as undesirable because it takes people away from a rural way of life that is thought to be intrinsically better and because of the economic and social problems found in cities.

Urban poverty has been viewed as one of the ills of growth. Recurrent assertions that overurbanization has taken place refer to environmental problems in cities, wage policies that induce excess rural-to-urban migration, and government and tax policies that favor cities. The arguments have usually been made in qualitative terms and have seldom been related seriously to the functioning of markets.

This chapter offers an analysis of these and other issues that bear on urbanization performance. The causes of urban poverty are first brought out, and the relation of poverty to overall economic development considerations is discussed. Attention then turns to accounting problems that distort measured changes in national income when urbanization occurs.

Overurbanization is then analyzed. Models with measurement are developed that deal with pollution, congestion, urban wages, and government fiscal policies. The models give indications of the city size effects and income losses attributable to urbanization externalities. The reasons why large cities in developing countries are so large and are getting larger are analyzed, with special attention to transport.

The final section draws policy implications. Possibilities are considered for internalizing externalities as a first-best policy that is superior to trying to influence urbanization as such. Options and pitfalls in undertak-

ing direct spatial policies to influence urbanization are discussed, and the role of externalities in nonspatial policies is considered. The need to improve methods for evaluating the spatial effects of infrastructure decisions is emphasized, and evaluation needs in the management of urban services in general are brought out.

### Urbanization and Poverty

The "rings of thorns"—or shantytowns on the outskirts of many of the large cities of the developing world—have been the subject of much concern, and they are, at least subconsciously, partly responsible for the view expressed by some that excessive urbanization has occurred. The growth of the numbers of poor people is an especially noted symptom of urbanization, and there may be a lurking hope that curbing urbanization would reduce poverty.

Interest persists in estimating the numbers of people in poverty according to some definition. Aside from the problem of choosing the level of income below which poverty is deemed to exist (a workable procedure if the choice is recognized as arbitrary), the problem arises of measuring the extent of poverty among different groups within the population. That task is made difficult by the varying circumstances of people, which make simple comparisons of money income invalid.

Explaining the extent of urban poverty requires an understanding of what people earn and is more difficult than explaining the overall degree of urbanization (discussed in chapter 2). Yet a foremost reason why poor people are found in large numbers in cities of developing countries is the sheer overall magnitude of poverty in those countries and its distribution among urban and

rural areas because of migration. Through migration, rough equalization of real income levels between urban and nonurban areas can be expected to be achieved within each income stratum, from rich to poor. Lack of knowledge of opportunities in different areas is not an apparent barrier to migration, nor is lack of such contacts as friends and relatives who can help the migrant get settled in a new environment.

Equalization of real earnings between rural and urban areas may not be exact. Whether equalization takes place cannot be proved or disproved by looking at urban and rural money incomes at any one time. Observed money incomes are not directly commensurate in rural and in urban areas. People place different values on differences in the availability of goods, prices, and general environments. Thus the equalization is not between money incomes but rather between real incomes, including the values people place on differences between rural and urban areas.

The hypothesis that a rough equalization of real earnings takes place is supported by two arguments. The first is introspective: it does not stand to reason that people, in making choices about where to live, will ignore opportunities in other places. If they pay some attention to alternatives, differences in real earnings between places will not get far out of line. The second argument is empirical: if people did in fact ignore real income differences, the erratic pace of demand for products and supplies by people in different places, as determined by demographic considerations, would produce wide swings in observed relative differences in money incomes. Although some annual fluctuations in the ratio of money incomes between rural and urban areas take place, over the years the fluctuations are limited. Thus, explaining poverty in cities becomes a matter of explaining how the demand for the services of poor people in urban areas interacts with their conditions of supply. The same type of explanation applies to people in rural areas. The supplies in each case are determined by the total number of persons of a given type in the economy.

Suppose we let  $N_u$  and  $N_r$  refer not to the total number of persons in urban and rural areas, as in chapter 2, but to numbers of persons who supply unskilled or common labor in the two types of area. Then the same kind of explanation of the rate of urbanization of the total population that was given in chapter 2 can be used to explain the rate of urbanization of unskilled labor. The income elasticities will reflect the income elasticities of the final products in which the unskilled labor is used, and they will also reflect the effects of the accumulation of physical capital and the growth in numbers of skilled laborers on the demand for unskilled labor. Whereas the demand for unskilled labor in rural areas consists primarily of demands for field work and lower-income

farming, the demand in urban areas consists of demands for common labor in a variety of activities, including factory work and services. Services in urban areas are supplied through work in formal firms and as domestic help as well as in casual and informal employment. Common labor is itself an aggregation and simplification that masks many gradations of productivity within the unskilled category. In speaking of the poor, one is speaking primarily of families whose breadwinners are at the lower ends of the unskilled distribution.

Although, as noted, the demand for unskilled labor is affected by accumulations of physical capital and growth in the number of skilled laborers, these two types of productive factors will also tend to grow as the output of urban commodities rises. The number of unskilled laborers, the number of skilled laborers, and the amount of physical capital will all tend to grow with urbanization. Performance levels of skilled laborers may rise, owing to education and to experience with more sophisticated techniques, and physical capital may accumulate more rapidly than population. These developments may affect the demand for unskilled labor, but they operate to some extent in rural as well as in urban areas.

Changes in the numbers of unskilled laborers in urban and rural areas often go along, more or less, with changes in urban and rural populations. Without claiming an exact proportionality, it seems that changes in unskilled labor and population in an urban area tend to move similarly. If so, the explanation of quantitative changes in percentage terms for the urbanization of unskilled labor is similar to that already considered for the population as a whole. In some countries lower-income and subsistence agriculture become outmoded, leading to a low or even negative apparent income elasticity for unskilled labor in agriculture. In this case the general approach to the explanation of urbanization of the unskilled population remains the same as for the total population, but the quantitative outcome is different. The implication in this case is that the rate of urbanization of the unskilled population will be more rapid than that of the population as a whole.

There are three reasons why the absolute level of poverty diminishes during development. First, the application of new techniques and capital raises the marginal productivity of unskilled labor. This phenomenon is in many ways the essence of development. If development does not succeed in raising the well-being of all strata, it does not succeed in its purpose. Moreover, looking at the various gradations of development around the world, one sees that a primary distinguishing feature among countries is the extent to which the level of well-being of all strata appears to be raised.

It may be hypothesized that absolute poverty tends to decline at first slowly and then at an accelerating rate. If

labor is paid its marginal product, the percentage rate of increase in wages of unskilled labor—which determines the rate of decline of poverty—will equal the percentage rate of increase in marginal productivity. The percentage rate of increase in marginal productivity depends in turn on changes in techniques and in quantities of associated factors of production that directly raise marginal productivity and on changes in the proportions of labor and other factors of production used in a given activity. If more unskilled labor is used with the other factors of production, its wage will go down, and if less is used its wage will go up. For the economy as a whole, any increase in unskilled labor in one activity must be accompanied by decreases elsewhere, assuming that there is a given amount of unskilled labor in the economy at any one time. Moreover, because real incomes for unskilled labor tend to be equalized among activities, the percentage rate of change in wages of unskilled labor will tend to be the same in different activities.

Combination of these conditions gives insights on how wages of unskilled labor in an economy change over time. The percentage change in the wage,  $\dot{w}$ , will equal the increase in marginal productivity  $\dot{A}$  (attributable to improvements in techniques and changes in amounts of associated factors) plus the product of the elasticity of demand for labor in an activity,  $\beta$ , and the percentage change in employment of labor in that activity,  $\dot{n}$ . The economy is then divided into urban and nonurban activities, and the conditions for the percentage change in wage are  $\dot{w} = \dot{A}_u + \beta_u \dot{n}_u$  and  $\dot{w} = \dot{A}_a + \beta_a \dot{n}_a$ . The first condition is multiplied by the proportion of the labor which is urban,  $n_u$ , and the second condition by the proportion which is nonurban,  $n_a$ , and the results are added. Since the total of the proportions is unity ( $n_u + n_a = 1$ ), the left-hand side of each equation, after adding, equals  $\dot{w}$ . On the right-hand side, the elasticities of demand for labor,  $\beta_u$  and  $\beta_a$ , will be equal if the factor shares of the labor are the same in urban and nonurban activities, if their elasticities of substitution with other factors are the same, and if the price elasticities of demand for the products produced are the same. These assumptions are not likely to be fulfilled exactly, but they serve as a useful benchmark approximation. The assumptions will in fact be fulfilled if the product prices are given, as may happen if the economy is importantly involved in international trade and if the aggregate production functions for the two types of activities are similar as regards shares and elasticities of substitution—which again may not be unreasonable. If the elasticities are equal, one part of the right-hand side will consist of two terms which combine to form the expression  $n_u \dot{n}_u + n_a \dot{n}_a$ . But since  $\dot{n} = dn/n$  and  $dn_1 + dn_2 = 0$ , in view of the given amount of the labor in the economy,  $n_u \dot{n}_u + n_a \dot{n}_a$  equals zero and the expression

disappears. All that is left on the right-hand side is the weighted sum of the increases in marginal productivity that would occur in the absence of changes in amounts of labor in the different activities. The percentage change in wage is

$$(4-1) \quad \dot{w} = n_u \dot{A}_u + n_a \dot{A}_a.$$

Refinements to allow for differences in elasticities of demand for labor could be introduced, but the terms shown would probably still dominate the results. The result shown generalizes to many activities and can be written  $\dot{w} = \sum n_i \dot{A}_i$ .

The percentage change in the marginal product of labor in agriculture,  $\dot{A}_a$ , may be small, either because progress in agriculture is slow or because a low-income or subsistence sector makes up a large part of agriculture and is not progressing. At the early stages of development, when most labor is in agriculture,  $n_a$  will be close to one and  $n_u$  will be close to zero. Even though there is rapid progress in urban productivity, and hence a large value of  $\dot{A}_u$ , the percentage rise in the wage will be small because of the low weight of  $\dot{A}_u$ . As urbanization proceeds during development, the fall in  $n_a$  and the rise in  $n_u$  will increase the weight of  $\dot{A}_u$  and decrease the weight of  $\dot{A}_a$ . As a result, the rise in the wage rate will accelerate until, if the extreme of complete urbanization occurred, the rise in the wage would be fully equal to the upward shift in urban productivity.

Besides the rise in the return to unskilled labor, owing to improved techniques and the use of associated factors of production, two additional reasons for a decline in absolute poverty may be noted. First, people acquire skills and move out of the unskilled labor category. They do this because economic development, by raising the marginal productivity of skilled labor, increases the returns to education and other human capital investments by which people transform themselves from unskilled to skilled laborers. All persons who do so lift themselves from poverty categories to a higher-income status.

The second reason is closely related to the first. After some people have transformed themselves into skilled laborers, there are fewer unskilled laborers in the economy. Unskilled labor becomes scarcer than if the same proportion of the population had remained unskilled. In view of the elasticity of demand for labor, the greater scarcity of unskilled labor will lead to a rise in its wage.

### Urbanization as a Source of Growth: An Illusion

The fact that per capita money incomes are generally higher in urban than in rural areas leads to an increase in measured per capita income under urbanization. In

growth source analysis, which relates changes in national output to changes in inputs, the rural-urban shift shows up as a contribution to growth because differences in factor earnings are assumed to reflect real differences. The result is sometimes taken as an indication that urbanization is itself a source of growth, and thus that urbanization leads to additional development beyond that attributable to the demand and supply shifts that have been analyzed.

That the rural-urban shift is a source of growth may be partly or even wholly an illusion. A prerequisite of meaningful interpretation of the rural-urban shift is that education and other sources of differences in the average quality of labor in rural and urban areas be adequately measured. Two concomitant changes in economic development are, ordinarily, a shift of less-educated labor from rural to urban areas and an increase in education levels. If the increase in education levels is greater in urban than in rural employment, the shift from rural to urban employment will appear to give rise to growth unless the differences in education and other reasons for skill differences are carefully controlled. What is really a contribution to growth by increases in human capital may appear as a contribution by the rural-urban shift.

Even after differences in quality of inputs have been taken into account, measured contributions of the rural-urban shift may remain which are illusory. The fact that nonurban areas have more home production and other production which does not pass through the marketplace and therefore cannot be measured is often pointed out.

A perhaps more important consideration is that the prices of local goods tend to be higher in urban than in nonurban areas. A prime example is residential rent, but the phenomenon also pervades services. Higher money wages that make labor as well off in urban as in rural areas—as will tend to be brought about by migration—will have a multiplier effect and will raise prices of local goods produced by the labor. The higher money wages have a feedback effect and raise money wages still more. Thus, in a market economy with labor mobility, the prices of local goods can be expected to vary, and differences in money wages will appear even if real wages are equated in different locations. Differences in prices between areas may be wrongly interpreted as causing changes in total output when there is a rural to urban population shift.

Whether real effects of the rural-urban shift remain after measurement is carried out correctly is moot. It is sometimes said that a differential in real earnings between rural and urban areas is required to induce the shifts in resources that occur during growth. Expectational considerations should be recognized, however. A differential may be required only if the shifts are not

foreseen. If the shifts are foreseen—and the fact that the rural-urban shifts are large and of long duration during development gives reason to believe that they are—people can still make migration choices that equate real earnings. A differential might be required to compensate for the costs of moving, but empirically these costs appear to be small. A factor that could actually make urban wages lower in real terms than earnings in rural areas is the prospect of greater opportunity for increases in earnings in cities. If there is little opportunity for career advancement in rural areas, lifetime rural income profiles will be flatter. The steeper profiles in urban areas could lead migrants from rural areas to accept lower entry-level earnings in urban areas.

A final possibility, and perhaps the only actual source of real effects of a rural-urban shift, is the delayed decline of subsistence-type agriculture. As emphasized earlier, whole subsectors may become outmoded, in which case there is likely to be a generational process of replacement of human capital. Assuming that this large-scale change was not foreseen, older persons may be found in low-income career paths in subsistence-type agriculture even though their earnings are lower than earnings available elsewhere. Their offspring who replace them in the labor force then migrate to cities and find more productive employment. This intergenerational shift from low-productivity to high-productivity employment leads to an increase in total national income. It should be noted that this is not an independent source of growth but rather a manifestation of a lag in adjustment; it reflects a failure to respond promptly to past sources of growth.

### Has Overurbanization Occurred?

So far in this chapter no reasons have been found to question market performance in response to the demand and supply changes that determine urbanization. This section considers possible benefits and costs of the process of urbanization that people's responses do not take into account.

### *Environment*

As noted above, differences in living conditions between rural and urban areas cause money income differences because of people's responses to these living conditions. The question now is whether people, in their migration responses that influence the degree of urbanization, impose costs or gains on other people which those making the decisions have little or no incentive to consider. Two of the many costs which are not fully internalized are pollution and congestion.

Commuting costs, pollution, and congestion are among the many reasons that money wages are generally observed to be higher in urban than in nonurban areas. New entrants to a city pay the full costs of longer commuting distances in the form either of long travel to work or of high rents if they do not commute from the agricultural edges of the city. New entrants likewise bear higher costs of pollution and congestion than in rural areas, but only the average of such costs—not the extra costs that they impose on all others in the city because their presence leads to additional air pollution and slowing of traffic.

A way of approaching the question of market performance is to ask how much higher money wages must be in a larger city than in a smaller one to compensate for the increased costs in the former. As a simplified numerical example, take a developing-country city with a population of 600,000 and a work force of 150,000 persons employed full-time at an average annual wage of \$2,000.

Consider first the commuting distance effect. When a laborer migrates from farm to city, the edges of the city are moved farther out into the agricultural periphery. Anyone living in from the periphery must pay higher rents because of the commuting advantage of interior residential sites over peripheral locations. As the city grows money wages tend to increase in relation to money wages in nongrowing places. Suppose that in such a city each worker commuting from the periphery of the city spent, on average, 15 minutes longer a trip traveling to work than he would from peripheries of typical smaller cities. If the extra time is multiplied by 250 work days a year and two trips a day, and each hour is valued at \$0.60, the added cost of travel time in this city is \$75 a year. To compensate, the wage in the large city would have to be high enough to yield at least this much additional income.

Next, consider the pollution effect. Particularly in the drier summertime, dust and other suspended particulates are an important source of pollution. Much of this is natural, but human activity also causes much of it, and smaller cities are therefore generally less dusty than larger ones. A typical level of suspended particulates would be 150 micrograms a cubic meter of air, whereas in smaller cities and towns a representative level would be perhaps 100 micrograms. Studies of the relation between property values and air quality indicate that individuals are willing to, and in fact do, pay higher rents to escape unfavorable air quality. There have been numerous more direct estimates of the physical damage caused by pollution. A rough estimate of the annual cost of pollution is \$1 per microgram for a family of four persons, a figure that reflects damage to health, shortened life spans, deterioration of durable property, and

impairment of aesthetic values. This damage value would imply that the cost of pollution to a family in the larger city is \$50 more (50 micrograms difference multiplied by \$1) than the cost to a counterpart family in a smaller city with better air. If air pollution were the only unusually adverse characteristic of larger cities, and if the population were stable, the wage in the city of 600,000 would have to be \$50 more than in smaller places.

If the pollution effect is added to the distance effect, the estimated amount by which wages in the larger city must exceed wages in smaller places to compensate for the disamenities of the larger city is \$125. This is a conservative estimate because many of the numerical estimates used in the derivation are conservative and because a number of disamenities (such as other pollutants and noise, visual blight connected with slums, and crime) have been neglected. If these figures were correct and if no other factors influenced decisions to migrate, people would disperse themselves among cities in such a way that the wage rate was, say, \$2,000 in smaller cities and \$2,125 in the larger city to equalize the advantages of living in one or the other city.

The factors considered so far represent average costs per worker or per family, which are higher in larger than in smaller cities. The question becomes whether the addition of a worker in a city adds the same amount to costs, or a greater amount.

One reason the costs may be greater is that travel for everyone else in the city may be slowed by the extra congestion. Suppose that in a city of 600,000 half of the labor force, or 75,000 workers, commutes to the central employment location. Each travels an average of 5 miles per trip over congested streets at a speed of 20 miles per hour (the speed would be 35 miles per hour if there were no traffic congestion). If half the new entrants commute to the center of the city (the same assumption as for the existing labor force), the average new entrant into the city's labor force makes five rush-hour trips per week over this highway system. The effect on traffic speeds and on commuting times for the other motorists can be estimated through the use of the engineering relationship  $v = \bar{v} - kq$ , where  $v$  is the speed under congested conditions,  $\bar{v}$  is the speed under noncongested conditions (35 miles per hour),  $q$  is the traffic volume (75,000 cars per hour), and  $k$  is a constant (here,  $2 \cdot 10^{-4}$ ), giving  $v$  equal to 20 miles per hour (the existing speed noted above). Now recalculate  $v$  with  $q = 75,001$  cars per hour to estimate a new speed. Subtract the old from the new speed to obtain the change in speed. Divide this change into the 5-mile trip on congested streets and multiply by the value of time, \$0.60 per hour, to obtain the added cost that an extra trip imposes on a trip taken by someone already in the

city,  $\$1.50 \cdot 10^{-6}$ . Multiply this result by 0.5, since the commuting assumption implies that only one-half trip per new worker is added to the congested streets. Then multiply by the 75,000 affected workers and by the 500 work trips they each make per year (250 work days times two trips per day—home to work and work to home) to arrive at a yearly figure of \$28 extra commuting cost that a new worker imposes on the population already in the city.

The impact of the additional family on damages from air pollution can be computed with the use of data already cited. An inference from observations on air quality and metropolitan population size is that the average particulate level is likely to rise by  $4 \cdot 10^{-5}$  micrograms a cubic meter as a result of adding one family. If the damage value of \$1 a microgram is applied and it is assumed that there are 150,000 families in the urban area, the overall damage value is 0.00004 times \$1 times 150,000, or \$6.

Taking account of both congestion and pollution, the additional costs imposed on all other families when a family moves into an urban area of 600,000 population amount to \$28 plus \$6, or \$34 a year.

These results should be considered in relation to the previous conclusion, that wages in this city would be about \$125 higher than in other places of smaller size. When a family moves into the area, if there is a free market equilibrium of city sizes generally, the personal gain is the \$2,125 wage, for which \$125 urban costs are incurred in this city and a \$2,000 wage elsewhere is forgone. The other families in the large city suffer a combined loss of about \$34. Thus there is a net loss of \$34.

Although the example is only suggestive, it gives a perspective on the magnitude of the effects connected with the physical environment. The \$34 annual cost neglected in migration decisions is 1.6 percent of the wage rate. The demand for labor in any one city may be quite elastic owing to possibilities for undertaking production elsewhere. For example, if the elasticity of demand for labor is unity, the number of people in the city is made 1.6 percent larger by the neglected costs. If the elasticity is as high as 3, the city is made about 5 percent larger.

### *Protected Employment*

A phenomenon that has received much attention is the attraction to cities of persons who are willing to accept unemployment in return for being in a position to obtain protected employment should the opportunity arise. Examples of protected employment include employment effectively covered by legally decreed minimum wages, government employment in which wages

of employees are set through a political or negotiating process, and situations in which foreign corporations may decide for a number of reasons to pay more than the market-clearing wage. The wage in protected employment is higher than elsewhere in the economy.

A simple, well-known condition that has been used to analyze protected employment is based on the hypothesis that workers will distribute themselves so that the wage received in unprotected employment ends up being just as great as expected or prospective earnings in protected employment, taking account of the greater probability of unemployment that workers face if protected employment is sought. People are willing to face the unpleasant prospect of greater unemployment, which requires drawing down savings or asking kin for largesse, in return for the hope of higher earnings if they eventually succeed in finding a protected job. A queue of people seeking protected jobs forms, and it grows until expected well-being in the protected employment sector is driven down to that in unprotected employment.

In urban protected employment, which is not all urban employment and may be only a small part of it, the probability of employment is sometimes approximated as the ratio of total protected employment,  $n_{ue}$ , to the total number of people seeking protected employment,  $n_u$ . The rationale is that the total supply provides a pool from which those seeking employment have a chance of being drawn. The wage condition is that the wage in unprotected employment equals the expected wage in protected employment, which in turn is the probability of employment times the wage received if employed in protected employment, giving  $w_a = (n_{ue} / n_u)w_u$ . In terms of percentage changes the wage condition is  $\dot{w}_a = \dot{n}_{ue} - \dot{n}_u + \dot{w}_u$ , which is the first of five conditions for this model.

Meanwhile, the change in total employment in the economy as a whole as a consequence of a rise in the protected employment wage is the weighted sum of percentage changes in protected and unprotected employment. This implies as a second condition  $n_e \dot{n}_e = n_a \dot{n}_a + n_{ue} \dot{n}_{ue}$ . As in the earlier analysis, the weighted sum of percentage changes in people seeking employment is zero, given the total population in the country. This implies as a third condition  $n_a \dot{n}_a + n_u \dot{n}_u = 0$ . Other conditions from the earlier analysis are that the percentage changes in employment equal the percentage change in wage multiplied by the elasticity of demand for labor. The fourth and fifth conditions are thus  $\dot{n}_a = \beta_a \dot{w}_a$  and  $\dot{n}_{ue} = \beta_u \dot{w}_u$ .

The five conditions determine the percentage effects of  $\dot{w}_u$  (the percentage excess of the protected employment wage over the unprotected wage) on the unprotected wage, total employment, unprotected employment, protected employment, and numbers of people

seeking protected employment ( $\dot{w}_a$ ,  $\dot{n}_e$ ,  $\dot{n}_a$ ,  $\dot{n}_{ue}$ , and  $\dot{n}_u$ ). The effect on the total number of people in urban areas is given by the number of persons seeking protected employment. If that number increases, migration from rural to urban areas may be induced as people come to cities in hopes of finding a job at higher real wages. Solution of the five conditions for the percentage change in numbers of people seeking protected employment gives

$$(4-2) \quad \dot{n}_u = \frac{1 + \beta_u}{1 - n_u/n_a\beta_a} \dot{w}_u.$$

A conclusion apparent from this result is that the existence of the higher wage may either raise or lower the total number of people who offer themselves for protected employment. The elasticity of demand for labor in protected employment,  $\beta_u$ , is negative, and if it is greater than one in absolute value, the sign of the effect of the rise in protected wage as given by the numerator of equation 4-2 is negative. With constant factor shares and with other product and factor prices unchanged, the absolute value of the elasticity of demand for labor will be the reciprocal of labor's share and will thus be greater than one, which ensures a negative effect. If product prices are raised, the absolute value of the elasticity will be even greater. With an absolute value of elasticity in the vicinity of one, which is not too unlikely, the effect on the number of people seeking protected employment is about zero. The maximum increase in the number of people seeking protected employment is obtained in the unlikely event that the elasticity of demand for protected labor is zero and the elasticity of demand for unprotected labor is infinite. In that case the coefficient in equation 4-2 reduces to one, an indication that the increase in the number of people seeking protected employment is proportional to the excess of the protected wage over the unprotected wage. Wages are likely to be raised substantially only in cases in which the amount of protected employment is rather small. Only in very unusual cases would there be a large effect on wages in protected employment if the protected employment were a sizable fraction of urban employment. Thus, even in the extreme case of zero elasticity of demand for labor, the effect of wage protection on the number of people in urban areas would be limited. It has been shown that because of labor turnover, the elasticity required to produce an increase in people seeking protected employment is even smaller than unity. The effect on the number of people in urban areas is further limited by the fact that in the usual case, in which the protected wage applies to only a fraction of the workers in urban areas, the people seeking protected employment come not only from rural areas but also from other employment in urban areas, and to the extent that they come

from other urban employment, there is no net effect on urbanization.

The unemployment induced by wage protection, as a percentage of total protected employment, is derived by subtracting the percentage change in protected employment from the percentage change in the number of people seeking protected employment. It is  $n_u/n_{ue}$  times the equation 4-2 coefficient plus the absolute value of the elasticity of demand for labor. Whether the unemployment is substantial depends on the magnitude of the wage increase and on the amount of protected employment. The unemployment, again, may be limited. To the extent that workers rotate or are in the same families, the unemployment does not affect the distribution of income.

Still more refined models might allow for workers to be doing something else while they are looking for jobs. It seems possible that the refinements would lead to even smaller estimated effects of wage protection on urbanization and unemployment. Reliable empirical evidence is difficult to obtain because of the difficulty of obtaining and interpreting unemployment statistics.

#### *Benefits from Government Expenditure and Tax Policies*

The benefits of such urban services as water supply, publicly provided health facilities, unemployment compensation, social security, and maintenance of low food prices may make cities attractive if, on net, the charges and taxes resulting from residence in urban areas are less than the benefits. The net benefits of urban residence that result from government actions can be expected to be reflected in a compensating wage differential, along with the other differences between urban and rural areas that go into determining urban-rural wage differentials.

The role of the tax structure in determining net benefits of urban residence deserves special analysis. Taxes imposed solely on the site value of land can be expected to be neutral with respect to location decisions, but taxes on land improvements or on capital structures, such as property taxes, may affect incentives to add to capital. If agricultural and urban property are taxed differently, the location of capital may be affected. Since the owner of nonhuman capital does not have to reside where the capital is employed to collect the income, a condition for location neutrality is that the marginal tax rate on nonhuman capital be the same in different locations in money terms. The owner of human capital (the person himself) must reside where the services are rendered. Thus, a condition for location neutrality for human capital is that marginal tax rates on human capital in different places be the same in real terms.

The urban wage tends to be lower than it would be without net government urban benefits by the percentage at which the net benefits are valued. This gives a first condition that may be expressed in percentage terms as  $\dot{w}_a = \dot{w}_u + \dot{g}$ , where  $\dot{g}$  is the percentage of the wage at which the government benefits are valued. That condition may be combined with three earlier conditions—that the total population is distributed either to rural or to urban areas ( $n_a \dot{n}_a + n_u \dot{n}_u = 0$ ) and that the changes in employment are governed by the elasticities of demand for nonurban labor ( $\dot{n}_a = \beta_a \dot{w}_a$ ) and for urban labor ( $\dot{n}_u = \beta_u \dot{w}_u$ ). Four conditions thus determine percentage changes in urban and nonurban wages and employment ( $\dot{w}_u$ ,  $\dot{w}_a$ ,  $\dot{n}_u$ , and  $\dot{n}_a$ ) as a function of the percentage of the wage at which government benefits are valued,  $\dot{g}$ .

Solution for the percentage increase in urban employment yields

$$(4-3) \quad \dot{n}_u = \frac{-\beta_u}{1 + n_u \beta_u / n_a \beta_a} \dot{g}.$$

Since the elasticity of demand for labor,  $\beta_u$ , is negative, the effect of government-supplied urban benefits is unambiguously positive. Furthermore, in contrast to the effects on urbanization (considered above) of environmental externalities and protected wages, it appears that government benefits could substantially affect the number of people in urban areas, at least on a one-time basis. For example, if the elasticity of demand in absolute terms is about 1 and if the proportions of people in rural and in urban areas are about the same (that is, the denominator is 2), the value of the coefficient in equation 4-3 will be about one-half. One reason the effect may be substantial is that it applies to all residents of urban areas, on the assumption that none of them can be effectively excluded from government-supplied urban benefits. If these benefits are valued at 10 percent of the wage, the number of people in urban areas will increase by 5 percent.

### Primal Cities

In the developed countries there are signs of decentralization not only to suburbs but to smaller towns and even rural areas. But developing countries are moving toward even greater urban concentrations, evidence not only of rapid overall urbanization but also of more pronounced tendencies than in developed countries to have a preponderance of population in one or a few cities.

The tendency to concentrate population in one or a few cities is as much a policy concern as urbanization as

such, and perhaps it is a greater concern. The reasons for, tendencies toward concentration are in part the same as those for urbanization in general; natural population growth is a prime reason. Still, the question remains why population should be so heavily concentrated in a few places. Perhaps the single most important reason for tendencies toward large cities is that it is economical to centralize production to avoid the costs of transporting goods back and forth. Only when these advantages are overcome by rising costs of various kinds within a city does it become economical to decentralize. This is a pervasive factor the world over, in developed and in developing countries.

A possible reason for the greater concentration in developing than in developed countries is that the effect of the urbanization externalities considered above may be greater. In particular, net benefits from government actions could act to promote concentration in the largest cities. It would appear that the question of urban bias (if it exists) should be rephrased to deal not with exploitation of rural by urban people but rather with the effects of urban bias on city size. The existence of net benefits in urban areas encourages people to move toward those benefits. That movement dissipates temporarily induced rural-urban differences in well-being but may have significant effects on city size. But although uninternalized externalities may contribute somewhat to larger city size, it does not appear possible to make a really strong case that they are the prime cause of pronounced large urban concentrations in developing countries. It seems necessary to look elsewhere for a full explanation.

Another hypothesis, which appears more likely, is that intercity transport, which would permit decentralization to smaller cities, is a relatively expensive commodity in developing countries. Provision of basic road facilities is highly capital-intensive, as is the operation of trucks and other transport equipment. If this hypothesis is valid, a gradual tendency toward less centralization may be observed as development becomes more complete and the price of human inputs rises in relation to the price of physical capital. This tendency toward a lowering of the relative price of intercity transport, however, may set in only at a late stage of development.

A final, related, possibility is put forth more tentatively. Infrastructure decisions may not provide as adequately as would be beneficial for development outside large urban centers. Development of more adequate evaluation tools that would allow more accurate assessments of infrastructure investments might lead to more investment outside large cities.

These hypotheses need to be tested as part of an investigation of the fundamental determinants of city

size, as represented by the study by Henderson (chapter 7).

### Economic Policy and Urbanization

Most of the reasons for rapid urbanization in developing countries are closely related to the economic development process itself. Wholesale reversal of urbanization is surely unthinkable and would amount effectively to a great interference with development, if not its virtual denial. The more relevant question becomes whether something marginal should be done to try to affect urbanization and in particular the concentration of urbanization in a few very large places.

#### *Externalities*

Three types of policies toward urbanization may be identified. The first is the internalization of externalities. If pollution, congestion, and related environmental externalities increase with city size, as suggested earlier, their internalization will lead to consideration of environmental costs in private decisions and will encourage decentralization. To affect decentralization, the policies must distinguish between the severity of environmental costs in different places. For example, if pollution abatement measures are tailored to local circumstances, greater controls will be imposed where control costs are smaller and damages are greater. In environmental policy as now practiced, however, uniform environmental standards are often imposed across all or a large part of a nation, so that control requirements are the same irrespective of city size, and thus incentives to decentralize are thwarted. If environmental controls are to play a role in decentralization policy, a less procrustean approach (which would also contribute to the efficiency of environmental policy) is needed.

The earlier analysis suggests that the externalities connected with high wages in protected employment may either increase or decrease urbanization. The most important consequences may be unemployment and overinvestment in search costs. It is not clear that wage policy and urbanization policy as such are as closely connected as might appear at first sight. At any rate, public choice considerations make it difficult to eliminate wage protection once it is instituted. A promising approach may be to reduce the effects of protection by keeping annual nominal rises in protected wages low in relation to the inflation rate.

The analysis suggests that the largest opportunities for internalizing externalities may be in the provision of public services, the tax structure, and the operation of

government transfer and price programs. With regard to public services, there would be no externalities if recipients of the services were charged marginal costs and if strong attempts were made to carry provision of services to the point at which marginal cost equaled marginal benefits. Quite possibly, neither of these two conditions can be fulfilled. In countries where public services are markedly more available in large urban areas than elsewhere and where the services are provided free or below cost, raising the charges or extending the services more widely throughout the country are possible actions. Extension of services may be the more realistic policy direction and may have distributional benefits.

#### *Infrastructure Decisions*

One method of providing infrastructure is to concentrate facilities at places where demand for infrastructure services is growing, by mechanically extrapolating recently observed trends in demand. The method is safe because it has a self-fulfilling element: provision of infrastructure services in an area encourages people to reside in that area. Demand then increases, and more infrastructure is built. Provision of infrastructure thus feeds on itself. For example, road networks that make the larger cities ever more powerful hubs of transport may be progressively reinforced in response to the signal of growing traffic brought on by previous investments and further contribute to the large size of the cities.

At the opposite extreme, and less frequently, infrastructure may be provided to stimulate areas of lagging development or low population. The degree of success may be disappointing because natural conditions and other locational characteristics are not overcome by the new infrastructure, with the result that the infrastructure is not used fully and the areas do not grow as much as was hoped.

Neither extreme appears optimal. Development of improved methods of evaluation could have important consequences for the spatial distribution of population. More attention could be given to the effects of infrastructure on the location of activity and to the quantitative estimation of benefits and costs that result from alternative patterns of development. In particular, evaluation techniques for transport facilities are sophisticated in some ways but need further development. For example, the travel savings method, which is widely used to evaluate the benefits of road investments, is to be recommended if the origin and destination of trips is not affected. More reliable methods, however, are needed to estimate benefits if changes in trips are brought about when road investments induce changes in the location of activities. Development of improved methods could

lead to identification of alternatives that are preferable to expansion in the largest cities; investments might be made, for example, in favorably situated medium-size cities.

### *Direct Interventions*

Nearly every country has policies that try to alter incentives to choose particular locations. Often such policies are explicitly concerned with preventing the concentration of the population in large cities. The policies take many forms, including land use controls and subsidies. With notable exceptions, such as the studies by Lee (chapter 8) and Reif (chapter 9), few serious attempts have been made to quantify the effects of the policies. Many policies appear to be largely nominal, without any real teeth, but in a few cases they may be overeffective.

Direct interventions may be decided on either because the best economic policies, connected with externalities or infrastructure, are not politically feasible or because deliberate actions to influence location, going beyond externalities and infrastructure, are mandated—desirably or undesirably—by political considerations. In either case, there are efficient ways to design the interventions. A Pigovian tax subsidy approach may be mentioned as a way to achieve efficiency in interventions. It should be noted, however, that Pigovian approaches—despite a century of advocacy—have seldom, if ever, been found to be amenable to practical application. Under a tax subsidy approach that is designed to encourage a coherent amount of decentralization, a shadow price for additions to population in different places, such as in cities of different sizes, would be established. Firms would face subsidies or taxes according to the amount of their employment in different places, and the tax or subsidy would be proportional to the difference in shadow price of having people in different places. Alternative forms of inducements (such as interest-rate and energy subsidies) and quantitative controls (such as land use restrictions) would be judged according to how well they approximated the outcome that would be attained under the shadow price approach. The same standard could be used to design and evaluate more general overall strategies for spatial development.

Deliberate spatial interventions as described above are controversial. Realistically, in the absence of badly needed objective quantifications of the effects of spatial intervention policies, it is difficult to be optimistic that effective or desirable interventions will occur soon. Efforts to design interventions rationally and even bring in cultural diversity and national defense as spatial considerations, along with more usual externalities, are to

be commended, but this achievement so far remains elusive.

Another kind of intervention is inevitable: the public management decisions required to provide urban services, regulate and facilitate land use, and provide for financing. Various facets of these problems are dealt with in the studies by Bahl and Linn (chapter 10), Ingram (chapter 11), Pachón (chapter 12), Thomas (chapter 13), and Keare (chapter 14).

Because public management decisions for cities go far beyond spatial considerations, they are at once easier and harder to deal with. A large body of ready-made knowledge and tools can be brought to bear on these more traditional subjects, but there is a worldwide lag in coordinating expertise and combining economics with engineering, planning, law, and other disciplines required to make public management decisions. Multiple goals also make public management decisions difficult. Income distribution considerations are difficult to deal with at best and may run counter to spatial objectives, to say nothing of more general economic development objectives. One of the hardest questions is how much effort to put into city management decisions. The mainsprings of development are not found in many of these decisions, but failure to deal with them adequately may put drags on development. Which ones will take care of themselves somehow and which ones call for special efforts are concerns for economic development.

Yet there is a way to improve urban management at present from a spatial point of view. It is quite practical to introduce quantification of the value of externality effects into the evaluation of projects. Orders of magnitude of the value of externalities at the margin can be developed along the lines illustrated in this chapter. These can be used as add-ons in social benefit-cost decisions concerning roads, water projects, and other expenditures that affect the location of activity. In this way, spatial externalities could be taken into account in the workings of government, without adopting explicit spatial policies. The required quantification would probably reveal the externalities to be small in most cases. Quantification would help keep externalities from being blown out of proportion by the emotional beliefs and political distribution motives that too often figure in discussions of externalities.

### **Bibliography**

- Brueckner, Jan K. 1981. "Labor Mobility and the Incidence of the Residential Property Tax." *Journal of Urban Economics*, vol. 10, no. 2 (September), pp. 173–82.

- Chappie, Mike, and Lester Lave. 1982. "The Health Effects of Air Pollution: A Reanalysis." *Journal of Urban Economics*, vol. 12, no. 3 (November), pp. 346-76.
- Cooley, Thomas F., and C. J. La Civita. 1982. "A Theory of Growth Controls." *Journal of Urban Economics*, vol. 12, no. 2 (September), pp. 129-45.
- Cropper, M. L. 1981. "The Value of Urban Amenities." *Journal of Regional Science*, vol. 21, pp. 359-74.
- Cropper, M. L., and A. S. Arriaga-Salinas. 1980. "Intercity Wage Differentials and the Value of Air Quality." *Journal of Urban Economics*, vol. 8, no. 2 (September), pp. 236-54.
- Diamond, Douglas B., Jr., and George S. Tolley, eds. 1982. *The Economics of Urban Amenities*. New York: Academic Press.
- Fare, Rolf, Shawna Groskopf, and Bong Joon Yoon. 1982. "A Theoretical and Empirical Analysis of the Highway Speed-Volume Relationship." *Journal of Urban Economics*, vol. 12, no. 1 (July), pp. 115-21.
- Fogarty, Michael S., and Gaspar Garofalo. 1980. "Urban Size and the Amenity Structure of Cities." *Journal of Urban Economics*, vol. 8, pp. 350-61.
- Graves, Philip E. 1980. "Migration and Climate." *Journal of Regional Science*, vol. 20, no. 3 (November), pp. 227-37.
- . 1983. "Migration with a Composite Amenity: The Role of Rents." *Journal of Regional Science*, vol. 23, pp. 541-46.
- Grieson, Ronald E., and Richard Arnott. 1981. "Optimal Fiscal Policy for a State or Local Government." *Journal of Urban Economics*, vol. 9, no. 1 (January), pp. 23-48.
- Harris, John R., and Michael P. Todaro. 1970. "Migration, Unemployment, and Development: A Two-Sector Analysis." *American Economic Review*, vol. 60 (March), pp. 126-42.
- Haurin, Donald R. 1981. "Local Income Taxation in an Urban Area." *Journal of Urban Economics*, vol. 10, no. 3 (November), pp. 323-37.
- . 1983. "The Effect of Property Taxes on Urban Areas." *Journal of Urban Economics*, vol. 7, no. 3 (May), pp. 384-96.
- . 1984. "The Effects of Output and Factor Subsidies or Taxes on an Urban Area." *Regional Science and Urban Economics*, vol. 14, pp. 533-46.
- Henderson, J. Vernon. 1977. *Economic Theory and the Cities*. New York: Academic Press.
- . 1982. "Evaluating Consumer Amenities and Interregional Welfare Differences." *Journal of Urban Economics*, vol. 11, no. 1 (January), pp. 32-59.
- . 1982. "The Impact of Government Policies on Urban Concentration." *Journal of Urban Economics*, vol. 12, no. 3 (November), pp. 280-303.
- . 1985. "Property Tax Incidence with a Public Sector." *Journal of Political Economy*, vol. 93 (August), pp. 648-65.
- . 1986. "Efficiency of Resource Usage and City Size." *Journal of Urban Economics*, vol. 19, no. 1 (January), pp. 47-70.
- Izraeli, Oded. 1977. "Differentials in Nominal Wages and Prices between Cities." *Urban Studies*, vol. 14 (October), pp. 275-90.
- . 1983. "Federal Individual Income Tax and Allocation of Labor between SMSAs." *Journal of Regional Science*, vol. 23, pp. 105-13.
- McConnell, Virginia D., and Mahlon Straszheim. 1982. "Auto Pollution and Congestion in an Urban Model: An Analysis of Alternative Strategies." *Journal of Urban Economics*, vol. 11, no. 1 (January), pp. 11-31.
- McDonald, John F. 1983. "An Economic Analysis of Local Inducements for Business." *Journal of Urban Economics*, vol. 13, no. 3 (May), pp. 322-36.
- Porell, Frank W. 1982. "Intermetropolitan Migration and Quality of Life." *Journal of Regional Science*, vol. 22, pp. 137-58.
- Power, Thomas M. 1981. "Urban Size (Dis)amenities Revisited." *Journal of Urban Economics*, vol. 9, no. 1 (January), pp. 85-89.
- Rosen, Sherwin. 1979. "Wage-Based Indexes of Urban Quality of Life." In P. Mieszkowski and M. Straszheim, eds., *Current Issues in Urban Economics*. Baltimore, Md.: Johns Hopkins University Press.
- Tolley, George S. 1974. "The Welfare Economics of City Bigness." *Journal of Urban Economics*, vol. 1, no. 3 (May), pp. 324-45.
- Tolley, George S., Philip E. Graves, and John L. Gardner, eds. 1979. *Urban Growth Policy in a Market Economy*. New York: Academic Press.

## *Urban Development Policies in Developing Countries*

Bertrand Renaud

The recent world recession—some might even call it a depression—is testing the vitality of urban policies in developing countries. One is reminded of Dickens's *Tale of Two Cities*: "It was the best of times, it was the worst of times." On the one hand, the pressures to move the economy again, to confront the problems of sluggish exports, and to lower high unemployment rates, together with the lack of government resources for state and local investments, make it the worst of times for urban development policies. On the other hand, it is generally in periods of strain or crisis that the significance of the urban sector for the national economy is most truly appreciated. Austerity is forcing people to rethink the fundamental issues of urban development and the types of policies that can be most fruitfully pursued. This may be the best of times to review policy priorities and to look for robust solutions to urban problems.

Before the prospects for urban development policies are outlined, some questions must be answered. What are the characteristics of policies toward urbanization and the policy tools that have been used in developing countries? What rationales are given for those policies? What do we know about the effects and effectiveness of policies? Without answers to such questions it is not possible to examine properly the future direction of urban development policies.

### Basic Trends in Urban Policy

The forces that shape urbanization and urban policies in developing countries can be summarized in three broad propositions. First, the wave of urbanization that

has been sweeping through developing countries since World War II will be sustained through the end of this century. Second, there are significant quantitative differences between contemporary urbanization in developing countries and the historical patterns observed for today's advanced economies. Third, the policy problems encountered in developed countries have clearly been diverging from those of developing countries. Simple extrapolations of past trends in developed countries will be of limited value for the formulation of urban policies for developing countries.

### *The Main Characteristics of Urbanization in Developing Countries*

There were approximately 800 million people in 1750, about 1.3 billion by 1850, 2.5 billion by 1950, and over 3.6 billion in 1970; there will be between 6.2 and 6.5 billion people by 2000. Most of this population growth is taking place in developing countries. For instance, the 1982 census shows that the population of China is more than 1 billion and is growing by 17 million a year. These demographic changes have been associated with a rapid redistribution of population from rural to urban areas and a significant shift in the proportion of the urban population that lives in large or very large cities. Before 1850 no country was predominantly urban, and in 1900 only the British population was more than 50 percent urban. By 1920, 14 percent of world population was urban. Now that proportion is close to 40 percent, and it is anticipated that by 2000 half of the world's population will be urban. The urban areas of developing countries absorbed about 600 million people between 1950 and 1980. During the final quarter of this century there are

expected to be 1 billion people in developing-country cities, which is 64 percent more than the total population of about 610 million in the advanced market economies today.

Demographic and economic forces render contemporary urbanization different from the earlier experience of advanced economies. First, population in developing countries is expanding rapidly. A country that grows at the rate of 2.4 percent a year, the current group average for middle-income countries, will experience an 81 percent increase in population in twenty-five years. Second, the middle-income countries have experienced high economic growth rates so far. After twenty-five years an economy that is growing at 5.5 percent a year—the GDP-weighted average growth rate for the middle-income countries for 1970–79—will be 3.8 times larger than it is today. Rapid demographic and economic growth contributes to average growth of 3.8 percent a year in the urban population; in many countries the growth rate is much higher. Institutionally, the role of government is more important in developing countries now than it was for advanced economies when they were at similar levels of urbanization. The slowdown of the world economy in the 1980s will reduce the pace of urbanization in developing countries. Nevertheless, the historical contrast remains sharp between the slow urbanization at relatively high levels of income that advanced economies experienced in the past and the rapid urban population concentration at much lower levels of income in developing countries today.

World urban trends can be described further in a series of propositions.

- Urban areas will play an increasing role in absorbing large shares of the world's population. There will be a marked increase in the level of urbanization over the next twenty years. An increasingly large number of countries will have become more urban than rural. (Already, more countries are predominantly urban than are predominantly rural.)

- The distribution of urban population over broad regions of the world has changed dramatically. At the beginning of this century the largest share of the world's population living in cities of over 100,000 was found in Europe. By 1950 Europe was already behind Asia and America, and by 2000 Asia will dominate the world urban picture, with about 45 percent of the world's urban population. The level of urbanization will rise most rapidly in Africa.

- There is a trend toward concentration of urban population in large cities (over 1 million), very large cities (over 2 million), and supercities (over 5 million). In 1950 the world's urban population was 393 million, and 45.1 percent of that number was in cities of over 1

million. In 1975 the corresponding figures were 983 million and 51.4 percent. For 2000 some projections yield 2.2 billion and 63.0 percent.

- The striking growth of urban areas should not distract attention from two facts: many countries have very large rural populations, and a very large proportion of the urban population still lives in towns of less than 100,000 population. Because very large cities dominate the urban systems, there is a great difference between the statistical average size of settlements and the typical size of settlements. There are very large urban places, but the typical place is still rural oriented.

- Even though developing countries are urbanizing rapidly, only a few have declining rural populations. This is because of their high rate of population expansion—the average crude growth rate is 2.4 percent a year, about five times that of industrial countries, which register an average 0.4 percent a year. The reservoir of potential rural-to-urban migrants continues to expand.

The evidence suggests that the policy context in developing countries differs markedly from the better-known historical pattern of Western countries. In industrial countries urbanization took many decades and occurred relatively slowly in comparison with the urban transformation now occurring in developing countries. In today's developing countries the pace is more rapid, population growth is higher, income levels are lower, and the opportunities to relieve domestic population pressures through migration are limited.

### *Diverging Urban Trends in Developing and Developed Countries*

Since the early 1970s urban trends have been moving in divergent directions for developed and developing countries. It is worthwhile to briefly compare these tendencies because of the influence that urban policies and analyses in developed economies have had so far on the formulation of urban policies in developing countries. The contrast between the two groups of countries may facilitate the discussion of urban policies in developing countries.

The difference between urban conditions in developing countries and in advanced economies is most easily brought out by reducing the complex socioeconomic processes of urbanization to a single demographic indicator, the percentage of total population that is urban. Long-term urbanization, which involves an increase in this percentage, can then be seen as the process of moving along a logistic curve, with the urban share of total population rising from about 1 or 2 percent to a saturation level somewhat less than 100 percent. The dynamics of urbanization in a given country can be

characterized by the slope of the urban logistic curve, which reflects the tempo of urbanization, and by the urban saturation level, defined as the highest percentage of urban population that is likely to be urbanized at very high levels of economic development.<sup>1</sup> In comparing national urbanization strategies, much confusion could be avoided by referring to this simple and familiar curve as a quick way of differentiating between countries, since the level of urbanization and the level of economic and institutional development are closely correlated.<sup>2</sup>

Until the 1960s advanced and developing countries alike were seen as moving along this urbanization curve, with the first group preceding the second. In the 1970s two major structural changes became apparent in developed countries, and their impact on urban policies is becoming pronounced. The first structural change is the end of traditional urbanization: the national populations of advanced economies have become fully urbanized, rural-to-urban migration has slowed to a trickle, fertility rates in many countries are falling below the replacement level, and the populations of some countries are even declining. The steadying effect of internal migration on the growth of most cities is now missing in many advanced countries. In addition, national governments are trying to limit international migration flows. Urban systems such as those of the United Kingdom and the Federal Republic of Germany have stopped expanding. Demographic urban growth often takes on the appearance of a zero-sum game in which the demographic growth of one city comes more and more at the expense of another city, whether it is defined as a central city, suburban center, metropolitan area, or medium-size or small town. For the first time advanced countries are discovering what it means to have reached the urban saturation level at the end of the logistic curve.

The other major structural change is economic rather than demographic. It is marked by the transition from an extended period of rapid and sustained growth during the 1950s and 1960s to the present period of slow and uncertain growth and is tied to the emergence of a world market over the past thirty years, the accompanying shifts in the international division of labor, and, in particular, the impact of these shifts on the demand for and production of standardized manufactured goods. During the 1970s the triple impact of the slowdown in economic growth, structural change in the demand for manufactured goods, and shifting terms of trade with the oil-producing countries accelerated the pace of change of urban policies in advanced economies, partly but not exclusively because of the concentration of old-line manufacturing activities in certain regions and cities.

A stylized way to describe this economic structural change is to say that we are witnessing the beginning of

the third industrial revolution in the advanced economies. The first industrial revolution, which started everything, was based on coal, the steam engine, railroads, and textiles. The second industrial revolution was based on petroleum products, chemicals, the automobile, and electrical and mechanical industries. The current revolution is based on new energies, electronics, information industries, bioengineering, and services. The first two revolutions, especially the first one, strongly favored urban concentration; the third one does not. According to this interpretation one could hypothesize that the 1980s are not the beginning of a period of permanent slow growth but rather a transition period toward new forms of growth. The policy implications of these two structural changes for urban policies in developed countries are becoming clear.

Until the late 1960s the policy paradigm in advanced economies was based on urban growth under these conditions: a moderately rising national population; continuing transfers of population from the countryside (associated with the development of manufacturing and service employment); a differential growth that favored larger cities and the formation of city regions; an increased demand for space per capita (associated with significant gains in per capita income); and a long-term concentration of population in favored economic regions.

Confronted with urban growth, advanced countries developed policies incrementally and responded to problems as they emerged. At first, only limited attention was paid to the context within which urban issues were debated and evaluated. Then rapid urban growth, combined with abundant fiscal resources, led to an increase in the scope of objectives, in the breadth of the population to be served, and in the number and variety of instruments to be tried. A major shift occurred, from an early and almost exclusive consideration of physical planning problems to a more comprehensive focus on the socioeconomic problems of cities. Increasing emphasis on income redistribution in an urban context was accompanied by a concern for economic efficiency in solving problems. The rapid growth of urban programs frequently led to problems of coordination and bureaucratic congestion. During this period parallels were drawn between urban policy issues in developed and developing countries.

The factors that now differentiate urban policy issues in advanced economies from those of developing countries are many: the stagnation or decline of the largest metropolitan regions in developed countries, greater inequalities within large urban areas than between regions in developed countries, a shift of manufacturing activities in advanced economies from the production of standardized goods toward the production of goods with

a high proportion of services input, and the dominant role of multilocal, multifunctional corporations that do not simply react to their spatial and economic environment but attempt to shape it.

In this new context of stabilizing urban systems and fiscal retrenchment, urban policies and instruments in developed countries will differ considerably from those of developing countries. In a surprisingly large number of Western countries the policy paradigm is shifting from urban growth to an emerging urban zero-sum game. Because of the nationwide impact of industrial restructuring, which affects social groups selectively, the relative balance between social (people-oriented) policies and urban (place-oriented) policies is shifting in favor of the former in developed countries. New urban policies for developed countries must now be defined against four major issues:

1. How to improve the economic base of cities, many of which confront major problems of reconversion from an old manufacturing base to services and high-technology activities
2. How to improve social conditions, since industrial restructuring is leading to sharp income inequalities within cities.
3. How to improve urban services to accommodate the changing size and composition of resident populations while meeting major maintenance needs
4. How to allocate resources and responsibilities among levels of governments, given the heterogeneity of local conditions among cities.

To manage simultaneous urban growth and decline, the emphasis should be on selectivity, flexibility, smaller-scale interventions, greater diversity of actions, and increasing reliance on local government to coordinate activities at the city or neighborhood level. For the advanced economies there also remains the problem of declining cities, underutilized infrastructure, and the role of the central government in influencing such trends.

### Urbanization Policies

The rapidly shifting patterns of population distribution and economic activity in most developing countries create inequalities in economic growth rates, industrial structure, employment conditions, household incomes, wages, and levels of services which are keenly felt by decisionmakers and social groups. These inequalities, which include those between the rural and urban sectors, regions, cities of different sizes, and social groups

within large cities, have prompted policymakers to experiment with spatial considerations in their national economic policy schemes. For example, they have become more concerned with the composition of investment projects and their differentiated impacts on the growth of regions and cities.

### Objectives

In many countries, particularly middle-income countries, the decentralization of economic activity from large, congested capital regions and the closing of the welfare gap between rich and poor regions are the stated policy objectives. The justifications for initiating urban decentralization policies generally given by developing-country governments are, first, the necessity of reducing the severe managerial problems and the economic and social costs associated with congestion, pollution, and the difficulty of providing adequate services in large cities; second, the need to reduce wide regional income disparities within the country; and third, the belief that more investment outside the largest urban centers in favor of the poorer outlying regions will foster national growth. Thus, urban decentralization policies are defended as both equitable and economically efficient. The national urbanization policies of many developing countries are based on the premise that it is possible and right to decentralize population away from the largest cities.

A national population decentralization strategy actually covers many objectives. This multiplicity of objectives calls into use a comparable multiplicity of instruments, and the problem becomes one of identifying the combination of policies most appropriate for a given country at a given level of development.

Some of the more explicit objectives encountered in developing countries are:

1. *The integration of peripheral regions to increase the size of national domestic markets and the intensity of regional demand.* Raising regional incomes is greatly dependent on rural development, agricultural policies, and the development of transport and communication networks.

2. *The integration of peripheral regions and the opening up of new resources to raise national output.* This objective applies to mineral-rich countries as well as to countries which have had an important resource frontier to exploit, such as Malaysia and Brazil.

3. *The reduction of interregional disparities.* In a well-integrated society, concern about inequality among individuals and households should be more important than concern about inequality between places. The reduction of regional disparities is, however, a legitimate concern in developing countries. In Latin America, for

instance, the output per capita of one region may be as much as ten times the output per capita of another, while in fully integrated advanced economies the apparent differences are on the order of three or two to one.

4. *The improvement of national political integration and social cohesion within the nation.* It would be a conspicuous mistake to assume that national policies have purely economic objectives such as increasing output or redistributing economic opportunities. In heterogeneous societies marked by important cultural, political, and linguistic differences, a dominant element of a national urbanization strategy is to maintain the cohesion of the state and to prevent regional minorities from, in some fashion, leaving the national coalition of social groups. Special institutions such as regional development corporations are often used as an expression of the central government's concern for the region.

5. *The rapid development of border regions for reasons of national security.* This objective has been important in Latin America, where disputes over the exact location of national boundaries in underdeveloped regions have been frequent in the past. Similar situations also exist in Asia and the Middle East.

6. *The improvement of the national system of cities.* This is a more recently stated objective of national policies in middle-income countries. The city system has a dominant role in the transmission of economic impulses and the diffusion of social and economic innovations over the national territory. Raising the level of economic activity in medium-size cities will help them attain economies of scale and will improve the interregional diffusion of economic growth. It is also expected that a more developed system of cities will help in equalizing access to important services such as education and public health which can increase the human capital of a country.

### *Some Questions*

The above objectives are still too broadly defined. In particular, the objectives of decentralizing population away from the largest urban centers and narrowing interregional differences beg many questions. Can an urban strategy operate independently of other national, social, and economic strategies? Is decentralization really desirable from the viewpoint of economic efficiency? Is there a tradeoff or a convergence between higher rates of national growth and greater regional equality? Is it possible to channel economic activities to preselected cities? Under what conditions? At what cost?

### *Urban Problems and Policies*

Urban policies in developing countries have been formulated and tried in the same incremental and un-

structured way that has typified the experience of advanced countries during their decades of rapid urban growth. In fact, they have been derived from experiments in advanced countries.

These policies stress the decentralization of economic activity and population as a means of relieving congestion and solving the environmental problems of large cities. Common objectives are to create employment and service centers, to build new towns, and to implement large investment programs in transport and infrastructure. Technical interest in the design of new facilities and new urban technologies, however, often overtakes the understanding of the effective demand for services by urban groups. This supply-side, design-dominated approach to urban development is only slowly being corrected through a better understanding of the engineering, economic, and institutional constraints on urban development at low levels of income.

At the metropolitan level the overwhelming preoccupation is with the absorption of a large population expansion every year. The supercities with populations above 5 million may no longer be the fastest growing, and the current economic slowdown may marginally reduce their growth rates. Still, they are likely to grow at rates between 3 and 6 percent, or by 150,000 to 300,000 a year. Most metropolitan governments have great difficulty in developing and implementing well-structured investment programs to confront such problems. Urban areas are outgrowing their governments' capabilities. Local governments have inadequate authority to coordinate activities beyond their own boundaries, and their political power and resources are inadequate for project implementation. Interventions by local governments need the support of central governments. New local tax bases must be created and existing ones redesigned. Increasing reliance on transfers from central governments seems desirable but is not necessarily feasible. The issues are familiar and can be listed briefly:

- How to improve the economic base of the city and provide productive employment to a rapidly expanding labor force that often has low skill and educational levels? How to influence employment location to ease the growth of the city?
- How to provide transport that is affordable for the majority of the population? How to ensure a dynamic balance between the public and the private sectors when a 10 percent minority of car owners can and does impose severe congestion costs on the other 90 percent of daily commuters by overwhelming an inadequate infrastructure?
- How to develop land use patterns that generate a better balance between home and workplace? How to recapture some of the increment in land values gener-

ated by urban growth to finance city infrastructure investments?

- How to improve housing supply when 90 percent of the housing stock is provided by the private sector without the necessary urban infrastructure which would make new neighborhoods fully functional? Put another way, how to take advantage of low-income progressive housing investment patterns to develop affordable but efficient and healthy neighborhoods?

- How to mobilize financial resources to meet the massive infrastructure needs of cities for road networks, transport systems, water and sanitation systems, education, local health services, safety, and fire protection?

- How to provide assistance to the poorest residents, who are most in need of help and least able to get it?

As noted earlier, rapid demographic growth at low levels of income is characteristic of urban development in developing countries. Because of severe resource constraints, resource mobilization is as important as the distributive aspects of urban growth: resources must be found before they can be allocated equitably. In this new period of severe capital scarcity, it is clear that to receive support from central government planners, urban investments will have to contribute demonstrably to increasing the productivity of cities. This is a new challenge for urban analysts. How can it be shown that a certain urban investment package raises the productivity of a city? How can such a package be defined? What are the analytical methods available for establishing this link? Can they be implemented easily and routinely?

### *Three Dimensions of Urbanization Policies*

Urbanization policies can be viewed along three dimensions. First are national economic policies which have unintended effects on the urban development of a country. Second are regional policies that operate at the national level and aim at an equitable and efficient allocation of population and resources among regions. Third are policies for the internal management of cities.

The *unintended spatial biases of national economic policies* in favor of some urban centers are commonly products of trade policies which protect the manufacturing sector. For example, policies regarding credit allocation, public investment, and prices give preferential treatment to economic activities which are concentrated in a few cities and regions. The management practices of the central government and its regulation of economic activities require location of these activities close to the capital and contribute to the urban vortex.

It is not necessary to make the unrealistic claim that spatial considerations should prevail over economic growth considerations in the formulation of national economic policies. But there are several reasons why a

good understanding of the unintended spatial effects of national economic policies—of implicit spatial policies—is essential for developing countries. First, these policies have a definite impact on where people live and work. Second, there is a growing awareness and specific evidence that the effects of implicit incentives on business location decisions are much stronger than the publicized explicit incentives that favor decentralized location. Third, if national economic policies were adjusted so that their spatial bias in favor of the most advanced regions were moderated, reliance on the invisible hand of the market would reduce the need for specialized teams of spatial planners and avoid the waste of scarce managerial and administrative talent on the fine-tuning of regional decentralization incentives of doubtful effectiveness or efficiency. The likelihood of such an adjustment is very remote, however, because coordination of the execution of better-designed policies presents great problems. Here we are mostly interested in understanding the urban impacts of national policies and the dynamics of urban development in the manner defined by Kelley and Williamson in chapter 3.

*Regional policies* can strengthen promising secondary urban centers through such actions as better investment and management policies for transport, industrial estates policies, and more important, the systematic development of organized informational networks between these cities and the capital region, such as banking networks, industrial association networks, and better administrative structures. As is now well established, growth center strategies are more a way of thinking about such policies than a precise methodology.

The *appropriate internal management of cities* is important to the success of national spatial policies. In the case of very large cities, policies to limit or stop population growth are not substitutes for policies that directly address the correction of congestion and pollution and the provision of adequate services. If other cities are not efficiently and effectively managed, their chances of attracting industries and migrants from the largest urban centers will be small.

### *Analytical Issues*

Progress in the formulation of urbanization policy is possible by integrating more closely the three dimensions of policy discussed above. Research findings, however, lag far behind the needs of policymakers. Designing workable programs to moderate urban concentration, to narrow the gap between rich and poor regions, and to address more effectively the needs of lower-income groups strains our knowledge of national urban development processes.

Much urbanization policy in developing countries and elsewhere is based on presumptions rather than on established findings. For instance, there is always much interest in coordinated regional schemes in developing countries, but the actual motivation is the perception that trickling-down does not work fast enough to reach poor regions and, particularly, poor people. It is argued that, on equity grounds alone, it is necessary to increase the level of investment in peripheral regions, or at least in nonmetropolitan areas within these regions. In addition, it is felt that both equity and economic efficiency objectives can be met jointly by investing in peripheral regions because such investment opens up new resources to the national economy.

The justification of national spatial policies on equity grounds is not a problem. It is accepted that public policy should help achieve a certain minimum standard of living in every region of a country. The minimum standard of living could be defined on the basis of the market demand curve of a given percentile of the national population, say the twentieth percentile, for important goods and services such as housing, medical care, nutrition, and public education. For instance, in figure 5-1 the private demand curve for nutrition leads to an equilibrium level of consumption equal to  $OQ$  in a peripheral region. National intervention is considered justified to bring that level up to at least  $OR$  for all social groups. The net social benefit of a transfer policy will be a triangle  $MNT$  which is the difference between  $MNST$ , the public good externality generated by providing everyone a minimum standard of consumption up to  $OR$ , and  $NST$ , which is the usual measure of the efficiency cost of providing a subsidy of up to  $ST$  to the target group. Exactly the same analysis is repeated in a more general form in figure 5-2, in which the shaded area  $MNSU$  represents the social gain from the externality created by the regional transfer policy and the triangle  $NST$  represents the gain from the externality which

Figure 5-1. The Private Demand Curve for Nutrition

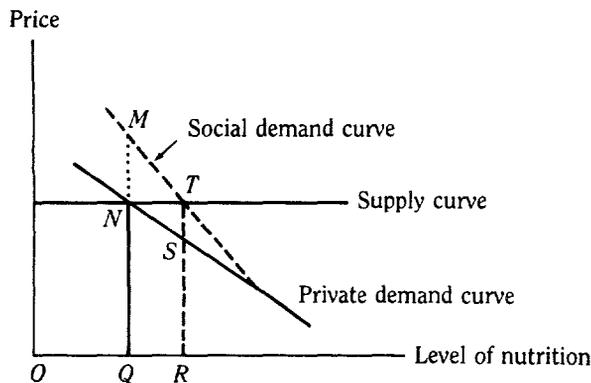
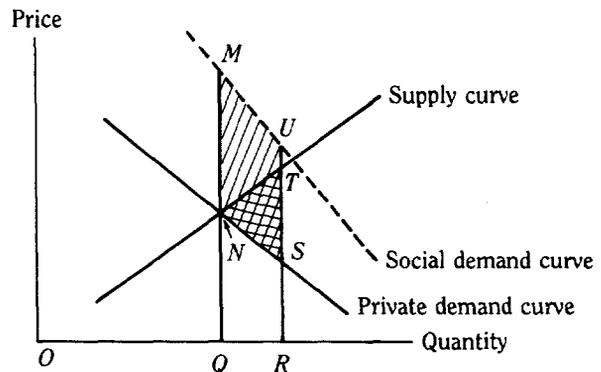


Figure 5-2. The Equity Rationale of National Spatial Policies



is offset by the loss in economic efficiency. Inherent in such spatial redistributive objectives is a certain amount of efficiency loss, which is accepted by society on equity grounds.

A major research task related to urbanization policy is the development of a methodology to determine whether the increasing concentration of population and economic development in a limited number of very large cities in developing countries is a result of faulty policies and market failure, or whether rapid urban concentration is an inevitable part of the development process. We need to know whether decentralization policies can be justified on economic efficiency grounds, or whether they are futile and economically inefficient.

### The Direction of Urbanization Policies

Rapid urban growth will continue to shape policies in developing countries in the coming years. The two traditionally dominant urban issues—decentralization away from large urban centers and reduction of inequalities between regions—will continue to shape the agenda. It is reasonable, however, to expect a more knowledgeable selection of urban development objectives and a greater degree of attention to the choice of instruments to achieve more realistic goals than in the past. Over the past decade the gap between the rhetoric of urbanization policies intended to prevent urban migration and the actual urban growth patterns (which include rapid concentration in major cities) has become painfully clear. Everyone hopes not to hear again the old question of whether urban growth can be stopped and would rather be asked what kind of urban growth should be encouraged. Misleading notions of optimum geography and optimal size distribution of cities based solely on population should not be used to select population

growth targets for which no feasible program can be found.

Significant progress has been made in understanding the impact of the spatial distribution of urban growth as a result of growth policies, trade regimes, economic regulations, the influence of federal systems of government, and fiscal policies. International comparisons also show how ill-adapted earlier policy advice has often been to the structure of existing institutions within a country. The World Bank, through its activities in about sixty countries, has gained substantial insights into the dynamics of urban growth in developing countries and into the actual capacity of national and local institutions to induce change. In addition, there is better understanding of the time dimensions of urbanization policy. A sobering experience for policy analysts with limited or no operational experience has been the sudden awareness that most urban projects take an average of eight years from original planning to full-scale operation. Thus, urban project cycles are considerably longer than the political cycles in many countries. The momentum behind new urbanization policies depends on market forces and the parallel growth of specialized urban institutions that are less sensitive to short-term politics. There are major differences in urbanization policies around the world, and it is to these policies that we now turn.

#### *Urbanization Policies in Sub-Saharan Africa*

The severity and complexity of the problems facing many countries of Sub-Saharan Africa in the 1980s will greatly constrain urbanization policies in the region. Because it is starting from the lowest urban base, Sub-Saharan Africa is the region that is undergoing the greatest change, and its postindependence period has seen spectacular urban growth. The average growth of urban population has been 6 percent a year—8.5 percent for thirty-five major capitals that are doubling their populations every nine years. There are now twenty-eight African cities with populations of over 500,000; twenty years ago there were only three (World Bank 1981). Because of this rapid growth, cities which were in good physical and financial shape two decades ago now face enormous difficulties.

The policy problems that are besetting African countries in the postindependence period are severe, and the same problems can be identified almost everywhere.<sup>3</sup> There has been a conspicuous failure to predict and plan for rapid urban growth. Most city governments have been unable to grasp the implications of a population that doubles every nine years. Overcentralization has compounded the problems, as many new central institutions have assumed or claimed more responsibility than

they could hope to handle. The consequence has been hypertrophy at the center and atrophy at the local level. There has been significant deterioration of services, the two most common examples being national housing corporations and water systems which are managed directly by ministries. Overcentralization has also affected the financial health of cities, where a common error in English-speaking countries has been the assumption by the central government of revenue bases which had historically been available to cities. In addition, the financial responsibilities of cities, particularly in education and health, have increased drastically, but new revenue bases have not been established.

There has also been a systematic failure to evaluate urban income distributions and to plan for them. Before independence most cities had high employment rates and few poor families, since urban migration was conditional on the availability of a job. In the postindependence period urban policies failed to plan for services that the incoming migrants could afford. Thus, achieving realistic living standards has been a pervasive problem. New urban migrants do not have a realistic sense of what is economically possible, and there has been a strong political demand for high standards of services that are unattainable at present income levels. The typical result has been high-quality services for a few and no services for most. This in turn has led to serious physical problems of unserved and sometimes unserviceable squatter settlements, as well as overcrowding in planned and unplanned areas.

This extreme polarity in income distribution has been maintained through the growth of public sector employment, which inspired the formulation of the Harris-Todaro model of rural-to-urban migration. Whereas formal-sector employment accounts for 6 to 15 percent of the working population, public-sector employment increased during the postindependence period and now constitutes 40 to 75 percent of salaried employment.<sup>4</sup>

Not surprisingly, political problems have arisen in the day-to-day financing of cities. In most countries it has been difficult to levy adequate charges for the sustained development of such urban services as water supply, waste disposal, electricity, and road maintenance. Inadequate charges have led to lagging and deteriorating services and have thereby increased the resistance to paying for them. Current flat or negative growth rates of gross national product, coupled with a projected annual population growth of 3 percent, are having a sobering impact on urbanization policies in Africa.

Nigeria, the largest country in the region, well illustrates present problems. Its position in 1973 as a poor overpopulated country with a per capita income of \$150 changed dramatically as oil revenues brought in an estimated \$100 billion in foreign exchange earnings

during 1973–81. But despite substantial petroleum revenues, exports, and a reported per capita GDP of \$670 in 1981—as against the regional average of \$410—Nigeria still ranks among the world's least developed countries. Its urbanization has been greatly accelerated for the worse by the oil revenues. Nigeria also provides a clear example of the dominant impact of trade and industrial policies on urban growth and concentration. An analysis by Bertrand and Robertson (1978) used the relatively standard estimation of the sectoral net subsidies provided by trade and industrial incentives to each industry and estimated the spatial distribution of these subsidies for nineteen states in Nigeria. They found that as a result of the spatial distribution of economic activities 89 percent of total net subsidies granted to industries benefited Lagos. The magnitude of these industrial incentives overwhelmed those proposed by a contemporaneous official report on industrial dispersal, which suggested that industrial decentralization be pursued to promote social equity, to avoid congestion at the center, to avoid inflationary pressures on products and wages because of urban congestion, to improve regional employment opportunities, and to foster political cohesion. Budgetary resources allocated for these explicit decentralization objectives could not be precisely calculated, but it was estimated that they were less than a tenth of the estimated industrial subsidies that benefited Lagos. Concurrently the agricultural sector, which has great potential, performed poorly as a consequence of inappropriate government policies, particularly unfavorable domestic terms of trade. Agricultural exports declined substantially, and the output of food crops, except rice, became inadequate; food imports increased tenfold, from \$200 million in 1973 to \$2 billion in 1980. The instability of oil revenues has accentuated the difficulties of managing the economy and developing urban policies.<sup>5</sup>

The case of Nigeria well illustrates the dominance of national economic policies over the direction and location of urbanization in Africa. Remarkably, however, the concentration of new industries in a few urban centers is making national economic planners, who are traditionally indifferent to urban policy, more aware of the crucial role of effective management. Even in this period of economic slowdown and scarce public resources, there is greater support for effective urban management than there was a decade ago. Sporadic breakdowns in urban public transport, which prevent workers from reaching manufacturing plants in the capitals of smaller African countries, also have had a clear demonstration effect.

Regarding the future of urbanization policies in Sub-Saharan Africa, the priority should be on laying solid

foundations for effective management in large and medium-size cities. Given the scarcity of central resources, a high-priority item is the development of a local tax base through the implementation of a cadastre and an effective system of land administration. Policies less biased in favor of the largest urban centers—in particular, agricultural price policies—are more likely to be implemented now than in the past two decades. It is less clear whether trade and industrial policies will be modified in a way that will favor secondary urban centers. Given the generally weak institutional structures, the means of stimulating the economic base of the intermediate cities are even less apparent than is typically the case in higher-income countries.

### *Changing Urbanization Policy in India and China*

India and China have the two largest urban populations in the world, but by the end of this century they will be the only important countries with less than 50 percent of their populations in urban areas. In spite of their low per capita income levels they differ significantly from African countries in that they have extensive and resilient institutional structures.

Until now there has been a strong bias in India against the urban sector. In few other countries could one find a more firmly held view that activities to improve urban areas are by their nature antirural. Recently, however, there have been indications that this view of urbanization as an urban-rural zero-sum game has been giving way to greater understanding of the need for well-structured urban policies. The results of the 1981 census are contributing to this shift in opinion. Contrary to expectations, the total population growth rate did not decline during 1971–81; in fact, it increased by 24.75 percent. The rate of urbanization continues to accelerate. That rate grew at 2.58 percent a year during 1951–61, at 3.79 percent during 1961–71, and at 4.60 percent during 1971–81. The urban sector is expected to have 300 million people in the year 2000. At the same time, more than 50 percent of GNP is now produced in urban areas, even though close to 70 percent of the labor force remains in agriculture.

Given the federal structure of India and the size of its states, problems with urban policies are discussed more effectively at the state level. In the three most urbanized states of Maharashtra (Bombay), Gujarat (Ahmadabad), and Tamil Nadu (Madras), as well as in West Bengal (Calcutta), the management of large cities is being improved. The greatest progress has been in the acceptance that quality standards in the residential sector have to be tailored to levels of household income. But the strong

Indian preference for regulations and for control of the entire economy over the use of incentives to achieve desired social objectives is still felt. Bombay's costly plan to freeze city growth and to shift development to New Bombay is only partially helped by spontaneous suburbanization. In every state the desire to relocate industry in small towns and to stop its spontaneous growth in larger urban centers contributes to expanding the underground economy that is stimulated by conflicts between licensing and the needs of business firms. As in most countries, India's coordination of urban policies is complicated by the split between the secretaries of commerce and industry, who regulate industrial incentives and location, and the secretaries of public works and housing, who better understand urbanization. Serious problems must be solved regarding the internal management of cities in India, with special concentration on the land use legislation of 1976, the long-standing rent control in the housing sector, and the improvement of the fiscal strength of the cities.

Urbanization policies in China differ considerably from those in India. By substituting detailed regulations for market mechanisms, China has so far bypassed the problem of balancing explicit decentralization incentives against the implicit incentives of trade and growth policies. In addition, China appears to have succeeded more than most developing countries with similar income levels in improving the internal management of its cities.

According to the 1982 census, China's population has passed the 1 billion mark.<sup>6</sup> The annual rate of increase is 2.1 percent, and total population has increased by 313.6 million since 1964. The population in cities and towns grew at an annual rate of 2.7 percent, implying that 206.6 million people are now living in cities and towns, or 20.6 percent of the total population, compared with 18.4 percent in 1964. Since the first plan (1953-57), the policy has been to suppress the growth of large cities and to relocate industry from the coast to the interior. Strict control over migration through work permits and control over housing, as well as forced out-migration of young workers, have prevented the growth of cities, particularly the largest ones. For instance, the population in Shanghai province grew only from 10.6 million to 11.8 million people between 1960 and 1982. The five-year plan for 1980-85, which projected the expansion of both internal and external trade, is likely to have speeded up the growth of the urban sector, but there remains a strong objection to the growth of the largest cities.

While the normal growth of cities through migration and natural increase has been controlled, considerable efforts have been made to improve the internal manage-

ment of cities. Given present income levels, municipalities provide high-quality urban services in garbage collection, street cleaning and maintenance, water supply, and waste disposal. Three areas which require particular attention in the largest cities are public transport and traffic management (with the growth of motorization), environmental pollution, and the upgrading and expansion of the housing stock. The size of the country, the level of development of the transport system, and the government controls over trade and migration all contribute to the substantial regional income disparities that still remain among the twenty-six provinces and autonomous regions.

#### *Decentralization Policies and Regional Development in Korea and Malaysia*

Two interesting urbanization policy efforts are being made in Korea and Malaysia. Since Korea has one of the highest population densities in the world, and only 20 percent of its land is usable, it has developed one of the most comprehensive national land-use planning policies anywhere, with the dual aim of controlling the rapid growth of Seoul and reducing regional economic disparities. Those policies attempt to redress the differences in employment opportunities and in public and private services between Seoul and other regions that fuel rapid growth in the capital region. Under the land-use policies, licensing for manufacturing firms in Seoul was established, regional industrial estates were developed, and firms already in Seoul and considered foot-loose were ordered out of the city. Educational disparities were tackled systematically by regulating school registration in Seoul, by increasing the budgets and enrollment quotas of provincial universities, and by regionalizing educational budgets. The extremely rapid growth of the past two decades and the dispersion of heavy industrial investments around the country contributed to an equalization of regional opportunities, but the rapid rise in the level of urbanization caused by export-oriented growth policies and industrialization ensured that the largest urban centers, especially Seoul and Pusan, would continue to grow. The inevitable geographic decentralization of heavy industries for purely physical reasons did not compensate for the fact that 80 percent of manufacturing employment is generated by small and medium-size firms located outside planned industrial estates, as is the case in most countries. The unanswered questions in Korea at present are the extent to which concentration would have been worse in the absence of the policies, and the costs and benefits of control of industrial location.

Some of the most extensive efforts in Asia at develop-

ing lagging regions are found in Malaysia, because of its federal structure. Analytically, the most influential project is the Muda River irrigation project in the northwest section of the peninsula near Thailand, which was started in 1967 and has been the object of extensive studies since 1972. A major contribution of these studies has been the identification and measurement of the direct and indirect effects of large agricultural investments. In addition, they show how the effects are distributed among the people of the region and examine the extent to which project benefits leak outside the region. The work of Bell, Hazell, and Slade (1982) is one of the few which give a detailed perspective of rural-urban interactions. The authors found that, given regional land tenure patterns and the social structure, the indirect effects of the project were large. In particular, about 80 percent of additional value added was generated in the region for each dollar of value added directly generated by the project. This indirect income accrued mostly to nonagricultural households in the region that worked at nonfarm enterprises.

As the project matured, the net capital outflow to other regions increased significantly. Seven years after the start of the project the annual net capital outflow was estimated at 72 percent of regional household savings. The leakage was attributed to the lack of investment opportunities in the region compared with alternatives elsewhere in Malaysia. The relative isolation of the region and the high cost of transport limited the growth of an industrial base. Although large direct benefits to the region were provided, the results showed that self-sustaining growth was often more limited than one might expect. The analysis showed the importance of induced household demand in generating downstream benefits and the crucial role played by the sectors that produce nontradable goods and services. It was found that assistance to these sectors in the form of credit, skilled personnel, and access to markets might limit net capital outflows.

Another recent review of Malaysian regional policies examined the cumulative impact of major programs in the northeastern states of Kelantan and Trengganu, where economic development lags behind the rest of the peninsula and the per capita regional GDP is 56 percent below the national average. In addition to the truly regional and location-specific disadvantages of the states, a major finding of this review was that inadequate coordination of programs and projects implemented in the region has limited their impact. This finding for Malaysia illustrates the major problem of urbanization policies: often there is no regional plan or any formally stated development strategy because there is no political or administrative unit coterminous with the area that

requires the strategy. Each national agency develops its own programs at its own pace and on its own terms.

### *Problems of Large Cities in Latin America*

The level of urbanization in Latin America is high compared with that of other developing regions; it ranges from more than 80 percent in the southern zone (Argentina, Chile, and Uruguay) to less than 25 percent in Haiti. Urbanization in Latin America has been faster, more concentrated, and on a larger scale than in European countries. Latin America's experience is similar to that of other developing areas, but because urbanization took place earlier in Latin America than in other regions, urban institutions and policies are already well developed. In spite of the spontaneous shift of growth to intermediate cities—annual growth rates for which are now equal to or greater than those of the largest cities—the population problems of the supercities still dominate. Venezuela's industrial deconcentration policies illustrate the explicit urbanization policies in the region (see chapter 9, by Reif). In Brazil environmental regulation in São Paulo has had a significant but selective impact on suburban relocation of manufacturing (see Thomas 1981). Henderson, in chapter 7, reports on the impact of labor market structure on industrial growth and location in Brazil. A World Bank research project on spontaneous industrial decentralization in the state of São Paulo shows that, except for suburban relocation across the rings of the metropolitan region, only limited firm relocation takes place in intermediate cities (Hamer 1985). In addition, the local structure of the labor market is an important factor in the emergence of new firms. Economies of scale within an industry were found to be more important than city size in influencing firm location. These findings cast a new light on past work by regional economists, which focused on subsidies and tax holidays. They also indicate that in future work the gap between economic analysis and actual firm behavior and decisionmaking should be narrowed.

In Mexico efforts are being made to shape growth patterns within the capital region through a combination of transport investment, provision of infrastructure services, and regulation of investment. The dominant role of the capital region will not change; the priority is to make the region more efficient.

### **Conclusion**

Dominant considerations in the coming decade in the formulation of urbanization policies will be the general

economic slowdown, the scarcity of public resources, and the continuing high pace of urban growth in spite of relative declines in demographic growth rates. Comprehensive national urbanization policies dominated by a strong preoccupation with population distribution and with improvement of the urban system are not likely to be effective and should be de-emphasized. They are not realistic on several counts. First, they typically assume a complex structure of simultaneous programs which government agencies are unable to execute and which simply require too much program coordination and execution. Second, as has happened in several advanced economies, especially those with a federal structure, the desire to correct the unintended effects of national economic policies sets spatial policies on a collision course with other policies and policymakers. If urban development ministries attempt to displace finance ministries, their efforts will be futile, especially if their policies rest on weak analytical foundations.

On the positive side, the past decade has seen progress on two fronts. First, analyses of urbanization policies have improved noticeably. Second, policymakers have clearly become aware of the essential role of the urban economy in the national economy. Cities contain the leading economic sectors, they are the incubators for many innovations, and they provide a network for the flow of goods and services within a nation. As the example of India shows, even when a nation is still overwhelmingly rural, with 70 percent of its labor in agriculture, more than half of its GNP is produced in cities. Policymakers are now well aware of the crucial need for efficiently run cities.

Although studies such as that of Kelley and Williamson (chapter 3) are not yet an adequate basis for the formulation of population distribution policies, they are beginning to map out the links between economic growth strategies and urbanization. Such analyses improve the quality of predictions of future urban change and thereby facilitate the task of the public and private institutions responsible for urban affairs. Other recent work also shows why urban strategies will vary according to levels of development and urbanization.

The management of urban growth should continue to define the policy context in developing countries. The two priorities for the decade should be better internal management of cities by local governments and improvements in cities' resource bases and mobilization. Urban policies should incorporate more realistic sectoral policies regarding housing, transport, and the provision of utilities. They should concentrate every effort on seemingly unglamorous but fundamental managerial activities at the local level such as land registration, land cadastres, and mechanisms for cost recovery.

## Notes

1. U.N. urban population projections are derived from the estimation of urban logistic curves. See United Nations (1977).
2. See the recent cross-section and time-series estimation of an urban logistic curve by Mills and Becker (1982).
3. This profile of African urban policy reflects useful discussions with George Beier of the World Bank and draws from some of his recent unpublished work.
4. See World Bank (1981), table 4.1, page 41.
5. The development of the new federal capital in Abija and the revision of the constitution to increase the number of states are also likely to complicate the execution of urban policies.
6. The official figure of 1,031,882,511, as of July 1, 1982, includes Taiwan, Hong Kong, and Macao. The mainland population is 1,008,175,288, and the average annual mainland increase has been 17.4 million. Shanghai province has 11.8 million, Beijing has 9.2 million, and Tianjin has 7.7 million. These three largest cities represent only 12 percent of the urban population.

## Bibliography

- Bell, Clive, Peter Hazell, and Roger Slade. 1982. *Project Evaluation in Regional Perspective: A Study of an Irrigation Project in Northwest Malaysia*. Baltimore, Md.: Johns Hopkins University Press.
- Bertrand, Trent, and James Robertson. 1978. "An Analysis of Industrial Incentives and Location in Nigeria." Western Africa Regional Office, Projects Department, Industrial Development and Finance Division, World Bank, Washington, D.C. Processed.
- Hamer, Andrew Marshall. 1985. *Decentralized Urban Development and Industrial Location Behavior in São Paulo: A Synthesis of Research Issues and Conclusions*. World Bank Staff Working Paper 732. Washington, D.C.
- Henderson, J. Vernon. 1977. *Economic Theory and the Cities*. New York: Academic Press.
- Kelley, Allen C., and Jeffrey Williamson. 1984. *What Drives Third World City Growth? A Dynamic General Equilibrium Approach*. Princeton, N.J.: Princeton University Press.
- Lee, Kyu Sik. 1981. "Intra-Urban Location of Manufacturing Employment in Colombia." *Journal of Urban Economics*, vol. 9, no. 2 (March), pp. 222-41.
- Linn, Johannes F. 1983. *Cities in the Developing World: Policies for Their Equitable and Efficient Growth*. New York: Oxford University Press.
- Mills, Edwin, and Charles Becker. 1982. "Urbanization and Economic Development." Princeton University, Princeton, N.J. Processed.
- Renaud, Bertrand. 1981. *National Urbanization Policies in Developing Countries*. New York: Oxford University Press.
- . 1984. "Structural Changes in Advanced Economies and Their Impact on Cities in the 1980s." In Robert D. Ebel,

- ed., *Research in Development and Public Finance*, vol. 4. Greenwich, Conn.: JAI Press.
- Richardson, Harry W. 1977. *City Size and National Spatial Strategies in Developing Countries*. World Bank Staff Working Paper 252. Washington, D.C.
- Ruane, Frances. 1981. "On Modeling the Influence of Sectoral Policies on the Spatial Concentration of Industrial Activities." World Bank, Washington, D.C. Processed.
- Thomas, Vinod. 1981. *Pollution Control in São Paulo, Brazil: Costs, Benefits, and Effects on Industrial Location*. World Bank Staff Working Paper 501. Washington, D.C.
- Tolley, George, Philip E. Graves, and John L. Gardner, eds. 1979. *Urban Growth Policy in a Market Economy*. New York: Academic Press.
- Townroe, Peter M. 1979. *Employment Decentralization: Policy Instruments for Large Cities in Less Developed Countries*. Vol. 10, pt. 2, in D. R. Diamond and J. McLoughlin, eds., *Progress in Planning*. Elmsford, N.J.: Pergamon.
- Townroe, Peter M., and Vinod Thomas. 1983. "Pollution Control Policies and the Location of Industry." In T. O'Riordan and R. K. Turner, eds., *Progress in Resource Management and Environmental Planning*, vol. 4. John Wiley and Sons.
- United Nations. Population Division. 1977. *Manual 8. Methods for Projections of Urban and Rural Population*. Population Studies no. 55. New York.
- World Bank. 1979. *World Development Report 1979*. New York: Oxford University Press.
- . 1981. *World Development Report 1981*. New York: Oxford University Press.
- . 1982. *World Development Report 1982*. New York: Oxford University Press.

## *Urban Policy in Centralized Economies: China*

William L. Parish

The proper shape of Chinese cities has been a matter of intense policy discussion since 1949. There have been some constants: since the early 1950s government leaders have in principle consistently promoted small over large urban centers, tried to narrow the income gap between city and village, and sought to provide a secure floor of basic needs—health, education, housing, and essential food supplies—for everyone.

There have also been significant changes. Restriction of the growth of cities of any size was emphasized more in 1962–76 than before or since. Production rather than consumption was emphasized more before 1978. And housing, which is defined as a consumption good, was downplayed before 1978.

The task of this chapter is to trace the constants and fluctuations in Chinese urban policy since 1949 and to assess the net impact of these policies on city size and stability, productive urban employment, and adequate urban services.

### City Size and Stability

Post-1949 Chinese policy called for restricting the rate of urban growth in general, but policymakers were concerned particularly with the growth of such large coastal cities as Shanghai. The coastal cities, many of which blossomed after contact with foreign powers in the nineteenth century, epitomized to many Chinese what was wrong with the old social order. As centers of sin and decadent bourgeois habits, they violated the new spartan socialist ethic. Because they concentrated wealth on the coast, those cities also fostered a regional inequality that the new leaders wanted to alleviate. And,

given the tensions with the United States in the 1950s, large coastal cities were seen as vulnerable to military attack and as improper bases for vital industries. Starting in the 1950s investment funds were siphoned from coastal cities and reinvested in the interior. Factories and even some universities were dispersed from the east coast and reestablished further inland. New railroad lines were built to serve interior cities and to provide an effectively linked interior urban network (Kirkby 1985; Lardy 1978; Leung 1980).

Also, from the late 1950s through 1976, Chinese planners tried to limit the growth of cities of any size. In the mid-1950s Chinese cities continued to be plagued by persistent unemployment, and national planners increasingly came to feel that providing schools, parks, housing, and other amenities for new urban residents consumed scarce resources that could be better invested in industry and long-term economic growth. By 1958 a tough new migration law was passed that forbade people to enter cities unless they had a residence permit. The permit could be obtained only by persons who had secured a state job, and even then the employee's spouse and children could not accompany the new employee but had to remain in the countryside.

A number of measures helped make this antimigration law effective. Industries and work units came under government control and were prohibited from bidding freely for labor. All new positions had to be approved and allocated by the government labor bureau. Also, the structure of cities became increasingly cellular, with each local cell or neighborhood centered around a police station. The system remains largely the same today; each local police station registers every household allowed to live in the neighborhood. Periodically police

and local neighborhood officers conduct residency checks to root out those who lack a valid residence permit.

Finally, the system of rationed goods provided restraints against urban residence by anyone missed by the police. Until recently, Chinese cities had one of the longest lists of rationed goods in the world, typically including such basic necessities as grain, meat, cooking oil, bean curd, sugar, laundry soap, cotton cloth, coal, kindling, toilet paper, bicycles, sewing machines, major pieces of furniture, and, occasionally, other items that were temporarily in short supply. With strict restraints on private markets, black market prices were so high as to make it extremely difficult for a person to live in the city without proper registration for any length of time.

Without these kinds of extreme administrative measures, cities would likely have grown much more rapidly, given the large gap in incomes and services between city and countryside. Despite pro-peasant verbiage, the gap between urban and rural incomes remained considerable throughout the first three decades of socialist rule. The new leaders raised the prices paid to farmers, and rural taxes declined, but at the same time rural resources were siphoned off through the high prices farmers had to pay for industrial goods (Xue 1981, pp. 177–85). The government gave little in return, spending only a little over 10 percent of its total investment on agriculture (“Commentator” 1979).

The results of these policies are obscured by the hidden consumption urban residents enjoy over rural residents in subsidized housing, food, medical benefits, and other social services, all of which have increased over the years. One set of national account statistics, however, shows that per capita consumption for the nonagricultural population compared with the agricultural population continued to climb through the 1970s and declined only with new agricultural policies in the 1980s (State Statistical Bureau 1985b, p. 552):

*Consumption ratio, nonagricultural  
to agricultural population*

1952	2.4:1
1957	2.6:1
1965	2.4:1
1975	2.6:1
1978	2.9:1
1984	2.2:1

Thus, despite its origins in a peasant revolution, China seems to have suffered some of the same problems of urban bias found in many other developing societies. Even in 1984, when conditions had improved for the rural population, the urban-rural consumption gap remained as large as or larger than that in India, Pakistan, the Philippines, and Thailand (Lipton 1977, p. 430).<sup>1</sup> Throughout the past three decades, then, there have

been many reasons for peasants to want to desert village for urban life. If that desire was to be curtailed, it had to be done through administrative measures which denied peasants the benefits of urban life.

Chinese statistics on urban growth were previously clouded by changes in the definitions of cities—in the inclusion or exclusion of farmers living in or adjacent to towns, and in the size of towns counted as urban. We have made some progress in sorting out these changes, however, and the general trends are now clear (see Chan and Xu 1985; Kirkby 1985). In the 1950s city growth was very rapid, nearly 8 percent a year. In the 1960s and early 1970s, when strict migration controls were in effect, urban growth slowed to a mere 2 percent a year. But since 1977, with the relaxation of urban controls and the promotion of small towns, urban growth has accelerated to over 5 percent a year (Chan and Xu 1985; State Statistical Bureau 1984).

China’s simple urban growth rates, then, have been distinctive only in the 1960s and early 1970s. Both before and since, China has been similar to other developing societies. Table 6-1 captures part of this comparison: during the past two decades, which include both pre-1977 and post-1977 trends, cities in China grew more slowly than in most developing countries, but the post-1977 return to growth rates of over 5 percent brought China’s urban growth rates back into line with those of most low-income developing countries.

China has been more consistently distinctive in the distribution of urban growth among cities of different sizes. Much like India, which also has a large internal market and a long history of cities scattered throughout the country as much for administrative as for trade purposes, China began the modern era with several large cities rather than a single primate city that dominated the urban landscape.<sup>2</sup> Even after Western contact, cities of different sizes were spread relatively evenly across the national landscape, with only a slight tilt toward the large coastal cities favored by foreign traders (Berry and Horton 1970, pp. 67–75; Skinner 1977).

The policies of removing even that small tilt toward large coastal cities was in many ways successful. The largest city, Shanghai, with a population of 6 million in 1949, accounted for only 11 percent of the total urban population in China. In subsequent years, Shanghai grew slowly, and its share of total urban population declined steadily until it accounted for only 5 percent of total urban population in the 1960s and 3 percent in the 1980s (table 6-1).

This pattern was repeated, by and large, for other large cities. A comparison of 1953 and 1983 census results shows how smaller cities tended to catch up with larger cities (State Statistical Bureau 1983; Ullman 1961):<sup>3</sup>

Table 6-1. *Urbanization in Selected Countries*

Country or group	Urban population			Percentage of urban population		
	As percent- age of total population		Average annual growth rate, 1970-80	In cities of over 500,000, 1980	In largest city	
	1960	1980			1960	1980
China <sup>a</sup>	18	21	2.7	53	5	3
India	18	22	3.3	39	7	6
Pakistan	22	28	4.3	51	20	21
Indonesia	15	20	4.0	50	20	23
Low-income countries <sup>b</sup>	11	19	5.4	40	25	28
Middle-income countries <sup>b</sup>	33	45	4.0	48	28	29

a. 1964 and 1982.

b. Averages are weighted by population. The low-income average excludes China and India.

Sources: China, State Statistical Bureau (1985a), pp. 54, 58, 90, 549; World Bank (1982), table 20.

1953 population	Annual growth, 1953-83 (percent)	Number of cities
2,000,000 and over	1.8	4
1,000,000-1,999,999	2.1	5
500,000-999,999	2.5	15
200,000-499,999	3.1	29
100,000-199,999	4.5	42
50,000-99,999	4.8	41

In line with these growth rates, the overall rank size distribution of cities began to more closely approximate the even spread of cities of different sizes that had characterized China before Western contact in the nineteenth century. And, as the policymakers intended, cities in the interior grew more than cities along the coast (Goldstein 1985, p. 35; Kirkby 1985, ch. 5).

We should not overstate the trends, however, nor assume that the issue of China's future direction is settled. The shifts to the interior and toward smaller cities were modest and there were important exceptions—for example, the 3.5 percent annual growth rate of the second largest city, Beijing.<sup>4</sup>

In addition, there were many restraints on the growth of the smallest places, those with less than 20,000 population. Despite protestations of support for small industry in rural towns, central planners tended to favor larger industry. And the hostility to private marketing meant that much of the informal activity that provided the base for small market towns was lost. Chinese scholars now complain that many small market towns stagnated (see, for example, Fei 1984). In official statistics, most of these small places even lost their urban classification after 1953, making it nearly impossible to trace their precise development in the intervening years. It is only in the past few years that these places have been allowed to reinvent themselves, to open more industry on their own, and to be recertified as

urban in official statistics.<sup>5</sup> All of this helps to explain why the proportion of China's urban population in places with more than 500,000 persons is so high in relation to that in other developing countries (table 6-1). The figure is in part a definitional fluke stemming from the exclusion of many small places, but it also correctly suggests that it was mostly the middle-level cities that functioned as centers of administration and of state-owned industry which prospered in the middle years of the new Chinese social order.

Officially, the current policy is to promote the growth of the smaller rural market towns which were discriminated against previously, but there is a continuing academic debate within China about whether this is the wisest course. Some argue that a poor country like China cannot afford to support inefficient small towns (see Kirkby 1985, pp. 230-42). And even if official policy continues to promote small towns in principle, the unusually high costs of raw materials, energy, and transport faced by producers in these places make their viability uncertain. All this may account for the high business failure rate reported in some sources (Fei 1984). Much as in other developing societies, failure to pay attention to—or an inability to deal with—the many indirect policy consequences of taxation, public investment, and controlled prices may be what distorts well-intentioned but eventually unsuccessful programs of small-town development (Linn 1983).

If the small-town policy is unsuccessful, large cities could be flooded with new migrants. Already, with the return to family-based farming, a third of the agricultural labor force has abandoned simple grain production. In addition, the reduction of the number of goods rationed in cities and the opening of urban free markets have decreased the number of administrative tools for keeping peasants out of cities. Indeed, a few cities have

already begun to issue temporary residence permits to peasants (Solinger 1985). These permits continue to deny peasants the many subsidized rations, housing, and other services offered to normal urban residents, but they do attempt to regularize some of the inevitable movement into cities.

China, then, has followed a distinctive course in urban development, but its course has been neither a simple one nor one which is certain to remain constant in the future. This is a time of transition in which there is likely to be frequent fine-tuning to try to meet evolving and sometimes unanticipated economic conditions.

### Useful and Productive Employment

China's new socialist leaders wanted able-bodied members of society to have secure, productive jobs. A worker's future career was no longer to be endangered by the whim of fickle private employers but rather was to be guaranteed by the state. No longer was labor to be wasted on bourgeois consumption; it was to be invested in productive activities that led to rapid growth of industry and the provision of basic staples for everyone regardless of income. The pursuit of these goals has been successful, and the result is a pattern of urban employment that is very different from that found elsewhere. Nevertheless, some of the problems of unemployment and poor services encountered elsewhere have also emerged in China.

### Secure Employment

Urban employment has increased sharply over the past three decades. This is seen imperfectly but distinctly in surveys of families of workers employed in state and urban collective enterprises. Judging from the inverse of the number of dependents in these families, the employed population has almost doubled since the 1950s (State Statistical Bureau 1985b, p. 561).<sup>6</sup>

	<i>Percent of nonagricultural population employed</i>
1957	30
1964	29
1978	49
1984	58

Several factors contributed to this trend. One was the policy of excluding peasants who might have come to seek nonexistent jobs and of purging cities of those without proper employment. For example, in 1957 and 1962 there were drives to send peasants and "vagrants" back to their former home villages or to state farms

where they could be usefully employed. In 1962 alone 20 million people returned to the countryside (Hu 1982).

A second factor was the increasing employment of women. Though this change is not immediately obvious in the table above, female employment apparently began to increase in the late 1950s and became almost universal by the late 1960s. By the 1980s the vast majority of women in their twenties and thirties worked outside the home—not in part-time jobs, as in many other societies, but in full-time jobs that kept them busy eight hours a day, five and a half to six days a week.

The 1982 census results for working-age men (16–59) and women (16–54) in cities (mostly over 100,000 population) and towns (mostly under 100,000 population) suggest the pattern. The employed population of both males and females was extremely high (State Statistical Bureau 1985a, tables 40, 58):

	<i>Employed males (percent)</i>	<i>Employed females (percent)</i>
Cities	94	84
Towns	92	78

A third factor, which is related to increased female employment, is the sharp drop in the urban total fertility rate, from a high of 6.2 children in 1963 to a low of 1.1 in 1980 (Coale 1984, p. 59). By the 1980s urban fertility was extremely low, and there were far fewer children for each family to support. Women were freed for full-time labor, and the number of unemployed dependents declined.

Finally, the most recent rise in employment is tied to efforts to create new employment opportunities (see "New Solutions," below).

While employment increased, associated security and fringe benefits also increased steadily for the majority of the urban population. By 1957 the state (including central, provincial, city, and county authorities) had begun to take over most major industries. In that year the state already employed 68 percent of the nonagricultural work force, and by 1982 that proportion had risen to 77 percent. Also during this period the private work force of artisans, peddlers, and other individual workers declined steadily, and by 1982 it constituted only 1 percent of the nonagricultural work force.<sup>7</sup> Those who joined state-owned work units enjoyed many benefits, such as health insurance, which paid virtually all their medical expenses and half the expenses for their dependents as well. A pension system provided an average 70 percent (now 75 percent) of preretirement pay. A hardship allowance took care of those whose family dropped below a certain poverty floor. There was disability insurance, lengthy sick leave, and many supplements that took care of special clothing and transport needs. Finally, it was

virtually impossible for anyone to be laid off from jobs in the state sector, and life for workers was far more secure than it had been in the past. These conditions still obtain today. Workers in a similar sector in other developing societies are sometimes referred to as the aristocrats of the labor force. In China these aristocrats are more numerous and probably enjoy more wide-ranging benefits than in most other developing societies.

### *Informal Sector*

In spite of these advances, not all workers made it into the well-paid, secure state sector. Another sector of collective enterprises, run by neighborhoods, towns, and even larger state enterprises, has persisted to a degree that seems remarkable in a socialist state. These collective enterprises tend to be smaller labor-intensive units and to have lower wages, less security, and fewer benefits than the average state enterprise. The contrast between the two sectors is reminiscent of the distinction between formal and informal workers in other developing countries.

Current statistics show that the average income for workers in the collective sector in China is only 78 percent of that of workers in the primary, state sector (State Statistical Bureau 1985b, p. 558). The common absence of pension, medical, and disability benefits for collective workers increases the gap. Also, in ways analogous to the situation in other developing countries, the secondary, collective sector attracts a disproportionate share of women; some are older women with less education, but some are younger. In 1984 women were only 32 percent of the state work force but constituted 48 percent of the collective work force (State Statistical Bureau 1985b, pp. 230–32).

In other ways, however, the formal-informal distinction in China is unlike that in other developing countries. In China the secondary sector is smaller. It contains only one-fourth of the nonagricultural labor force, as compared with one-half in many other developing societies.<sup>8</sup> In China workers in the secondary sector are not much poorer than the primary labor force. They make four-fifths as much as those in the state sector, whereas in other countries informal workers make less than one-half the wage of workers in the formal sector. Further, the age, sex, and educational differences between the two sectors in China seem modest in comparison to those in other countries (Whyte and Parish 1984, p. 32). The most striking characteristic, however, is that employment in the second sector remains relatively secure in China. Employees are seldom laid off, just as they are not laid off in the state sector. Thus, in spite of certain similarities, the degree of dualism between mod-

ern formal employment and traditional informal employment seems much more muted in China than elsewhere.

### *Services*

The distinction between urban work in China and elsewhere is also seen in the types of job pursued. In many developing societies menial service jobs such as washing cars, shining shoes, peddling trinkets, and working as household help have been saturated by peasants in search of nonexistent jobs in the cities. These are exactly the kinds of bourgeois nonproductive activity that Chinese leaders set out to eliminate in the 1950s. The change in the composition of the labor force since then shows how successful they have been.

Today in developing societies the distribution of the nonagricultural labor force tends to mimic that of developed economies such as the United States (see table 6-2). The United States has come to have a large proportion of its labor force in retail trade, restaurants, hotels, government, education, and other services late in its economic development, whereas the average developing society, with a large sector of informal service workers and a large government bureaucracy, has tended to leapfrog the gradual evolutionary process. No matter what the maturation process, however, most developed economies today, like the United States, have over half their labor forces in information and service activities.

Socialist societies, in contrast, have tended to emphasize "productive" work in manufacturing, mining, and construction and have downplayed such "consumption" activities as finance, retail trade, and restaurants. Even government has shrunk as a proportion of the labor force, which is surprising given the increased administrative needs of centrally planned economies. China has rapidly adopted the socialist program. By the late 1970s China's labor force had been remodeled into the productive mold that China's leaders desired. By 1982 over half its nonagricultural labor force was in manufacturing and mining, while only about a third was in finance, retail trade, government, and other service activities.

This emphasis on production instead of bourgeois consumption had substantial costs. Because of the pre-1978 emphasis on heavy producer industries rather than light consumer industries, supplies of consumer goods remained minimal. Although production of radios, watches, vacuum bottles, and bicycles increased greatly, many consumer goods such as televisions, clothing, food, and books remained severely restricted in both number and variety. As demand outpaced supply, many of these goods, including bicycles, some foods, and the better watches, had to be rationed by cumbersome pro-

Table 6-2. *Nonagricultural Labor Force, by Economic Activity and Type of Economy*  
(percent)

<i>Economic activity</i>	<i>Market economies</i>		<i>Socialist economies</i>	
	<i>United States</i>	<i>Developing economies<sup>a</sup></i>	<i>China</i>	<i>Eastern Europe<sup>b</sup></i>
Manufacturing and mining	28	27	52	46
Construction	5	8	8	10
Transport and public utilities	8	9	8	11
Finance	5	2	1	1
Trade, restaurants, and hotels	21	21	11	10
Government and other services	33	33	21	22

*Note:* All data are for about 1970 except those for China, which are for 1982. Figures may not add to totals because of rounding.

a. Twelve societies with 1970 per capita GNP of less than \$1,200, ranging from the poorest, India, to the richest, Argentina.

b. German Democratic Republic, Hungary, Poland, and Romania.

*Sources:* China, State Statistical Bureau (1985a), table 54; United States, U.S. Bureau of the Census (1976), table 597; other countries, United Nations (1973), table 10, and United Nations (1974), table 40.

cedures. Any increase in unrationed goods, or a change in the variety of goods, often drew a crowd of potential buyers to a store.

Consumer supplies and services were often distributed through cumbersome channels. Gone were the peddlers and artisans who had served Chinese neighborhoods. Consumers now had to go to centralized stores and service centers that might be some distance from their residences. And because of the sharp decline in the number of restaurants, tailor shops, bicycle repair shops, barber shops, and so forth, consumers not only had to shop all over town to find what they needed, but they also had to wait in long lines once there. Food market lines formed before and after work and especially on Sundays, when more people had the day off. This situation was particularly difficult for the increasing number of dual-career families who had no adult at home during the week.

By the end of the 1970s bad conditions were beginning to lead to many complaints in the press, including the cry that families spent so much time searching for supplies on their day off that they worked harder on Sunday than during the rest of the week. The time spent scouring for basic supplies and services left little time for recreation or other joint family activities. The goal of involving people in productive work was coming into direct conflict with the goal of improving the quality of life for families in cities. Chinese press reports suggest that because of the difficulties with supplies and services, many people began to perceive the quality of life in cities as deteriorating in the 1970s.

### *Youth Unemployment*

An additional problem caused by the emphasis on production and heavy industry was increasing youth

unemployment. State investment went largely into heavy industry, which generated only 94 jobs per 1,000 yuan investment, whereas a similar investment in light industry could have generated roughly 257 jobs. Thus, the state was doing little to maximize new employment opportunities (Zhao 1980). By the early 1960s newly expanded urban secondary schools were beginning to graduate more than 2 million new job seekers every year. At the same time, retirement rates were low within the young labor force that had flooded the growing state economy during the 1950s. The result was an employment crisis for educated youths that could be solved only by sending urban youths to the countryside, and during 1966–76, 17 million youths were sent to rural areas. This makeshift solution to the employment program proved problematic. Many youths who could not cope with rural living conditions slipped back into the city, where they could not legally get a job or proper urban rations.

By the start of 1979, when the program of sending youths to the countryside began to be phased out, there were 5 million to 12 million unemployed youths in the cities, or 5 to 11 percent of the nonagricultural labor force.<sup>9</sup> Thus, China was encountering a problem common to developing-country cities: the rapid expansion of the educational system up to the middle school (high school) level produced a new group of educated youths with high aspirations. They hoped to get the same type of good jobs that their older siblings and neighbors had obtained, but because the value of education had decreased, those jobs were not available to them. And because of China's investment policies, jobs available in larger cities were very scarce. The result was widespread frustration and disillusionment and an outbreak of petty crime that alarmed many parents and central leaders (see Whyte and Parish 1984, ch. 8).

### New Solutions

The problems with youths and with consumer goods and services helped bring about a reversal in several economic and educational policies after 1978. State investment shifted slightly from heavy industry to light industry, which not only produced more consumer goods but also provided more jobs. Total investment, which was roughly 36 percent of net material product in 1978, was reduced to release funds for current income and for an increase in consumption to promote the growth of job-producing light industry. Collective enterprises supervised by neighborhoods and towns were encouraged—employment in this sector rose from 22 percent of the nonagricultural labor force to 26 percent. And more people were allowed to go into business on their own: the share of self-employed persons in the nonagricultural labor force rose from 1 percent in 1978 to 3 percent in 1984 (State Statistical Bureau 1985b, p. 214). Peasants were permitted to come into the middle of cities to sell their goods in newly reopened peasant free markets, and the variety of goods available to consumers increased. Some full-scale middle schools that were turning out graduates without job-related skills were closed, and more technical schools to train students in such skills were opened.

The results of the changes have been a significant drop in unemployment and a considerable increase in consumer goods and services. Not everyone has gotten a state job. Two-fifths of those employed over the past few years have been employed outside the state sector, many in neighborhood collective enterprises or in individual activities that they started themselves, thus receiving more variable income and fewer fringe benefits than they might wish (State Statistical Bureau 1984, p. 130). Some employees ended up in temporary jobs without a secure future. Nevertheless, by the end of 1981 outright unemployment had declined to only 2.4 percent of the nonagricultural labor force, and the remaining unemployed were mostly those who had graduated from middle school during the previous year (State Statistical Bureau 1985a, table 68; *Zhongguo Baike Nianjian* 1984, pp. 619–20).

Supplies and services also increased. With more emphasis on light industry, state stores began to carry more clothing, bicycles, tape recorders, and other consumer goods. Neighborhoods began to have more collective restaurants, laundries, bicycle repair stalls, and other service centers that catered to everyday needs. Peasant stalls on designated streets increased the variety and freshness of eggs, fish, peanuts, vegetables, and other goods available to urban residents. Many items were still rationed, and shortages continued to occur for some items, but the supply was considerably improved

over earlier years. With more colorful clothing available and more artisans and service activities in the streets, cities began to take on a livelier appearance.

### Health Care and Housing

The ambivalent attitude of many of China's leaders toward urban services has been noted. Leaders have been in agreement that urban dwellers should be supplied with such basic human needs as grain, essential cotton clothing, and an initial ten years or more of education. It is widely accepted that these should be available to everyone on a relatively equal basis. There has been less agreement on social services and supplies beyond the basic needs. The volume of goods and services supplied and the degree of insistence on equal distribution of those goods and services have shifted with the political winds.

#### Health Care

In China, as in other socialist states, health care has been consistently defined as a basic human need and a human right, but the degree of insistence on equal distribution of the services, particularly to the countryside, has varied. Throughout, provision of ample medical care has had a high priority. As a result, medical services have grown steadily. By the 1970s the number of persons for each doctor, nurse, or hospital bed was much lower in China than in other developing societies, as the table shows (n.a. signifies not available).<sup>10</sup>

Service unit	Population per service unit		
	China	Low-income countries	Middle-income countries
Western doctor	2,470	9,900	4,310
Doctor, any type	1,172	n.a.	n.a.
Nurse	2,306	8,790	1,860
Hospital bed	503	1,730	570

With this wide distribution of services, the reported infant mortality rate in China dropped to only 49 per 1,000 live births—less than half that of India—and the life expectancy at birth rose to age 68, eight years more than in the average middle-income developing country and eighteen years more than in the average low-income developing country (see note 10 for data sources). Health care in cities was even better than these national averages suggest.

Cities are particularly well served by this system. In large cities provision of care is three-tiered. Large work units and many neighborhoods have a first aid station staffed by nurses and paramedics to treat minor injuries.

People with more serious illnesses go to the local hospital that is assigned to their neighborhood or work unit. If the hospital evaluates a problem as being more serious, it can send the patient to a city or provincial hospital for specialized care. The system is not without its problems. Since those employed in the state sector receive essentially free medical care, the patient load is high at local hospitals. Emergency cases with raging fevers and other signs of trauma can be taken immediately, but others may have to line up early in the morning to take a number for an appointment that day. Complaints arise because of the lack of choice of doctor and the few minutes of attention the doctor provides. Another problem is that medicines are sometimes in short supply in relation to demand.<sup>11</sup> But, overall, the system seems to have worked well in supplying basic care, either free or at low cost, compared with the care available in most developing-country cities.

In large cities, preventive public health activities are also well organized. The city periodically organizes, through the cellular neighborhood and work-unit structures, inoculation campaigns to ensure that everyone is protected against disease. Neighborhoods and work units also help publicize the need for sanitation and for the eradication of pests and vermin. Cities organize the daily collection of human sewage from homes and public latrines as well as the daily sweeping of major streets. Neighborhoods help organize the sweeping of inner streets, garbage collection, and the seasonal cleaning of houses. For seasonal cleanings the neighborhood may distribute fumigating agents and then have everyone gather for a movie while fumigating is taking place. Neighborhood leaders may also inspect each house or apartment to make sure that it has been properly cleaned and leave a small colored sticker on the front door to indicate that the house has passed inspection. These public sanitation activities are less thorough in small towns and cities, but they help make China's large cities much cleaner than many other developing-country cities.

The conditions listed above, when combined with subsidized food supplies and ample nutritional support, dramatically improved urban health. The rural pattern is more mixed. Nationwide, in 1975 the average life expectancy for city dwellers was 72 years, while the figure for people in the countryside was only 57 years. In part, the fifteen-year gap documents the extraordinarily good urban health care and prevention system rather than the neglect of the countryside, but it is an astonishing gap, nevertheless, and one matched by few other developing countries. The gap is reflected also in rates of nutritional stunting as measured by the percentage of children less than 90 percent of the average height of their age group. In cities the 1979 rate was only 3 percent, while in

villages it was 13 percent. The causes of these differences rest heavily on food supplies and public sanitation—conditions that are difficult to change in widely scattered villages. But to some extent the differences are also attributable to state expenditure patterns that favor urban consumption and state medical expenditures (which, on a per capita basis, are almost ten times greater in cities than in villages), and to a similar bias toward cities in the supply of doctors, nurses, and facilities (Jamison and others 1984, pp. 11, 32, 93). In health, as in overall income, it has been difficult to remove some of the tilt toward cities, and China again repeats some of the patterns found in other developing societies.

### *Housing*

Housing has been subject to sharp fluctuations in policy. In the early 1950s there was an attempt to clean out major slum areas and erect new apartment buildings in their place. But by the mid-1950s housing was declared a consumption good, undeserving of major new investment. It was not until after 1976 that planners once again chose to provide major resources for housing the urban population.

The consequences stand out sharply in statistics on floor space per capita. From a high of 4.5 square meters per capita in the early 1950s, urban floor space declined steadily until it reached a low of 3.6 square meters in 1978. Only with massive urban building in recent years did housing reach a new per capita high of 4.6 square meters in 1983 (*Zhongguo Baike Nianjian* 1984, p. 622; Zhou 1979).<sup>12</sup>

By some indicators, urban housing was not all that terrible even in the middle 1970s. The stringent limitations on migration averted the growth of shanty towns crammed with peasants seeking urban work. And according to one set of statistics, based on some of the more prosperous regions of south China, Chinese urban housing was as good as or better than housing in many other cities of the developing world (Whyte and Parish 1984, p. 78). These statistics imply that Chinese urban dwellers were no more likely than people in other developing-country cities to be crammed three to a room and they were somewhat more likely to have electric lighting and a kitchen, although that kitchen might have to be shared with others in the same building. Nine-tenths lived in brick and concrete structures instead of makeshift wood and tin shantytowns.

Nevertheless, there were many serious problems with housing in the 1970s. The press complained that 35 percent of all urban families lived in overcrowded conditions, that 20 percent lived in slums, and that 5 to 6 percent had no proper housing at all ("Summary" 1979; Zhou and Lin 1980). Not only were individual families

crowded, but they also often had to share toilet and kitchen facilities with other residents in the same building. This sharing was and continues to be a frequent source of tension between families, with one family complaining that the other is leaving garbage scattered about in the kitchen or is monopolizing precious bathing and toilet facilities. Indeed, one account of crime in Shanghai claims that such disputes between neighbors continue to occupy an inordinate amount of police attention (Zhao 1984).

Another difficulty with public housing, common to many socialist societies, is that rents are set too low to cover yearly upkeep as well as building costs (*Zhongguo Baike Nianjian* 1981, p. 541). Rent was only 1.4 percent of 1984 urban family budgets, and expenditures for rent, electricity, water, and cooking and heating fuels totaled only 4.2 percent, considerably less than the 10 percent typical of other developing societies (Lluch and others 1977, p. 40; State Statistical Bureau 1985b, p. 567). As a consequence of the low rents and the inadequacy of state budgets to make up the difference, many existing buildings deteriorated (*Zhongguo Baike Nianjian* 1981, p. 541). One press source estimated that "more than 50 percent of the houses in the urban areas of China are in a bad state of repair and need maintenance, of which more than 10 percent are in dangerous condition" (Zhou and Lin 1980).

In the 1970s, there were additional complaints from professionals and administrators forced to share the space standards and amenities of persons in blue-collar jobs. Other complaints stemmed from inability to choose one's house or neighbors (or to move away from contentious neighbors) when housing was allocated through bureaucratic channels (Whyte and Parish 1984, pp. 76-85).

Since 1978 some of these problems have been alleviated. Floor space has increased, particularly for professionals and administrators. And in new buildings with a toilet and kitchen for each apartment there is less reason for friction among neighbors.

Some problems remain, however. One study in Beijing complains that, despite a massive building program in that city, the percentage of families living in severely overcrowded housing remained unchanged through 1982 (Liu 1984). "Severely overcrowded" is defined as a per capita living space of less than 2 square meters, no room of one's own after marriage, or three generations in the same room. Even while average floor space in Beijing increased from 4.6 to 5.7 square meters in four years, the percentage of severely overcrowded households remained at 14 percent.

Several reasons, in addition to a growing number of marriage-age youths, contributed to this problem. One was the increasing tendency for housing to be built not

by centralized city housing authorities but by work units (factories, schools, and bureaus). Nationwide, in 1983, 58 percent of housing was owned by work units, 26 percent by city housing authorities, and 17 percent by private individuals (*Zhongguo Baike Nianjian* 1984, p. 623). With the emphasis on work unit housing, richer worker units with housing that was relatively good to begin with provided even better new housing. Poor work units built little housing, as did most private individuals living in their old family homes. It was a situation of the rich getting richer while the poor stayed poor, at least in their housing.

Another contributing factor was low rent, which had two unforeseen consequences: rents provided insufficient money for new building or maintenance, and since rent was less than 2 percent of a family's budget, it placed no restraint on overuse. Families with more space than they needed felt no compulsion to move to smaller quarters, and local housing authorities were powerless to get them out in favor of larger families who desperately needed more space (Liu 1984, p. 28).

Some issues with housing remain, then. If rents continue to be low, they will provide little income for new building or for repair, and they will fail to ration scarce housing resources. If work units continue to build much of the new housing, some of the old inequalities will remain. And, as in the past, bureaucratic allocation of most housing will limit choices in both housing and neighbors.

## Conclusion

China's history since 1949 illustrates both the potential and the disadvantages of urban development in a centrally planned economy. With much greater control over economic resources than in the average market economy, the government was able to shift investment funds to promote the development of medium over large and interior over coastal cities. All of this helped reduce regional inequalities. In addition, the control over jobs and rationed consumer supplies meant that for a time the government was able to limit severely the growth of all cities, and funds that might have been spent on an elaborate urban infrastructure for waves of new migrants was spent instead on rapid industrial growth.

As in other centrally planned economies, the government has been able to put the urban population to work. Most able-bodied women and more than half the urban population have jobs. Few of those jobs are part-time or likely to be lost tomorrow; they are primarily full-time jobs that promise to last for a full career. In these ways Chinese cities have avoided some of the problems of employment instability that have afflicted other developing-country cities.

There have been policy mistakes, however, that have created in China some of the same problems found in other societies. It was a mistake to downplay the role of light industry and consumer services. Some developing-country cities may have too many people in informal service activities, but China illustrates the problems of the opposite extreme. With so little opportunity for light industrial growth and with bureaucratic restraints on small, individualistic enterprises, problems of youth unemployment began to appear. The unemployment problem was exacerbated by the rapid expansion of employment for women. Some jobs were created during the 1960s and 1970s, but since those jobs were taken almost as frequently by women as by men, the need to create additional openings was greater than in other developing societies. Furthermore, the rapid expansion of secondary school education, as in many other developing societies, contributed to a growing corps of unemployed educated persons. China avoided this problem at the university level but not at the secondary level. The unemployed youths were particularly frustrated because the same level of education had guaranteed good jobs just a decade before, when education was less common. This is a frustration shared with many youths in developing societies, but it was perhaps felt more acutely because of the socialist promise of secure jobs and rapid development. That frustration contributed to the outbreak of petty crime in the 1970s and continues to fuel the social alienation of some youths today. As we have already noted, the slighting of light industry and services, as well as the minimal investment in housing, also helped create a perception among many urbanites that the quality of urban life in the 1970s was declining and the socialist promise was going unrealized.

In brief, China's experience in virtually eliminating small informal service activities illustrates the necessity for these types of activities in developing cities. Hackenberg (1980) argues that because of high population growth rates and the difficulty of centralized provision of essential urban services, the informal service sector in poor societies is essential and should be embraced rather than shunned. Although they reject the extreme version of this position, China's leaders seem to be reluctantly moving in this direction. Smaller and more makeshift work arrangements, organized ad hoc by neighborhoods and individuals, with lower rates of pay and security, are now approved as a way of providing both employment and essential urban services.

China's experience also illustrates the difficulty of eliminating the gap between city and countryside. In many ways China seems to have worked much harder at eliminating this gap than have other societies. Through the encouragement of bootstrap operations, villagers have been induced to help build an infrastructure of waterworks, roads, level fields, schools, public health

care, and administration that is the envy of many other societies. The government has tried to help directly by raising prices for agricultural products. But a closer examination shows that before the 1980s the chimera of rapid industrial growth helped shape government investments and industrial prices in such a way that agriculture grew less rapidly than might have been expected. And when it comes to state-subsidized services, including housing, health care, education, and cheap food, it is still the urban sector that has reaped the most benefits. These characteristics that China shares with many other developing societies illustrate the continuing problems of urban and industrial bias even in a society that has overtly renounced this route to development.

China has begun to correct many of these difficulties. The government has induced a rapid rise in peasant incomes by further raising rural purchase prices and by removing many administrative constraints on farmer's activities. The promotion of small market towns and new collective enterprises is among the measures which have helped increase peasant incomes. As a result, some individual farmers have reaped rewards unimaginable even to the highest-paid urbanite. On a national basis as well, farmers have finally begun to narrow the gap between average rural and urban incomes.

Some of these policies still hang in the balance. Whether small market towns will indeed prosper and absorb much of the excess rural labor force remains uncertain. Investments in local transport, small trucks, cheap gasoline, and cheap electricity may have been delayed too long. Without these measures, farmers are likely to find economic activities in small towns unviable and clamor to go to large and medium-size cities.

Already, as a result of the increasing role of peasant free markets in cities, the reduction in the number of rationed goods, the undertaking of new construction that requires unskilled labor, and the general relaxation of administrative control, the number of quasi-legal migrants in cities is rising. Increasing foreign trade may also create more pressure for major port cities such as Shanghai to grow in ways that they have not in the past. Thus, while much of the basic structure may remain in place, we may again see a rapid growth of large cities that will make China more similar to other developing countries. Only close attention to the details of prices, taxes, and investment will provide either observers or Chinese planners a clue as to which way this set of development forces will tilt.

## Notes

1. The comparative data are mostly for the 1960s, and the ratios for these countries range from a low of 1.1 to a high of 2.2. With the addition of Chinese urban subsidies, which in one

source are said to have totaled 164 yuan per capita in 1981, the true 1984 Chinese urban-rural ratio could approach 3:1. This would make the comparison with other countries even more striking—although public expenditures and subsidies will have some urban bias in the other countries as well [see *Beijing Review*, no. 43 (1982), p. 7].

2. A primate urban structure is one in which much of the total urban population is in the largest city.

3. Sex ratios provide an additional indicator of the faster growth of small places. Chinese cities (mostly over 100,000 population) have 108 males for each 100 females, only slightly higher than the ratio for rural areas of 104, which suggests that a modest number of males have entered cities. In contrast, towns (mostly of 10,000–100,000 population) have ratios of 116, implying much more migration. In the 30–34 age group, which includes many new male laborers moving from the countryside, the ratios are 112 for large cities, 106 for rural areas, and 129 for towns (State Statistical Bureau 1985a, tables 35–37). Thirteen million peasants took work in towns and cities in the decade 1966–76. These peasants were probably moving to towns even while 17 million youths were being expelled from larger cities (see Hu 1982).

4. Among cities with over 100,000 population, the regression slope of city size on city rank (both logged) flattened only slightly, from  $-0.95$  to  $-0.88$ , between 1953 and 1982 (raw data from State Statistical Bureau 1983; Ullman 1961).

5. Between the 1953 and 1982 censuses, the number of designated urban centers was almost halved, from about 5,500 to 2,990. Many places regained their urban designation in 1983, and 10,000 places were eventually to be designated urban, in line with their new economic role (Goldstein 1985, p. 67). These changing designations probably had only modest effects on the growth rates reported above. Some of the reduction in centers occurred before the 1966–76 slowdown in growth, and the 1978–82 upturn in growth occurred before new towns were designated.

6. The results in this table are similar to the results of the 1982 census, which showed that 56 percent of the population in cities and 53 percent of that in towns were employed (State Statistical Bureau 1985a, tables 40, 58).

7. State Statistical Bureau 1984, pp. 45, 111. By 1984 the proportions in state, collective, and private enterprises were 71, 26, and 3 (State Statistical Bureau 1985b).

8. The comparisons here and below are based on Mazumdar (1976).

9. The low figure, which ignores urban youths still legally assigned to agricultural jobs in the countryside, is from Hu (1982). The higher figure, which includes urban youths who still were in the countryside but were scheduled to return legally to the city in 1979–80, is implicit in *Zhongguo Baikie Nianjian* (1981), pp. 627–28. There is a third and even higher estimate that may erroneously include all urban youths ever consigned to the countryside, regardless of their current residence and job status. This estimate, 20 million unemployed at the start of 1979, comes from a purported speech by Li Xiannian to the National People's Congress (Li 1979).

10. World Bank (1980a); World Bank (1980b); *Zhongguo Baikie Nianjian* (1980); Foreign Broadcast Information Service (FBIS), April 30, 1980, p. 19. "Doctor, any type" includes Chinese herbal doctors and secondary-school-trained equiva-

lents of the Russian *feldsher*, but no "barefoot" paramedics. The data for China are for 1979; data for developing market economies are for about 1978.

11. Evidence on how minimal costs to the user lead to overusage comes not only from interviews but also from Chinese drug use studies and government attempts to contain costs (see Jamison and others 1984, pp. 67, 99).

12. The floor space figure excludes kitchen, lavatory, and public corridors.

## Bibliography

- Berry, Brian J. L., and Frank E. Horton. 1970. *Geographic Perspectives on Urban Systems*. Englewood Cliffs, N.J.: Prentice-Hall.
- Chan, Kam Wing, and Xueqiang Xu. 1985. "Urban Population Growth and Urbanization in China Since 1949." *China Quarterly*, no. 104, pp. 583–613.
- Coale, Ansley J. 1984. *Rapid Population Change in China, 1952–1982*. Committee on Population and Demography, Report 27. Washington, D.C.: National Academy Press.
- "Commentator." 1979. "The Aim of Socialist Production Must Be Really Understood." *People's Daily*, October 20, pp. 1–2.
- Emerson, John Philip. 1965. *Nonagricultural Employment in Mainland China*. U.S. Bureau of the Census, International Population Statistics Reports, Series P-90, no. 21. Washington, D.C.: Government Printing Office.
- Fei, Xiaotong. 1984. *Xiao Chengzhen Da Wenti* [Small Town, Big Problem]. Jiangsu Province Small Town Research Group, Compendium 1. Jiangsu People's Press.
- Goldstein, Sidney. 1985. *Urbanization in China: New Insights from the 1982 Census*, Papers of the East-West Population Institute, no. 93. Honolulu: East-West Center.
- Hackenberg, Robert A. 1980. "New Patterns of Urbanization in Southeast Asia." *Population and Development Review*, vol. 6, pp. 391–420.
- Hu, Mengzhou. 1982. "Solution to Employment Problems." *Beijing Review*, no. 39, p. 23.
- Jamison, Dean T., and others. 1984. *China: The Health Sector*. Washington, D.C.: World Bank.
- Kirkby, R. J. R. 1985. *Urbanization in China*. New York: Columbia University Press.
- Lardy, Nicholas. 1978. *Economic Growth and Distribution in China*. Cambridge: Cambridge University Press.
- Leung, C. K. 1980. *China: Railway Patterns and National Goals*. Department of Geography, University of Chicago. Research Paper 195. Chicago.
- Li, Xiannian. 1979. "The Economy of the Whole Country Is Seriously Out of Balance." *Ming Bao* (Hong Kong), June 14, 1979, p. 4, in FBIS, June 19, 1979, pp. L11–13.
- Linn, Johannes F. 1983. *Cities in the Developing World*. New York: Oxford University Press.
- Lipton, Michael. 1977. *Why Poor People Stay Poor*. Cambridge, Mass.: Harvard University Press.
- Liu, Zhi. 1984. "Ten Contradictions Beijing City's Develop-

- ment Faces." *Shehuixue yu Shehui Diaocha*, no. 1, pp. 23-32.
- Lluch, Constantino, and others. 1977. *Patterns in Household Demand and Saving*. New York: Oxford University Press.
- Mazumdar, Dipak. 1976. "The Urban Informal Sector." *World Development*, vol. 4, pp. 655-79.
- Skinner, G. William. 1977. *The City in Late Imperial China*. Stanford, Calif.: Stanford University Press.
- Solinger, Dorothy J. 1985. "Temporary Residence Certificate Regulations in Wuhan." *China Quarterly*, no. 101, pp. 98-103.
- State Statistical Bureau. 1983. *Ten Percent Sampling Tabulation on the 1982 Population Census*. Population Census Office, State Council, and Department of Population Statistics. Beijing: Chinese Statistical Press.
- . 1984. *Zhongguo Tongji Nianjian* [Chinese Statistical Annual]. Beijing: Chinese Statistical Press.
- . 1985a. *1982 Population Census of China (Results of Computer Tabulation)*. Population Census Office, State Council, and Department of Population Statistics. Beijing: Chinese Statistical Press.
- . 1985b. *Zhongguo Tongji Nianjian* [Chinese Statistical Annual]. Beijing: Chinese Statistical Press.
- "Summary." 1979. Summary of *Guangming Daily* article in *Xinhua*, Beijing, November 17, in FBIS, November 20, 1979, p. L5.
- Ullman, Morris B. 1961. *Cities of Mainland China: 1953 and 1958*, Series P-95, no. 59. Washington, D.C.: U.S. Bureau of the Census.
- United Nations. 1973. *Demographic Yearbook, 1982*. New York.
- . 1974. *Demographic Yearbook, 1973*. New York.
- U.S. Bureau of the Census. 1976. *Statistical Abstract of the U.S., 1976*. Washington, D.C.: Government Printing Office.
- Whyte, Martin K., and William L. Parish. 1984. *Urban Life in China*. Chicago: University of Chicago Press.
- World Bank. 1980a. *World Development Report 1980*. New York: Oxford University Press.
- . 1980b. *World Tables: The Second Edition*. Baltimore, Md.: Johns Hopkins University Press.
- . 1982. *World Development Report 1982*. New York: Oxford University Press.
- Xue, Muqiao. 1981. *China's Socialist Economy*. Beijing: Foreign Languages Press.
- Zhao, Jian. 1984. "Criminal Offenses and Social Policy." *Shehui*, vol. 4, no. 5, pp. 16-18.
- Zhao, Lukan. 1980. "A Probe Into Some Employment Problems." *Beijing Review*, no. 43, p. 20.
- Zhongguo Baike Nianjian*. 1980. [Chinese Encyclopedia Yearbook.] Beijing: Chinese Encyclopedia Press.
- . 1981. Beijing: Chinese Encyclopedia Press.
- . 1984. Beijing: Chinese Encyclopedia Press.
- Zhou, Jin. 1979. "Housing China's 900 Million." *Beijing Review*, no. 48, p. 18.
- Zhou, Shulian and Senmu Lin. 1980. "On the Question of Housing." *People's Daily*. August 5.

# Part III

---

## *Concentration or Decentralization: Evaluation of Policies*

Part III is devoted to the issue of urban concentration and related policy initiatives. Although attention has focused in recent years on decentralization policy, that is by no means the only or major direction of government intervention. This part reviews policies to foster both concentration and deconcentration and brings out the cost of obtaining the desired results. Henderson (chapter 7) discusses the experience with concentration in Brazil, and Lee (chapter 8) and Reif (chapter 9) analyze decentralization policies in Colombia and Korea and in Venezuela, respectively.

### **Impetus to Urban Concentration**

Chapter 7 deals with the spatial concentration of industry and cities in Brazil and presents the essential principles of industrial location theory and their relevance to Brazil. Small urban areas in Brazil conform to theoretical expectations, but the findings suggest that the central government has in the past promoted the concentration of heavy industry in Greater São Paulo and Rio de Janeiro. According to the author political considerations may have contributed to the past bias in favor of the major urban centers, but there was also a belief that such a policy would foster economic development. The policy's desirability would hinge on whether scale economies exist in large metropolitan areas and, if so, whether they are urbanization or localization economies. Urbanization economies—the benefits of city size in itself, such as the size of the total labor market and the presence of communication networks—would justify concentration. Localization economies—efficiencies in labor markets and services specific to an industry and benefits arising from specialization by firms—would not.

A model is presented that incorporates both urbanization and localization economies, and the results of estimation are reported. The model assumes constant returns to scale with a Hicks-neutral shift factor external to the firm. The shift factor contains a variable that reflects localization economies, as measured by own-industry employment, and a variable that reflects urbanization economies, as measured by urban population.

Two estimating equations derived from the theory are tested for three industries in the region. One equation directly evaluates effects on production, where value added per capita is regressed on the capital-labor ratio, urban population, own-industry labor force size, distance from urban area, labor force quality, and average firm size (which is included to test the assumption of constant returns to scale). The second equation regresses the capital-labor ratio on factor prices, distance to urban area, property taxes, own-industry size, labor quality, and firm size, where the variable for own-industry size tests the assumption of Hicks neutrality. The author finds no evidence of significant urbanization economies, but he discovers strong support for localization economies, which tend to fall as a sector's labor force grows. Results also appear to justify the assumptions of constant returns to scale and Hicks neutrality. It is concluded that a relaxation of government-encouraged concentration in Brazil would be beneficial.

### **Locational Choice and Decentralization**

Chapter 8 addresses employment location policies. Deconcentration policies can be intended to modify location patterns within an urban region or at the

national level. The former involve intrametropolitan phenomena, the latter interregional issues. This chapter, which concentrates on intrametropolitan issues, investigates patterns of employment location, introduces an economic apparatus to explain firm location choice, and tests the importance of several policy instruments.

The investigation of employment location describes the spatial distribution of manufacturing, commerce, finance, and services in the Colombian cities of Bogotá and Cali. Evidence is found of industrial decentralization, although not in all sectors. A life-cycle classification of firm development (birth, death, relocation, and stability) reveals the important role of births and deaths in this context. The analysis provides an introduction to contemporary location theory, including developments in stochastic specification. The chapter lists site characteristics which are subject to policy manipulation and reports on statistical tests of their importance for three industries. It concludes with a discussion of a conscious decentralization policy with Seoul, the capital city of the Republic of Korea, as an example.

### Impact of Decentralization Policy

In light of its experience in the 1970s with government intervention in the spatial distribution of manufacturing industry, Venezuela provides an opportunity to examine the effectiveness of a decentralization policy. Chapter 9 asks whether government instruments, including incentives and restrictions, have contributed to the deconcentration of industry in Venezuela. This question leads to more specific questions. Has deconcentration of industry taken place? Is deconcentration a result of the policy instruments? What other factors have in-

fluenced industrial location, and what has been their relative importance?

The hypothesis that less protected industries are more concerned with locational factors is supported by the evidence. More generally, the model attempts to identify factors that affect the location decisions of firms. Alternative explanatory factors—such as changes in the Venezuelan economy and the unintended spatial effects of trade and macroeconomic policies—arose when little evidence was found that government incentives affected a firm's location. In particular, the study shows that entrepreneurs attach a relatively small weight to government financial incentives.

### Bibliography

- Haurin, Donald. 1983. "The Incidence and Effects of Factor Taxes and Subsidies upon Urban Areas." Ohio State University, Columbus. Processed.
- Henderson, J. Vernon. *Economic Theory and the Cities*. New York: Academic Press. 1977.
- Lee, K. S. 1982. "A Model of Intra-Urban Employment Location: An Application to Bogotá, Colombia." *Journal of Urban Economics*, vol. 12, no. 3 (November), pp. 263-79.
- Mills, Edwin S., and Charles Becker. 1983. "Indian City Sizes and City Growth." Princeton University, Princeton, N.J. Processed.
- Murray, Michael P. 1982. "Here, There, Where? A Strategy for Evaluating Industrial Relocation Policies in Korea." Urban Development Discussion Paper 6. World Bank, Washington, D.C.
- Renaud, Bertrand. 1981. *National Urbanization Policy in Developing Countries*. New York: Oxford University Press.
- Scott, Ian. 1982. *Urban and Spatial Development in Mexico*. Baltimore, Md.: Johns Hopkins University Press.

# *The Analysis of Urban Concentration and Decentralization: The Case of Brazil*

J. Vernon Henderson

The spatial patterns of industrial and urban development in Brazil have received considerable attention in recent years. Discussions have been characterized by swings in opinion and policy focus that have ranged from the view that the increasing agglomeration of activities has been desirable to the view that the evolving spatial concentration and congestion have been excessive. The relative emphasis on these views has changed over time and has varied with the overall fortunes of the economy; concern about concentration has often diminished during times of economic difficulty.

The differing opinions draw attention to the need for an economic analysis of the forces behind the trends toward concentration and decentralization. Neither phenomenon may be undesirable in itself, but government policies may have intended and unintended impacts on industrial location, and these impacts may affect economic welfare. This chapter examines concentration as it is influenced by government policies and attempts to evaluate the desirability of such influences when the existence (or absence) of certain types of economies of scale in urbanization and industrialization is taken into account. The data base for the core of the paper relates to 1970, and much of the discussion therefore concerns the historical evolution of policies which essentially promoted concentration. In the 1970s distinct efforts were made to achieve decentralization, and some aspects of those efforts are covered in part IV. An up-to-date evaluation of the net outcome of efforts to

promote concentration and decentralization is not the purpose of this chapter.

## Urban and Regional Patterns

The size and growth rates of the largest metropolitan areas, in particular the Greater São Paulo Metropolitan Area and, to some extent, Rio de Janeiro, have been matters of concern in Brazil. The annual growth rates of both urban centers have slowed: for 1970–80 Greater São Paulo had a growth rate of 4.4 percent—the same as the national average growth rate for urban areas—and Rio's rate was considerably less, 2.6 percent. Policymakers, however, continue to view the sizes and growth rates of these cities as excessive. This concern is expressed in a medium-size-cities program which is designed to attract rural migrants into medium-size rather than large cities.

Although Greater São Paulo, with 12 million people in 1980, and Rio, with 8 million, are large by international standards, they are not inordinately big metropolitan areas for a country as large as Brazil, particularly in view of the falling or stagnant growth rates of the two centers. Furthermore, the size distribution of urban areas in Brazil is similar to that of such other large decentralized countries as the United States in terms of the proportion of the total population living in big and small urban areas. Small and medium-size cities are growing at least as fast as larger cities, and there are many other flourishing metropolitan areas in Brazil in addition to Greater São Paulo and Rio.

Although the general size distribution of cities in Brazil may be reasonable by world standards, the spatial

*Note:* The funding of this research by the World Bank is gratefully acknowledged. The encouragement and support of Douglas Keare and Andrew Hamer was much appreciated. Tri Pham, as well as William Dillinger and David Keen, provided able research assistance. Responsibility for all errors is mine.

distribution of resources between large and small cities may be less than desirable. In particular, Greater São Paulo's industrial composition is unusually concentrated in relation to that of other large diversified economies. Greater São Paulo has a great concentration of heavy industry within its boundaries compared with London, New York, Los Angeles, and Tokyo, which, in comparison, are oriented toward services, high technology, and light manufacturing. In 1970 Greater São Paulo accounted for 43 percent of value added for all Brazil in iron and steel and fabricated metals, 71 percent in transport equipment, and 35 percent in chemicals. Such concentration deserves to be explained, and its implications for the efficiency of spacial allocation of resources need to be analyzed.

Greater São Paulo and Rio are located in the South and Southeast of Brazil, a region which comprises the states of Minas Gerais, Espírito Santo, Rio de Janeiro, São Paulo, Santa Catarina, and Rio Grande do Sul. It is the most developed region in Brazil and is larger than Spain, France, the Federal Republic of Germany, and the Democratic German Republic combined. Our sample consists of the 126 urban areas in this region which had populations over 20,000 in 1970. These urban areas now contain about 40 million people.

Highly developed and industrialized, this region is crisscrossed by major modern highways and rails, and its cities offer the basic range of modern utilities and services. Most of the cities are located either in fertile agricultural areas or in areas rich in mineral deposits. The region accounts for almost all of Brazil's manufacturing, and its standard of living is high for a developing country, despite notable deficiencies.<sup>1</sup> The principal shortcoming is in education—in the quality of the current schooling system and in the educational attainment of adults. In 1970 less than 25 percent of adults in the 126 urban areas had more than primary school education (four to five years).

The region produces a full range of manufactured products. Its large metropolitan areas are highly diversified and produce a wide range of goods and services, while smaller urban areas (under 250,000 population) are highly specialized. Specialization—where, say, 10 percent of local employment is in only one industry—occurs in such activities as production of steel, textiles, apparel, pulp and paper, chemicals, transport equipment, and machinery, in food processing, and in services (export activities in ports and retailing and wholesaling in towns that serve as traditional agricultural service centers). Almost half of the 126 urban areas are relatively specialized in the production of one manufactured product.

Brazil is partially a state capitalist economy. Government-owned banks control investment of most domestic

savings except for the investments of traditional plantation families. The government owns almost all public utilities, 50 percent of the iron and steel industry, and 35 percent of the chemical industry, including probably most petrochemicals. The government uses its ownership to influence the location of economic activity and the spatial and intersectoral allocation of capital. In addition, tariffs and taxes have indirectly influenced location and urban concentration.

Brazil has a federal system. Because of changing institutional arrangements it is difficult to determine the degree of local autonomy at any one time. In 1970 some funding for urban infrastructure investments such as municipal roads, water, and sewers came from national earmarked taxes. Much of the funding for both capital and operating expenditures, however, was from local property taxes and user charges, and localities determined their own levels of investment and maintenance. (Except for the state of São Paulo today, state agencies have taken over much of the municipal decisionmaking, but this was not so for the period of our data.) Electrical hookups were generally available in the 126 urban areas, although the quality of service as indicated by brown-outs and blackouts apparently varied considerably. Municipalities controlled their own planning, zoning, and parks but not the educational system or police and fire protection.

Thus, in 1970 localities seemed to have some degree of autonomy in determining their own physical appearance and structure, so that urban areas could compete on an equal footing for resources. Since city councils (although not mayors) were locally elected, local preferences for urban structure may have been represented and to some extent realized. It should be added, however, that discriminatory grants were made to localities, and the ability of local officials to get grants and good utility service was important.

### Government Influence on Industrial Location

Under a free market situation in which localities are fiscally autonomous, certain broad production patterns for cities can normally be expected to emerge (Henderson 1982). Smaller cities serve agricultural areas, act as administrative centers (state capitals), or engage in manufacturing where processing of raw materials or use of heavy processed materials is involved, such as in production of iron and steel, transport equipment, textiles, and pulp and paper and in food processing. Manufacturing centers tend to specialize and to be located near the materials suppliers, and thus, like the natural resources they use, they are spatially dispersed. Smaller urban areas in Brazil do conform to this pattern.

Large metropolitan areas may be expected to support clusters of footloose industries—service, high-technology, and light manufacturing activities—which find it advantageous to locate in large urban market areas and are not tied to the locations of natural resources. But Brazilian metropolitan areas, in particular Greater São Paulo, do not conform to this pattern. The reasons for the difference, and its implications for efficiency of resource usage, need to be explored.

Through its control of utilities, banking, and much steel production, the national government can strongly influence and to a considerable extent even determine the location of heavy manufacturing. It appears that the government has so acted and that, within the region under study, it is strongly biased toward locating investment in Greater São Paulo and Rio or on the Greater São Paulo–Rio axis. Since the government only recently started to evaluate the overall impact of its past decisions, it is difficult to prove the above contention, but several pieces of informal evidence, some relating to the past, others to the present, indicate a governmental role in promoting concentration in the past.

First, in a number of well-known cases locational decisions were made solely on the basis of governmental decisions—for example, the location of the steel works in Volta-Redonda and of major power lines and transformers. Second, at least in the period under review, the government appears to have opposed policies that would lead to decentralization. The pollution control regulations that would force some decentralization from Greater São Paulo have only recently received strong government support. The medium-size-cities program to encourage decentralization has never been effectively implemented. Third, by offering a high real rate of return on deposits, the government banks attract a high proportion of national savings (excluding perhaps those of the plantation families). Apart from agricultural activities, banks appear to be restricted to making long-term loans only to very large firms (those with more than 1,000 employees). Applications for these loans can only be made and pursued in state capitals such as São Paulo and Rio. There is effectively no stock market. Other long-term investments come primarily from the activities of international corporations and the large plantation families.

If the above appraisal of the spatial bias of government decisions concerning industrial location is correct, there remains the question of why the government was biased in favor of Greater São Paulo. Political considerations may have been one relevant factor, but there was also a widespread belief among the authorities and their advisers that Greater São Paulo provided an economic environment that was essential to industrial development and efficiency. For example, in the 1970, Brazil's

balance of payments problems caused by oil price increases were stemmed by the rapid growth of manufactured exports. The abandonment of pollution control policies and of the medium-size-cities programs appeared to have occurred in the past partly out of fear that decentralization of industry out of Greater São Paulo might lead to failure of the industries and stunt the growth of manufacturing. The same was true of longer-term goals of industrial and economic development. The strong belief in the critical role of the economic environment of Greater São Paulo persisted in spite of the success of private manufacturing enterprises in cities other than Greater São Paulo and Rio.

That belief essentially rested on the premise that metropolitan areas offer strong economies of urbanization. Economic theory holds that the agglomeration of economic activity in limited spatial areas occurs because of economies of scale in production. At the industry level there are urbanization and localization economies. Urbanization economies have to do with the scale of economic activity in cities in relation to efficiencies in urban labor markets, interindustry communication, and specialized support services and urban infrastructure. Localization economies relate to scale within an industry as a result of efficiencies in labor markets and services specific to that industry and a greater degree of interfirm specialization in detailed activity within the industry.

If the economies obtained are localization economies there is, from the production point of view, no obvious reason for encouraging the location of different unrelated industries in the same urban area, and so doing is inefficient on the consumption side (Henderson 1982). Specialization therefore occurs in smaller urban areas. Of course, if one industry uses inputs from another industry which are expensive to transport (as in the case for automobiles and steel), they will locate close together, although not necessarily in the same urban area.

Economies of urbanization do provide a basis for locating different industries together to enhance each other's level of efficiency.<sup>2</sup> The existence of urbanization economies as a result of the creation in Greater São Paulo of a large-scale economic environment for all kinds of manufacturing could then be grounds for promoting spatial concentration. To evaluate this issue we need to test whether the scale economies of the heavy manufacturing industries in Brazil pertain to localization or urbanization. If there are no economies of urbanization, but only economies of localization, it would appear that encouraging or forcing manufacturers to pay the relatively high wages and land rents in Greater São Paulo would yield no gain in efficiency. Concentration would be an error because urbanization economies are the rationale for encouraging, if not forcing, a con-

centration of different heavy industries in one metropolitan area. Furthermore, other considerations weigh against such concentration, as elaborated in "Conclusions and Policy Implications," below.

### Sources and Magnitudes of Economies of Scale

To examine the nature of economies of scale, production functions and factor ratio equations for different industries have been estimated for the region under consideration. For each industry the unit of observation is an urban area. The production function for value added is written as

$$(7-1) \quad X = g(S) X(K, L)$$

where  $X(\cdot)$  represents firms' own constant returns to scale (CRS) technology,  $K$  and  $L$  are inputs of capital and labor, respectively, and  $g(S)$  is a Hicks-neutral shift factor, external to the firm, whose arguments are scale measures defined at the urban area level. Since  $X(\cdot)$  is CRS, on the assumption that all firms in the city face the same technology, one can aggregate over firms and use observations on the industry and urban area levels. The external economy formulation and CRS are consistent with the assumption of perfect competition (Chipman 1970). The assumptions of Hicks neutrality and CRS to the firm have been tested, and the results are reported below.

Assuming CRS for  $X(\cdot)$ , equation 7-1 may be written

$$(7-2) \quad X/L = g(S) \tilde{X}(k)$$

where  $k = K/L$ . Taking logs, defining  $\log \tilde{X}(k) = f(\log k)$ , and performing a Taylor series expansion of  $f(\cdot)$  about  $k = 1$  yields a translog specification for equation 7-2:

$$(7-3) \quad \log(X/L) = C_0 + \log g(S) + f'(0) \ln(k) + \frac{1}{2} f''(0) [\ln(k)]^2.$$

The arguments in  $g(S)$  relate to measures of localization economies and of urbanization economies. The localization economies are measured by levels of own-industry employment in the urban area; the urbanization economies are represented either by the urban population or by the total employment in the urban area. Among the various functional forms tried, the best one, defined informally by such considerations as low multicollinearity and high adjusted  $R^2$ s, was

$$(7-4) \quad g(S) = \exp(\alpha/L) + \beta \log P$$

where

$$\epsilon = \frac{d \log(X)}{d \log L} = \frac{\alpha}{L}.$$

The value  $\epsilon$  is the elasticity of firm output with respect to industry employment, holding firm inputs fixed;  $L$  is own-industry employment;  $P$  is urban population; and  $\beta$  is the elasticity with respect to population. The specification that  $\epsilon$  declines as  $L$  rises is supported by a quadratic formulation. The interpretation of the  $\epsilon$  or  $\beta$  elasticities is that a 10 percent increase in own-industry employment or in city population, respectively, leads to a  $0.1 \cdot \epsilon$  or a  $0.1 \cdot \beta$  percent increase in output for any firm, if the firm's own inputs are held constant. That is, for the same inputs, a firm's output goes up because the external shift factor,  $g(S)$ , is larger.

Three other problems need to be addressed in the estimation of equation 7-3. First, both technology and labor force may vary among cities. These items are controlled by measures of labor force quality,  $Q$ , specific to an industry in an urban area. Measures of educational attainment, such as average years of schooling or percent of the labor force with three or fewer years of schooling in that industry in that urban area, and measures of experience, such as average age, were tested.

Second,  $X(\cdot)$  may not be homogeneous of degree one. To test for degrees of homogeneity different from one a measure of average firm size in that industry in that urban area was used. The measure was the average number of employees per firm,  $FS$ .

Third,  $X$  should be a measure of physical output, but in our formulation the variable is expressed in terms of value. The procedure would be satisfactory if output prices were invariant among urban areas, but in reality they are not. We therefore hypothesized that received price declines with distance  $D$  from the nearest major coastal metropolitan area. Where  $X^*$  is the measure of value added and  $X$  is the true quantity equivalent, it is hypothesized that

$$X = X^* p_0^{-1} (1 - tD)^{-1}$$

where  $p_0$  is price in coastal metro areas,  $t$  is unit transport costs, and  $D$  is distance in kilometers. On this basis it may be approximated that

$$(7-5) \quad \log X^* = \log X + \log p_0 - tD.$$

Using the above result and combining equations 7-3, 7-4, and 7-5, a final estimating equation is obtained:

$$(7-6) \quad \log(X^*/L) = C_1 + b_1 \log(k) + b_2 (\log k)^2 + b_3 FS + b_4 Q + b_5 D + b_6 (1/L) + b_7 \log P.$$

(+)
(-)
(0)
(+)
(-)
(-)
(+)

Expected signs are noted below the coefficients. For the  $\log(k)$  terms, the sign restrictions must hold in the

neighborhood of  $k = 1$  to have positive and diminishing marginal products.

The remaining question is whether scale effects are Hicks neutral. Combination of marginal productivity conditions on the basis of equation 7-1 yields the general equation

$$(7-7) \quad k = k(\omega, \mu, S)$$

where  $\omega$  is gross per employee costs (wages plus benefits) in an industry in an urban area and  $\mu$  is the cost of capital. It is assumed that either the pretax cost of capital is the same everywhere or that it increases with distance  $D$  from the nearest major coastal metropolitan area (which will also be a state capital and the administrative center for long-term bank loans). The posttax cost of capital varies with the effective local property tax,  $t_p$ , on equipment and structures in that industry in that urban area.

If scale effects are Hicks neutral,  $S$  should in fact not appear in equation 7-7. A measure of own-industry size was inserted to test whether this is the case. A significant positive or negative coefficient would indicate non-neutrality.

Finally, when firm size and labor force quality are controlled, the estimated form of equation 7-7 is

$$(7-8) \quad \log k = C_2 + a_1 \log \omega + a_2 D + a_3 t_p + a_4 \log L + a_5 Q + a_6 FS$$

(+)                      (-)                      (-)                      (0)                      (+/0)                      (0)

Expected signs are noted below the coefficients.

Equations 7-6 and 7-8 were estimated by ordinary least squares for three major two-digit industries located in Greater São Paulo, namely, iron and steel, transport equipment, and chemicals. The variables are defined as follows:

- $X$  Value added (value of production less total materials costs less production taxes). Production tax rates vary spatially, and their differences may not be passed on to consumers. The inclusion of production taxes in value added has a minimal impact on the results.
- $L$  Average monthly number of employees minus (a trivial number) owners and directors. Information on hours worked is not collected.
- $\omega$  Total salaries less payments to owners and directors plus firm contributions to social security, private insurance, and pension programs, all divided by  $L$
- $K$  Market value of capital stock. Census question

asks what the firm could sell its equipment, structures, and land for today. (Other questions ask book value and depreciated book value.)

- $FS$  Average firm size:  $L$  divided by number of firms
- $t_p$  Property tax rate: industry property tax payments divided by  $K$ . This varies by industry and urban area according to exemptions granted.
- $Q$  Percentage of labor force with three or fewer years of schooling, by two-digit industry, calculated directly from 25-percent long-form sample of 1970 Demographic Census
- $D$  Distance in kilometers to nearest coastal port. For all six ports the urban area is a major metropolitan area. There is only one major interior metropolitan area in the sample, Belo Horizonte. São Paulo is counted as a port although it is 75 kilometers from the sea and the actual port is Santos.
- $P$  Urban population

As a check on the results, a dual form of equation 7-6 was estimated, where

$$\log (X/L) = F(\log w, D, 1/L, P, Q).$$

The scale-economy results for the dual form are indistinguishable from those for the primary form.

The sample size for the industries involved was rather small, and in the final results presented here some insignificant variables are omitted. The results for equation 7-6 are given in tables 7-1 and 7-2.

### Economies of Scale

Table 7-1 provides essentially no evidence of urbanization economies, but there is strong evidence of signifi-

Table 7-1. Production Function: Results by Industry

Variable	Steel	Transport equipment	Chemicals
$1/L$	-109.652 (2.01)	-68.873 (1.95)	-119.224 (2.11)
$\log(P)$	-0.019 (0.24)	-0.003 (0.04)	0.091 (1.06)
Firm size	0.108 (0.65)	0.166 (1.78)	-0.155 (1.09)
Education (percent low-level)	-0.390 (0.41)	-0.310 (0.52)	-0.266 (0.48)
$\log(k)$	0.443 (3.51)	-0.024 (0.23)	0.160 (1.37)
Distance to port	0.082 (1.87)	0.013 (0.33)	0.040 (0.92)
Constant	1.677	2.654	2.854
Adjusted $R^2$	0.54	0.57	0.50
Number of cases	36	27	28

Note: The production function is  $\log (X^*/L)$ .

Table 7-2. *Own-Industry Scale Effects*

Variable	Steel	Transport equipment	Chemicals
$\epsilon$ (median)	0.129	0.153	0.264
$\epsilon$ (2,000 employees)	0.055	0.034	0.060
$\epsilon$ (5,000 employees)	0.022	0.014	0.024

Note:  $\epsilon$ , elasticity of firm output with respect to industry employment.

cant localization economies. Table 7-2 shows that localization economies tend to peter out by the upper tails of employment (in our samples, 5,000 employees). These results indicate that probably little benefit is to be gained by encouraging the location of these different industries in the same urban area to obtain scale economies. There appear to be strong but diminishing benefits from agglomerating a particular industry's employment in the same urban area.

### Other Factors

The  $(\log k)^2$  term is omitted in the final results because a quadratic effect could not be isolated.  $\log k$  is generally significant with the expected sign. The labor force quality measures had the correct signs but were disappointing in their lack of statistical significance, given their considerable variability among urban areas and the fine detail of the data. It could only be concluded that labor force quality is probably not a critical factor in these industries. Firm size performed in no consistent fashion and was not significant.

Distance to the nearest coastal major metropolitan area was either insignificant or had an incorrect sign. This indicates either that output prices are essentially spatially invariant or that this effect is offset by an effect in which efficiency increases with distance from the nearest coastal major metropolitan area. For example, for steel the extent of government control over and influence on production undoubtedly declines with distance from the nearest state capital and hence efficiency may increase with distance.

### Scale Biases

For steel and transport equipment, scale effects are clearly unbiased (table 7-3). For chemicals  $\log L$  has a positive coefficient which is weakly significant. Even that may represent not a scale bias effect but problems in aggregating into the two-digit level the capital-intensive petrochemical industry, with its high levels of employment per urban area, and the traditionally spatially dispersed chemical sector. Examination of the latter sector showed no evidence of scale biases.

Wage levels have their expected strong impact on use of capital. The wage coefficient can be interpreted as the

Table 7-3. *The Capital-Labor (K/L) Ratio*

Variable	Steel	Transport equipment	Chemicals
$\log(\omega)$	1.230 (3.09)	0.482 (1.17)	0.703 (1.72)
Property tax rate	-0.934 (1.54)	-19.115 (4.07)	-5.098 (1.69)
$\log(L)$	-0.031 (0.28)	0.013 (0.13)	0.123 (0.97)
Firm size	0.345 (2.02)	0.079 (0.49)	-0.016 (0.10)
Constant	-0.897	1.845	1.389
Adjusted $R^2$	0.43	0.59	0.40

Note: The  $K/L$  ratio is  $\ln K$ .

elasticity of substitution in production. The numbers for transport and chemicals are in the range of normally accepted values; for steel they are on the high side.

Property tax rates have unexpectedly strong impacts on use of capital, an indication that the tax significantly distorts investment decisions. The elasticities for the industries, evaluated at average tax rates, are -0.04 for steel, -0.66 for transport, and -0.10 for chemicals. Evaluated at the average opportunity cost of capital (assuming a pretax cost of 0.12), the elasticities are -0.15 for steel, -2.95 for transport, and -0.70 for chemicals.

Firm size had no impact except in steel. The results for steel may represent a state capitalism phenomenon, whereby state-owned firms are larger and have access to subsidized capital. Distance to the nearest port had the incorrect sign.

The above results indicate the absence of urbanization economies and the existence of strong localization economies. For steel, which is 50 percent state owned, there is a suggestion of general inefficiencies and distortions in use of capital. All this points to a gain from relaxation of state control over industrial location and production decisions and thus a gain from permitting decentralization.

### Environmental Considerations in Decentralization

Market forces in countries such as the United States, Canada, and the United Kingdom have tended to promote the location of the heaviest polluters in smaller cities, away from the largest and most densely populated urban centers. In Brazil the worst industrial polluters remain in the largest population centers. Although air quality ought not to be the sole consideration in industrial location decisions, São Paulo's air quality is abysmal compared with that of large metropolitan areas elsewhere; residents of the area view air quality as a

**Table 7-4. Air Pollution in Greater São Paulo, Compared with U.S. Metropolitan Area**  
(micrograms per cubic meter, unless otherwise indicated)

Metropolitan area	Number of stations	Sulfur dioxide			Particulates		
		Worst station	Median station	Days minimum standard exceeded <sup>a</sup>	Worst station	Median station	Days minimum standard exceeded <sup>b</sup>
Los Angeles County	13	53	32	0	164	100	3
New York City	14	115	69	8	82	63	1
Chicago (Cook County)	44	74	28	1	201	85	53
Greater São Paulo	13	n.a.	143 <sup>c</sup>	17	n.a.	115 <sup>c</sup>	121

a. Minimum standard is 365 micrograms per cubic meter.

b. Minimum standard is 260 micrograms per cubic meter.

c. Average.

Sources: United States: Environmental Protection Agency, 1977 data; Brazil: Thomas, Comune, and Rizzieri (1980), 1978 data.

prime urban problem (see chapter 13). Table 7-4 presents comparable data for Los Angeles, New York, Chicago, and Greater São Paulo, with the emphasis on sulfur dioxide and particulates, which are industrial pollutants. The average value for sulfur dioxide for the thirteen stations in Greater São Paulo exceeded by far the value at the worst reporting station among U.S. cities. It would seem that encouragement of decentralization of industry outside the Greater São Paulo valley would have important benefits.

### Conclusions and Policy Implications

This brief account of the Brazilian government's influence over industrial location in the past, against the backdrop of an empirical evaluation of scale economies, brings out some important conclusions. Localization economies are strong, and therefore agglomeration of firms into specialized cities to take advantage of such benefits as efficiencies in labor markets and in services specific to an industry and greater interfirm specialization within an industry is advantageous.

The results presented here do not show any significant urbanization economies at the scale of activities prevalent in the urban centers of the South and Southeast in 1970. The rationale for efforts to encourage industrialization of the largest urban areas rests on the putative net benefits for heavy industries from locating in areas with a large general scale of economic activity. The findings of this study do not, however, support this rationale. Rather, they indicate that efforts to limit or counter decentralization initiatives are not desirable.

In addition, negative externalities in the form of environmental degradation could constitute grounds for promoting some degree of decentralization. As noted earlier, however, the size distribution of Brazilian cities is by no means excessively skewed, and efforts to bring

about decentralization or a different distribution of city sizes for its own sake may not be justified. The provision of more uniform incentives to middle-size cities—which could simply mean the elimination of any special incentives, direct or indirect, to larger cities in the southern region—coupled with environmental restrictions in the highly damaged and builtup areas, could lead to some economically beneficial decentralization of activities. A more detailed review of environmental policies is provided in chapter 13. The implications of various governmental policies—direct and indirect, intended and unintended—for industrial location and spatial configuration of overall economic activities are further analyzed in chapters 10 through 14.

### Notes

1. I estimate that the average factory worker in this region takes home at least \$250 a month (1980), plus a full set of fringe benefits. In comparing this estimate with the United States, it should be realized that the cost of nontraded goods is much less in Brazil than in the United States.

2. In fact urbanization economies that differ across industries can also lead to specialization among cities because the optimal size of city for each industry differs.

### Bibliography

- Chipman, J. S. 1970. "External Economies of Scale and Competitive Equilibrium." *Quarterly Journal of Economics*, vol. 84, no. 3 (August), pp. 347–85.
- Henderson, J. V. 1982. "The Impact of Government Policies on Urban Concentration." *Journal of Urban Economics*, vol. 12, no. 3 (November), pp. 280–303.
- Thomas, V., A. E. Comune, and J. Rizzieri. 1980. "Control of Industrial Air Pollution in São Paulo: Evaluating Costs, Benefits and Spatial Effects." University of São Paulo. Processed.

## *Employment Location and Spatial Policies: Colombia and Korea*

Kyu Sik Lee

In the next two decades the urban population in developing countries is likely to continue to grow about four times as fast as in the industrial countries. Between 1975 and 2000 the urban areas of developing countries are expected to absorb about 1 billion people. In the mid-1970s the net annual addition to the populations of Mexico City and São Paulo, for example, was over a half million each; the number was over a quarter million each in Jakarta and Seoul. By 2000 the developing world will have forty cities with 5 million or more people; twenty are expected to have more than 10 million people. Until 1950 Buenos Aires was the only city in the developing world with a population of over 5 million (World Bank 1979, p. 72).

The rapid urbanization in these countries has produced a heavy concentration of population and economic activity in a very few large urban centers. This pattern of urban concentration has generated two main policy concerns. First, it is widely believed that the largest cities in these countries are getting "too big." This belief is not usually based on evidence that negative externalities such as pollution and congestion are greater than the benefits of agglomeration economies. It is more likely that the concerns about the size of the large cities stem from the decline in the quality of life of their high-income groups, from the frustrations of planners who have experienced enormous pressure in recent years to accommodate rapid urban growth, and from a fear that large cities may experience catastrophic failure

of management. The second and perhaps more easily understood policy concern focuses on regional equity, since in most countries it is politically and socially important to maintain a balance among regions in income, opportunities for education and employment, and urban amenities. Concerns about regional equity also often underlie programs and policies that attempt to redirect population growth from large cities.

In the middle-income countries of Latin America and East Asia, including Mexico, Brazil, Venezuela, the Philippines, and Korea, policymakers have paid a good deal of attention to various ways of reducing spatial biases toward large cities. More recently, the governments of those countries have initiated specific policies to decentralize economic activity away from the largest urban center to peripheral areas or secondary cities. Policy packages using diverse instruments have been initiated and implemented with varying degrees of success (Renaud 1981; Townroe 1979). Industrial location policies tend to be the most important of the decentralization policies. This is not surprising, since "the true determinants of urbanization and spatial concentration in developing countries are found in the forces that determine the location of employment opportunities: the nature and pattern of industrialization, the pace of agricultural development, and the growth of transportation and communication networks" (World Bank 1979, p. 76). More specifically, the decentralization policies include explicit policy instruments that are intended to relocate existing industries from the large urban centers or to induce new industries to new industrial zones, secondary cities, or lagging regions.

The rationale for policies to decentralize economic activity is difficult to justify solely on economic grounds

*Note:* This chapter has been published in somewhat different form as "Decentralization Trends of Employment Location and Spatial Policies in *Loc Cities*," *Urban Studies*, vol. 22 (April 1985), pp. 151-62. It has been edited slightly.

because these policies are an outcome of diverse social, political, and economic objectives: interregional equity, political cohesion, national defense, and prevention of a catastrophic failure of urban management. Because decentralization policies are deemed plausible responses to this complex set of policy objectives, they are likely to be pursued in the middle-income countries, and it is extremely important to help governments select policies that are least damaging to the overall welfare of the economy.

For analytical reasons it is convenient to address employment location policies at two levels of spatial aggregation: deconcentration policies that are intended to modify employment location patterns within the capital region and decentralization policies that are intended to influence spatial patterns at the national level. The former involves a study of intrametropolitan phenomena; the latter deals with interregional issues, including the development of secondary cities and lagging regions. This chapter is confined to intrametropolitan issues.

To formulate sound spatial policies and to implement plans and programs efficiently, it is essential for policy-makers to understand the trend of employment location patterns and how firms' location behavior tends to produce a trend. In the Bogotá City Study conducted by the World Bank, analytical and empirical results were established regarding firms' location decisions, but a study of policy effects was not undertaken because explicit policy instruments were not implemented in Bogotá. Measurement of such policy effects is the focus of World Bank research on employment location policies in Korea, a country with a long history of policy experiments. This chapter reports the empirical findings from the Bogotá City Study, which had a heavy behavioral focus, and draws some policy conclusions from the Korean experience with spatial policy.

### Observed Patterns of Employment Location

This section documents the employment location patterns of four major industry groups—manufacturing, commerce, finance, and services—which together account for about 85 percent of total employment in Bogotá and Cali.<sup>1</sup> Three main data sets are used: the 1978 household survey for both cities, the 1972 household survey for Bogotá, and the social security establishment files for Cali (1976) and Bogotá (1978). Ideally the data for the study of employment location should be based on the establishment. Although the social security files are such data sets, their coverage of small firms is poor. The 1978 household survey, which was jointly prepared by the World Bank and the Colombian National Statistics Department (DANE), had a module for obtaining informa-

tion on firms where the respondents worked. The questions covered the location of the workplace, the number of workers in the firm, the initial year of operation at the present location, the previous location (if applicable), and the type of business.

Making inferences about the characteristics of firms where the household members work on the basis of a household survey requires some justification. The household survey results can be compared ex post with the social security data regarding firm characteristics important for this study, because the household survey questionnaire asked the respondent whether his firm was affiliated with the social security system. The distributions given by the household survey regarding the firms affiliated with the social security system can therefore be compared with the corresponding distributions from the social security data. In terms of the percentage distribution of several firm characteristics—size, type of industry, and location—the two data sets are strikingly similar, which increases our confidence in inferences drawn from the household survey. Most of the analyses in this section are thus based on the information from the 1978 household survey, which covers all workers regardless of the size of the firms in which they work.

Employment shares of major sectors in Bogotá and Cali are similar to those of other cities of the world. As shown in table 8-1, the share of manufacturing employ-

Table 8-1. *Employment Share of Major Sectors in Selected Cities*  
(percent)

City	Manu- facturing	Commerce	Services <sup>a</sup>
Bogotá, Colombia (1978)	23.6	20.3	40.9
Cali, Colombia (1978)	31.1	21.6	32.6
Seoul, Rep. of Korea (1970)	22.7	28.9	33.0
Kuala Lumpur, Malaysia (1970)	20.5	17.6	35.5
Manila, Philippines (1970)	22.1	13.8	37.3
Abidjan, Côte d'Ivoire (1970)	22.5	18.6	42.8
Tunis, Tunisia (1972)	18.9	17.5	39.2
United States, small cities (1960)	25.1	28.4	28.8
United States, large cities (1960)	30.2	32.6	24.8

a. Includes financial services.

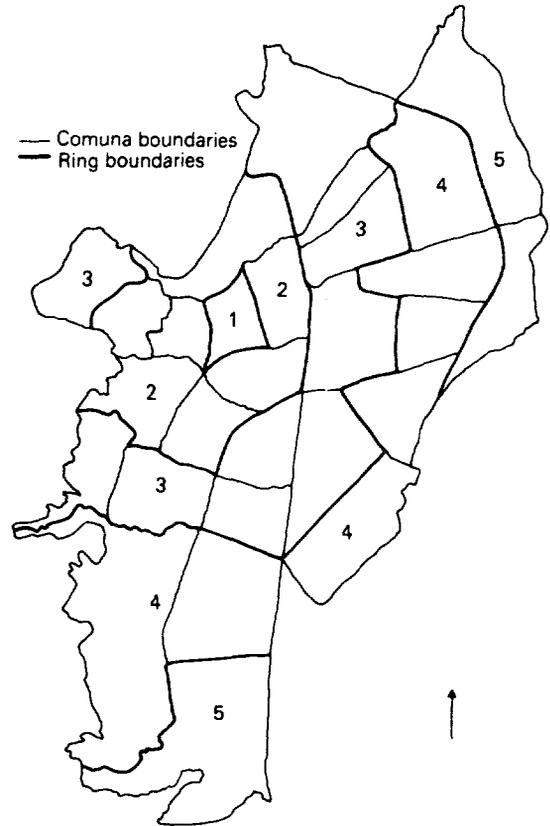
Sources: The values for Bogotá and Cali are based on the 1978 DANE household survey; those for other cities are from Renaud (1981).

ment is remarkably consistent among developing-country cities (from 20 to 25 percent). Bogotá's share of manufacturing employment falls in this range, but Cali's share, 31 percent, is similar to that of large U.S. cities. The employment shares in commerce—about 20 percent for both Bogotá and Cali—are comparable with those of other developing-country cities. Bogotá's share of employment in services (41 percent, including finance) is on the high side and Cali's 33 percent is on the low side, compared with other cities in developing countries. Bogotá's position as the nation's capital and financial center may contribute to its high share of services employment.

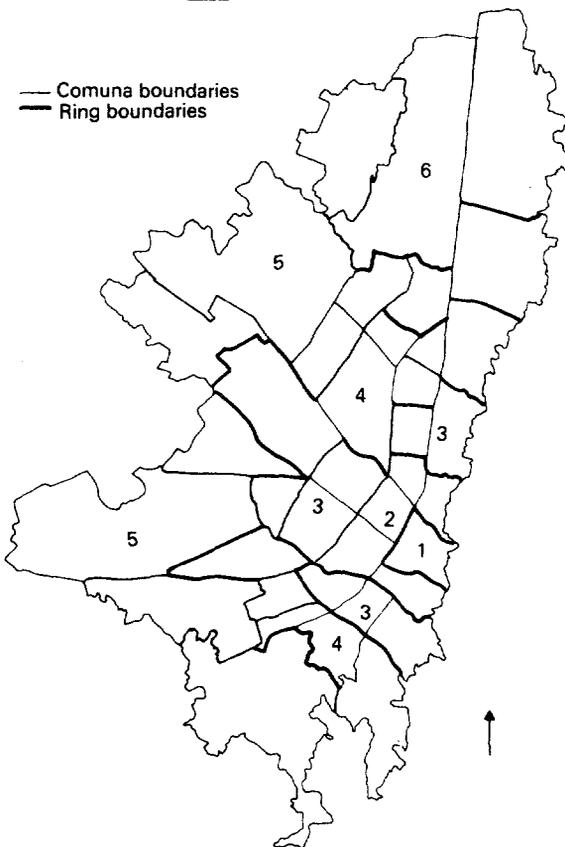
To study the spatial distribution of employment, we have devised a ring system for each city based on *comunas*, the administrative units in Colombian cities. Bogotá has thirty-eight *comunas* and Cali has twenty-eight. By aggregating *comunas* we divided Bogotá into six rings and Cali into five (maps 8-1 and 8-2). The ring system is useful for studying the relation between the spatial distribution of employment and distance from the central business district (CBD).

Table 8-2 shows the spatial distribution of employ-

Map 8-2. Cali: Ring System Based on Comunas



Map 8-1. Bogota: Ring System Based on Comunas



ment by ring for major industry groups in Bogotá and Cali. For all employment, Cali shows a greater central tendency than Bogotá; the peak concentration is in ring 3 in Cali and in ring 5 in Bogotá. In Cali the share of employment drops markedly beyond ring 3. The CBD's share of employment is not much different in the two cities—14 percent in Bogotá and 16 percent in Cali, which is comparable to that of large U.S. cities.<sup>2</sup> The central tendency is, however, substantially different among industry groups. In both cities the concentration of employment in the central area is highest in finance, followed by commerce, services, and manufacturing. For all industry groups, however, the central area in Cali has a much larger share of employment than that of Bogotá.

To examine the changing patterns of employment location, the spatial distribution of employment for at least two points in time must be examined. For the trend analysis for Bogotá two household survey data sets are used: the 1972 household survey conducted for the Bogotá Urban Development Study, Phase II, and the 1978 DANE household survey. Like the 1978 survey, the Phase II survey contained questions about the respondent's workplace, including location, firm size, and type

of business. For Cali the social security data file for 1976 and the 1978 DANE household survey were used.

Table 8-3 reports the spatial distribution of employment in Bogotá by major industry group for 1972 and 1978. We find evidence of decentralization of total employment during the period, but the sharp decline of the CBD's share (from 23 to 14 percent) is suspect. The apparent overstatement was largely attributable to the problem with the definition of the CBD in the 1972 survey. It should be noted that the share of the CBD and ring 2 together decreased only slightly, from 36.6 percent in 1972 to 31.7 percent in 1978.<sup>3</sup>

In Bogotá the extent of employment decentralization in manufacturing is very similar to that in commerce, except that the CBD kept a larger share of employment in

commerce. The service sector also shows a clear trend of employment decentralization. The location pattern of financial establishments is different from those of other sectors in that the central tendency persisted during the six-year period. For this sector, however, it should be noted that ring 2 had a substantial gain in employment share, which reflected the shift of employment to the International Center (see note 3).

Table 8-4 shows evidence of employment decentralization in Cali on the basis of the social security data for 1976 and the 1978 household survey data. As indicated earlier, there is a strong consistency between the social security data and the subset of the 1978 household survey data which represents those individuals whose firms are affiliated with the social security system. Since

**Table 8-2. Employment Distribution, by Ring and Major Industry Group, 1978**  
(percent)

Ring	All employment <sup>a</sup>	Manufacturing	Commerce	Finance	Services
<i>Bogotá</i>					
1 (CBD)	13.95	6.01	15.75	41.43	12.91
2	17.74	13.47	19.77	29.38	18.68
3	16.40	21.54	14.83	11.43	16.87
4	20.60	24.89	19.37	10.88	23.18
5	24.94	28.25	27.72	5.65	21.79
6	3.43	2.19	1.51	0.60	4.41
Location not specified	2.96	3.66	1.05	0.63	2.15
<i>Cali</i>					
1 (CBD)	16.36	9.83	25.23	67.41	14.09
2	26.16	25.71	23.68	19.69	34.57
3	32.67	39.64	31.64	4.70	30.24
4	14.37	11.81	11.84	1.57	15.26
5	2.78	1.38	4.53	0.00	2.12
Location not specified	7.66	11.62	3.08	6.64	3.73

a. Includes other sectors.

Source: DANE household survey, 1978.

**Table 8-3. Changes in Employment Location, Bogotá, 1972-78**  
(percent)

Ring	All employment <sup>a</sup>		Manufacturing		Commerce		Finance		Services	
	1972	1978	1972	1978	1972	1978	1972	1978	1972	1978
1 (CBD)	23.03	13.95	18.20	6.01	19.43	15.75	42.11	41.43	22.62	12.91
2	13.61	17.74	16.07	13.47	12.18	19.77	13.69	29.38	12.74	18.68
3	14.62	16.40	18.94	21.54	13.35	14.83	6.89	11.43	15.88	16.87
4	18.80	20.60	20.27	24.89	21.83	19.37	10.00	10.88	20.74	23.18
5	18.61	24.94	21.76	28.25	21.52	27.72	14.64	5.65	17.89	21.79
6	1.67	3.43	1.04	2.19	3.42	1.51	0.55	0.60	1.88	4.41
Location not specified	9.67	2.96	3.72	3.66	8.27	1.05	12.12	0.63	8.25	2.15

a. Includes other sectors.

Sources: DANE household survey, 1978; Bogotá Urban Development Survey, Phase II, 1972.

**Table 8-4. Changes in Employment Location, Cali, 1976-78**  
(percent)

Ring	All employment <sup>a</sup>		Manufacturing		Commerce		Finance		Services	
	1976	1978	1976	1978	1976	1978	1976	1978	1976 <sup>b</sup>	1978
1 (CBD)	31.51	26.19	20.19	14.94	48.38	54.97	45.68	84.90	29.15	25.34
2	37.12	27.61	34.28	27.02	34.50	24.40	16.67	15.10	55.47	32.19
3	28.40	31.33	41.40	46.27	11.09	15.65	37.55	0	14.79	16.45
4	2.95	13.82	4.09	11.03	6.03	4.97	0.11	0	0.54	24.37
5	0.03	1.05	0.05	00.70	0	0	0	0	0.06	1.65

*Note:* Based on employment in firms with ten or more employees.

a. Includes other sectors.

b. Excludes employment in the government sector.

*Sources:* DANE household survey, 1978; social security data files, 1976.

the social security files' coverage of small firms for all industry groups is poor, to increase the comparability between the two data sets table 8-4 includes only jobs in firms with ten or more employees.

For all industry groups together there is clear evidence of employment decentralization in Cali, as in Bogotá. The results reported in table 8-4 indicate that although Cali experienced employment decentralization in manufacturing and services, there is an increasing central tendency in commerce and finance. The sharp increase in the CBD's share of employment in finance during this period is obviously a result of sampling errors: the 1978 household survey must have under-sampled individuals working in financial firms, especially those located in the outer rings. Nevertheless, it is reasonable to conclude that these results do not support a decentralization trend in these two sectors in Cali.

The trend of employment decentralization having been examined, we now analyze the factors that influence changing patterns of employment location by investigating the location behavior of four groups of firms: firms that relocate (movers), newly established firms (births), defunct firms (deaths), and firms that remain at the same location but expand or contract. Evidence will be presented for the first two categories. The destinations of movers and the location patterns of births are important for understanding the changing patterns of employment location and for predicting the future spatial structure. Here we use the relevant information collected in the 1978 household survey to examine the direction of job movements and the location patterns for births.

The 1978 household survey asked when the firm began operations at its present location and what was the previous location if it had moved. On the basis of this information the origin and destination ratio was calculated for each industry group with the rings as subareas. The ratio was obtained by dividing the number of jobs moving out of an area by the number of jobs moving into that area during 1973-78.<sup>4</sup>

The results reported in table 8-5 indicate that in both Bogotá and Cali the CBD experienced a net loss of jobs in each industry group. The net outflow of jobs from the CBD was much greater in Bogotá than in Cali for each industry group except the service sector. In Cali all rings outside the CBD had net gains of jobs except for service employment in ring 3. Although Cali's ratio values suggest an outflow from the CBD to the other rings, within non-CBD rings the values do not indicate decentralization.

For Bogotá, the results show evidence of decentralization within the non-CBD rings, owing to relocation patterns. In Bogotá's manufacturing the ratio gradually declines as the distance from the CBD increases; rings 4 and 5, for instance, had a net gain in employment. The gain of jobs in ring 2 is also reflected in table 8-5 and indicates the growth of the International Center.

Table 8-6 reports the location patterns of jobs created by firms established during the five-year period (births). The strong regularity in the location patterns of new jobs is striking; in both Bogotá and Cali the employment share of new jobs by ring increases with the distance from the CBD for almost all industry groups except finance. This is an indication that newly established firms tend to locate in outer areas and thus contribute to the decentralization patterns. It is worth noting that in Bogotá's finance sector 43 percent of jobs created by new firms during 1973-78 were located in ring 2, a result that captures the northward shift of financial activities from the CBD to the International Center.

### Behavioral Underpinnings of Location Choice in Manufacturing

The previous section summarized, with the use of various data sets, the location patterns of employment in Bogotá and Cali and their changes over time. The results reveal a high degree of employment location dynamics and strong evidence of the spatial decentralization of

employment in both cities. To understand and explain these changing location patterns, in 1978 we conducted a survey of manufacturing establishments in Bogotá using the DANE industrial directory as the sample base. The survey questionnaire was designed to take no more than one hour to complete and did not require the respondents to refer to their accounting books, but it nevertheless yielded a large amount of information with nearly 300 computer-readable variables. The survey results, which provide the basis for understanding the

changing location patterns of manufacturing employment, were reported in Lee (1982a),<sup>1</sup> and a model of employment location was formulated and estimated using this survey data (Lee 1982b). The remainder of this section is based on Lee (1982b).

In this study the basic premise for modeling employment location is drawn from well-established theoretical and empirical foundations in the literature on housing and residential location. Mills (1972), Solow (1972), and Muth (1969) provided the basic theoretical foundations

**Table 8-5. Moving Patterns of Jobs, Measured by Origin-Destination Ratio for Rings, 1973-78**

Ring	All employment <sup>a</sup>	Manufacturing	Commerce	Finance	Services
<i>Bogotá</i>					
1 (CBD)	2.37	2.51	3.15	3.32	1.45
2	0.66	1.54	0.45	0.33	0.87
3	1.33	1.58	1.32	1.04	1.36
4	0.55	0.76	0.51	0.06	0.58
5	0.44	0.32	0.20	2.70	0.35
6	0.59	— <sup>b</sup>	— <sup>b,c</sup>	— <sup>b,c</sup>	0.29
<i>Cali</i>					
1 (CBD)	1.85	1.97	1.71	1.58	2.03
2	0.68	0.70	0.61	— <sup>c</sup>	0.45
3	0.86	0.93	0.84	— <sup>b,c</sup>	1.21
4	0.67	0.70	— <sup>b</sup>	— <sup>b,c</sup>	0.33
5	3.44 <sup>d</sup>	— <sup>b</sup>	— <sup>c</sup>	— <sup>b,c</sup>	— <sup>b</sup>

*Note:* The origin-destination ratio is found by dividing the number of jobs moving out of a zone by the number of jobs moving into the zone.

a. Includes other sectors.

b. No firms moved into the area.

c. No firms moved out of the area.

d. Even though the ratio was high, ring 5 lost only 651 jobs and gained 189, whereas the CBD lost 9,479 and gained 5,126. The total number of jobs relocated was 28,170 in Cali and 84,755 in Bogotá.

*Source:* DANE household survey, 1978.

**Table 8-6. Location Patterns of New Jobs, 1973-78 (percent)**

Ring	All employment <sup>a</sup>	Manufacturing	Commerce	Finance	Services
<i>Bogotá</i>					
1	13.78	7.55	14.09	29.20	18.61
2	17.80	13.47	20.31	43.04	14.05
3	14.42	19.98	8.15	13.86	17.38
4	20.87	24.66	23.00	9.43	17.55
5	27.67	30.62	31.97	4.47	26.78
6	2.64	2.82	1.72	0.00	2.32
Location not specified	2.82	0.90	0.75	0.00	3.32
<i>Cali</i>					
1	14.70	9.27	13.61	0.00	26.27
2	25.62	19.18	24.05	60.00	30.88
3	38.33	44.95	44.76	20.00	32.99
4	13.40	22.24	11.45	0.00	5.16
5	2.60	2.29	1.97	0.00	0.00
Location not specified	5.35	2.06	4.15	20.00	4.70

a. Includes other sectors.

*Source:* DANE household survey, 1978.

for the economic analysis of residential location and employment location in urban areas. Theoretical and empirical work on housing and residential location has advanced rapidly,<sup>5</sup> but comparable work in the field of employment location has not followed. Mieszkowski and Straszheim (1979, p. xiii) observe that "among the most important issues not covered in this volume is the relative importance of different factors in employment location and decentralization patterns within metropolitan areas, which deserves far more research. Principally because of data problems, little research has been conducted on employment locations or urban labor markets in recent years. Effective use of the censuses of employment is plagued by disclosure rules." Nevertheless, work by Leone (1971), Schmenner (1973), and Kemper (1973), followed by Struyk and James (1975), pushed back the frontier in developing an empirical basis for employment location study. Moreover, recent efforts in actual modeling of employment location and econometric estimation of such models have been promising: Hanushek and Song (1978) developed a framework for analyzing the spatial structure of employment in the Boston metropolitan area, Erickson and Wasylenko (1980) estimated a model of relocating firms in the Milwaukee metropolitan area, and Schmenner (1973, 1982) provided evidence from his econometric work on Cincinnati and New England. Carlton's work (1977, 1979) focused on the intercity location decisions of new firms.

Residential location studies usually assume that a consumer chooses a particular dwelling unit at a particular location at which he maximizes his utility, given the budget constraint. In a similar optimization framework, it is assumed that the firm, as a price taker, locates where it maximizes profits. In making the location decision, the attributes of the particular plant site and the lot size enter into the process of determining the optimum combination of inputs for production. This means that the site attributes enter directly into the firm's production function in the following way:

$$(8-1) \quad Q = f(L, X; Z)$$

where  $Q$  is output,  $L$  is lot size,  $X$  is a vector of variable inputs such as labor and plant and equipment, and  $Z$  is a vector of exogenous site characteristics.

The site characteristics are independent of lot size and represent local public goods available to that particular location. In a recent work on housing Burstein (1980) explicitly introduces local public goods into the consumer's utility function. Ellickson (1981) integrates the theory of residential location choice with that of local public goods by including in the consumer's utility function the attributes of the surrounding neighborhood and the public schools. For studies of firm location, local

public goods include the quality of public utility services such as electricity and water, the quality of municipal services such as police and fire protection, and zonal characteristics such as air quality and congestion.

The relevant cost components in calculating the optimum combination of inputs and hence in selecting a particular location are, then, wages, capital cost, input materials costs, delivery costs of inputs and output, and land rent. Following the standard theory in urban economics, a particular plant site is occupied by the firm which bids the highest price for it. The bid price depends on the attractiveness of the particular site (which in turn depends in part on the availability of local public goods) for a firm of the particular type. In locational equilibrium all firms in an urban area make the same profits and no firm has any incentive to relocate. This equilibrium configuration is an outcome of cost tradeoff calculations by individual firms, given the spatial variations of relevant costs. For example, a large manufacturing firm may choose a site in a low-rent area near the periphery to meet its need for more plant space, although delivery distance will be greater. Small firms may prefer a central location where the high rent is more than offset by externalities available for production and shipment. The high rent in the central area can also be explained by the high intensity of variable input use. The empirically observed rent gradient reflects the increasing marginal productivity of land as the distance to the CBD becomes shorter.

The stochastic specification of the model should provide a framework for predicting the probability that a firm of a particular type will occupy a site with particular attributes ( $Z$ ). Since a given site will be occupied by the firm with the highest bid, the relevant random variable for determining the probability is the maximum bid made within a group of firms with similar attributes. The probability distribution of a random variable associated with the maximum bid leads to a multinomial logit specification for the firm location model.

The application of the multinomial logit method to economic research became popular with McFadden's work on travel demand studies (1973, 1976). Such a model was used to predict an individual's choice of travel mode among a finite number of alternatives—car, bus, taxi, transit, or foot—given the characteristics of the individual or household. Subsequently, the multinomial logit framework was applied to housing and residential location studies by, for example, Friedman (1975), Lerman (1977), and Quigley (1976). These studies applied the travel mode choice model to variations of residential choice models: a utility-maximizing consumer of type  $t$  chooses a house of type  $Z$ , which is analogous to the consumer's choosing a travel mode  $Z$  among several alternatives.

In his recent work Ellickson (1977, 1981) makes an important departure from the above approach and offers an alternative multinomial logit specification for the residential choice model using the bid-rent theory. According to Ellickson (1981, p. 63), "the most natural way to interpret such models is in terms of a prediction of what sort of consumer is most likely to occupy a house with a specified set of characteristics"—the probability that a house with characteristics  $Z$  will be occupied by a household of type  $t$ .<sup>6</sup> Ellickson points out that his method has several advantages, such as specifying endogenously the properties of the disturbance terms (as the Weibull distribution).

Ellickson's approach is particularly relevant for modeling employment location. As discussed above, our problem is to predict the probability that, given a site with particular characteristics ( $Z$ ), a firm of particular type will occupy that site. Such a site may exist at any location in an urban area, and combinations of site characteristics can produce a large number of site variations. The application of the McFadden approach requires a definition of a finite number of locations (or subareas) from which a firm of a particular type is supposed to choose. Defining alternative subareas in that fashion is arbitrary, and the subsequent analysis will tend to be descriptive rather than analytical and to suffer from the idiosyncrasy introduced by the arbitrary definition of alternative locations.

### A Framework for Evaluating Policy Effects

A study of the effects of employment location policies requires a theoretical and empirical framework for testing the effectiveness of policy instruments with respect to the firm's location choice. The model described above provides such a framework. It predicts the probability that a firm of a particular type will occupy a site with particular attributes, and the site attributes are the independent variables that are subject to policy manipulations. The attributes include measures of access to output and input markets, employees' commuting distance, public services, zonal amenities, and external (or scale) economies. This means that the model can predict the incremental effects of alternative policy instruments on the likelihood that a firm will choose a particular site.

Policymakers in developing countries often attempt to relocate industries from the traditional industrial districts of a large city to outer areas or to smaller cities. The government's plans may include developing new industrial towns or estates or expanding existing ones to induce new or relocating firms to settle in a desired area. In all cases implementing such plans and programs

requires the selection of particular types of industries to occupy sites that have particular attributes. Hence, it is important for policymakers to understand the requirements of firms for attaining equilibrium at new locations and to be able to assess the level and costs of government subsidies and infrastructure investment needed to meet such requirements.

The Bogotá study described earlier did not test the effectiveness of explicit policy instruments, partly because such instruments were not implemented in that city. The behavioral underpinnings established in the study, however, provide sufficient clues regarding those policy instruments that are most appropriate for influencing the location choice of particular types of firms and for influencing aggregate location patterns.

The estimation results reported in Lee (1982b) show which independent variables (site attributes) are most important for the location choices of firms of particular types. In table 8-7 the independent variables are ranked by order of importance for each type of firm according to the values of the elasticity coefficients. For small firms in both the textile and the fabricated metal industries, accessibility to local input markets (INPUTBT), the commuting distance for production workers (WKSOUTH), and the population density of the zone of plant location (POPDENS) are important site characteristics. These three variables alone carry more than half of the total coefficient weight in the equation. Accessibility to local product markets (PRODSOLD) was the second most important variable for small fabricated metal firms, but it was not as important for small textile firms.

For large textile establishments the location quotient (LOCQT, a measure of scale economies in the zone of

Table 8-7. *Independent Variables in Firm Location Choice, by Rank, Bogotá*

Variable	Small textile	Small fabricated metal	Large textile
INPUTBT	1	1	4
WKSOUTH	2	4	7
POPDENS	3	3	8
ELECINT	4	8	2
LOCQT	5	7	1
PRODSOLD	6	2	5
ADMNORTH	7	5	3
DISTCBD	8	6	6

Note: INPUTBT, a measure of accessibility to local input markets; WKSOUTH, commuting distance for production workers; POPDENS, population density of the *comuna* where the establishment is located; ELECINT, frequency of electricity interruption; LOCQT, location quotient; PRODSOLD, a measure of accessibility to local product markets; ADMNORTH, commuting distance for administrative workers; DISTCBD, airline distance from the CBD.

Source: Lee (1982b).

plant location), the quality of electricity supply (ELECINT), and the commuting distance for administrative workers (ADMNORTH) are the three most important site attributes; together they carry more than two-thirds of the total weight in the equation. Access to local markets and population density are relatively unimportant for this group of firms.

The results for large textile firms are consistent with the findings that large firms are export-oriented and tend to locate near the periphery, where land needed for modern facilities is available at lower cost and highway access is better than in the central area (Lee 1982a). The evidence on small firms' local market orientation and their tendency to locate in high density areas is consistent with the incubator hypothesis, which was supported by a test in an earlier work (Lee 1981). The hypothesis states that small new manufacturing firms tend to locate in centralized areas that provide essential services for production and sales activities. The specific areas in Bogotá that were identified as incubator areas are not in the industrial district but are adjacent to the CBD.

From this analysis it is apparent that government policies intended to influence employment location patterns can be effective if such policies influence the site attributes which are important to firms. The analysis further suggests that different mixes of policy instruments should be considered for different groups of firms and industries.

### Policy Implications

The strong decentralization of employment in Colombian cities, summarized in "Observed Patterns of Employment Location," above, is comparable to the trend observed for large U.S. cities during the past several decades (Hoover and Vernon 1959; Leone 1971). But the main goal of spatial policy in the United States regarding intrametropolitan decentralization has been quite different from that in developing countries. In the United States the main policy objective has been to reduce urban decentralization in the hope of preventing central city decay. This policy objective is based on the belief that decentralization occurs because of deteriorating conditions in the central city; there is, however, little empirical evidence to support this view. Muth (1969) attributes most decentralization to transport improvements, suburban (new town) development programs, and federal subsidies for homeownership. Muth's findings suggest that federal programs intended to reduce decentralization, such as urban renewal programs that tend to lower housing density in the central city, have expedited the trend rather than reduced it. Attempts to reverse the tide of decentralization, which

occurs because of increased population and incomes, often result in economic inefficiency, since transport improvements have reduced the central city's comparative advantage for production and other economic activities.

Large U.S. cities also have municipal fiscal problems that encourage decentralization. The high per capita tax burden in central cities in relation to the suburbs provides incentives to move to the suburbs. It has been observed that "the lower the average income level of the central city relative to its suburbs, the smaller is the central city's population, and the larger is the land area occupied by the urbanized area" (Muth 1969). Evidence from Bogotá and Cali, however, does not suggest an increasing concentration of low-income population in the central city, and there is no sign of central city decay. In fact, policies to decentralize economic activity from large cities in developing countries stem mainly from the increasing concentration of economic activity in the central city and the perceived problems of congestion and pollution as the city's population grows rapidly. Also, municipal fiscal relationships in developing countries are different from those in the United States. In Korea, for example, electric utility service charges, which are uniform nationwide, tend to subsidize central cities over outlying areas and hence encourage centralization.

The case of Korea affords a chance to look at a conscious decentralization policy. During the past decade various spatial policies to control the growth of Seoul and to disperse its population have been implemented. In 1971 the greenbelt surrounding Seoul was established. The 1977 Industrial Location Act in effect prevented new manufacturing firms from locating within Seoul and enabled the government to issue relocation orders to establishments already set up there. That same year the government initiated a ten-year comprehensive plan for redistribution of population and industry from Seoul. The plan included many policy instruments in five principal categories: the relocation of industries from Seoul; inducements for people to relocate to southern provinces; the decentralization of educational facilities; the relocation of various urban functions within the capital region and the improvement of city plan implementation, particularly in dealing with clandestine housing construction in Seoul; and tax and credit incentive schemes to aid relocating firms. Most of the instruments stipulated in the plan address the location and relocation of manufacturing establishments. A comprehensive review of employment location policies in Korea appears in Choe and Song (1982), and a theoretical analysis of policy efficiency is discussed by Murray (1982). The empirical evaluation of such policies was the main task of a recently completed World Bank project.<sup>7</sup>

Several other recent policy measures in Korea and Colombia are discussed below.

The incubator hypothesis mentioned earlier states that small new manufacturing firms start in central locations that provide needed services and infrastructure and then move to less central locations as they grow and require more space for expansion (Hoover and Vernon 1959). This hypothesis was tested for Bogotá on the basis of industrial directory data (Lee 1981). Data from the establishment survey conducted for the Bogotá City Study do in fact indicate that small firms start up in central locations (Lee 1982a), and the logit results summarized earlier also support this hypothesis (Lee 1982b). There is evidence that Korean firms also follow this pattern (Meyer 1981, Lee 1985). The 1977 Industrial Location Act, which prohibits new manufacturing firms within the city limits, should have restricted such incubation. A related issue is whether incubation can be replicated in new industrial estates or existing towns outside Seoul.

In 1978, as an important measure for decentralizing population and economic activity away from Seoul, the government established a new industrial town, Banweol, less than 30 kilometers distant. Although nearly 1,000 plant sites for small and medium-size firms were prepared, the occupancy rate was less than 20 percent as of 1981. Many firms that moved to Banweol suffered excess capacity and financial losses as a result of overinvestment in plant and land area (induced by the government incentive schemes), increased operating costs after relocation, and the general recession of 1980.

A case study on Banweol (Choe and Song 1982) shows that the most serious problems facing the relocated firms have been reduced accessibility to product markets and input suppliers, the unavailability of production workers, and difficulties in obtaining day-to-day business information because of poor telephone service and limited person-to-person contacts. Poor access to Seoul and Incheon is largely responsible for these problems. A related problem is the reluctance of production workers to relocate to Banweol or to commute from Seoul. Attrition of skilled workers has been high, and it is difficult for firms to replace those who quit.

That such a seemingly short distance thwarted the development of Banweol is striking. The logit analysis used to study the Bogotá data helps explain the Korean experience: small and medium-size firms prefer central locations, and accessibility to local markets and proximity to production workers are the most important site attributes for them.

In contrast to the bias against decentralization in the United States, several developing countries have tried to decentralize economic activity away from the central city. Nevertheless, the desirability of decentralization

policies on economic grounds has not been established, and little is known of their effects or their welfare implications. The key policy question is how to guard against spatial policies that are excessive in relation to prevalent trends, since excessive measures may result in serious welfare losses. The lack of empirical information on decentralization<sup>8</sup> and policy effects in developing countries does not yet permit the formulation of more efficient spatial policies, but policies to decentralize population and economic activity are probably not good substitutes for better internal management of city growth. For example, the effect on air pollution or on traffic congestion of reducing the population or employment in a large city by a certain amount is likely to be small (Henderson 1980; Tolley, Graves, and Gardner 1979).

Notes

1. The material in this section is from Lee (forthcoming).
2. According to Bronitsky and others (1975), about 10 to 15 percent of total employment in large U.S. cities is in the CBD.
3. Moreover, on the basis of the two data sets, the total number of jobs in the CBD together with those in the International Center (*comuna* 81, the area directly north of the CBD) stayed almost constant during the six-year period.

<i>Comuna</i>	1972	1978
31 (CBD)	201,975	166,878
81 (International Center)	24,787	62,565
Total	226,762	229,443

4. Since the number of workers at the previous locations is estimated on the basis of the number of survey respondents, it is implicitly assumed that the number of jobs at the previous locations is the same as that at the 1978 locations. This assumption should not affect the conclusions of this study.
5. Major work in this area includes Kain and Quigley (1975), Straszheim (1974), and Ingram (1977).
6. The mathematical derivation of this result can be found in Lee (1982b).
7. "An Evaluation of Industrial Location Policies for Urban Deconcentration," RPO 672-58, RPO 672-91. See World Bank, 1984, *Abstracts of Current Studies 1984*, Washington, D.C.
8. Lee (1985) shows that in the 1970s Seoul's CBD experienced an annual net loss in manufacturing employment of 7.6 percent, while the periphery of Gyeonggi province, in which Seoul is located, gained employment at a rate of 34 percent.

Bibliography

Bronitsky, L., and others. 1975. *Urban Data Book*. U.S. Department of Transportation, Report DOT-TSC-0S5-75-45-1. Washington, D.C.: Government Printing Office.

Burstein, Nancy R. 1980. "Voluntary Income Clustering and

- the Demand for Housing and Local Public Goods." *Journal of Urban Economics*, vol. 7, no. 2 (March), pp. 175-85.
- Carlton, D. 1977. "Models of New Business Location." University of Chicago, Center for Mathematical Studies in Business and Economics, Report 7756. Chicago.
- . 1979. "Why New Firms Locate Where They Do: An Econometric Model." In William Wheaton, ed., *Interregional Movements and Regional Growth*. Washington, D.C.: Urban Institute.
- Choe, S. C., and B. N. Song. 1982. "An Evaluation of Industrial Location Policies for Urban Deconcentration in the Seoul Region." Urban Development Discussion Paper 7. World Bank, Washington, D.C.
- Ellickson, B. 1977. "Economic Analysis of Urban Housing Markets: A New Approach." R-2024-NSF. Santa Monica, Calif.: Rand Corporation.
- . 1981. "An Alternative Test of the Hedonic Theory of Housing Markets." *Journal of Urban Economics*, vol. 9, no. 1 (January), pp. 56-79.
- Erickson, R. A., and M. Wasylenko. 1980. "Firm Relocation and Site Selection in Suburban Municipalities." *Journal of Urban Economics*, vol. 8, no. 1 (July), pp. 69-85.
- Friedman, J. 1975. "Housing Location and the Supply of Local Public Services." P-5421. Santa Monica, Calif.: Rand Corporation.
- Hanushek, Eric A., and Byung Nak Song. 1978. "The Dynamics of Postwar Industrial Location." *Review of Economics and Statistics*, vol. 60, no. 4 (November), pp. 515-22.
- Henderson, J. V. 1980. "A Framework for International Comparisons of Systems of Cities." Urban and Regional Report 80-3. World Bank, Washington, D.C. Processed.
- Hoover, E. M., and R. Vernon. 1959. *Anatomy of a Metropolis*. Cambridge, Mass.: Harvard University Press.
- Ingram, G. K. 1977. *Residential Location and Urban Housing Markets*. Cambridge, Mass.: National Bureau of Economic Research.
- Kain, John F., and John M. Quigley. 1975. *Housing Markets and Racial Discrimination: A Microeconomic Analysis*. Cambridge, Mass.: National Bureau of Economic Research.
- Kemper, P. 1973. "The Location Decisions of Manufacturing Firms within the New York Metropolitan Area." Ph.D. dissertation. Yale University, New Haven, Conn. Processed.
- Lee, K. S. 1981. "Intra-Urban Location of Manufacturing Employment in Colombia." *Journal of Urban Economics*, vol. 9, no. 2 (March), pp. 222-41.
- . 1982a. "Determinants of Intra-urban Location of Manufacturing Employment: An Analysis of Survey Results for Bogotá, Colombia." Urban and Regional Report 81-3. World Bank, Washington, D.C. Processed.
- . 1982b. "A Model of Intraurban Employment Location: An Application to Bogotá, Colombia." *Journal of Urban Economics*, vol. 12, no. 3 (November), pp. 263-79.
- . 1985. "An Evaluation of Decentralization Policies in Light of Changing Location Patterns of Employment in the Seoul Region." Urban Development Discussion Paper 60. World Bank, Washington, D.C. Forthcoming in *Urban Studies*.
- . Forthcoming. "Employment Location in Colombian Cities." Water Supply and Urban Development Department, World Bank, Washington, D.C. Processed.
- Leone, R. 1971. "Location of Manufacturing Activity in the New York Metropolitan Area." Ph.D. dissertation. Yale University, New Haven, Conn. Processed.
- Lerman, S. R. 1977. "Location, Housing, Automobile Ownership and Mode to Work: A Joint Choice Model." *Transportation Research Board Record*, no. 610, pp. 6-11. Washington, D.C.: Transportation Research Board.
- McFadden, D. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." In P. Zarembka, ed., *Frontiers in Econometrics*. New York: Academic Press.
- . 1976. "The Theory and Practice of Disaggregate Demand Forecasting for Various Modes of Urban Transportation." Prepared for the Seminar on Emerging Transportation Planning Methods. U.S. Department of Transportation, Washington, D.C. Processed.
- Meyer, J. R. 1981. "Report on Proposed Korean Spatial Study." Urban and Regional Report 81-1. World Bank, Washington, D.C. Processed.
- Mieszkowski, P., and M. Straszheim, eds. 1979. *Current Issues in Urban Economics*. Baltimore, Md.: Johns Hopkins University Press.
- Mills, E. S. 1972. *Urban Economics*. Glenview, Ill.: Scott, Foresman.
- Murray, Michael P. 1982. "Here, There, Where? A Strategy for Evaluating Industrial Relocation Policies in Korea." Urban Development Discussion Paper 6. World Bank, Washington, D.C.
- Muth, R. 1969. *Cities and Housing*. Chicago: University of Chicago Press.
- Quigley, John. 1976. "Housing Demand in the Short Run: An Analysis of Polytomous Choice." *Explorations in Economic Research*, no. 3. Cambridge, Mass.: National Bureau of Economic Research. Available from University Microfilms, Ann Arbor, Mich.
- Renaud, Bertrand. 1981. *National Urbanization Policy in Developing Countries*. New York: Oxford University Press.
- Schmenner, R. 1973. "City Taxes and Industry Location." Ph.D. dissertation. Yale University. New Haven, Conn. Processed.
- . 1982. *Making Business Location Decisions*. Englewood Cliffs, N.J.: Prentice-Hall.
- Solow, R. 1972. "On Equilibrium Models of Urban Location." In Michael Parkin, ed., *Essays in Modern Economics*. London: Longman.
- Straszheim, M. R. 1974. *An Econometric Analysis of the Urban Housing Market*. Cambridge, Mass.: National Bureau of Economic Research.
- Struyk, Raymond J., and Franklin J. James. 1975. *Intrametropolitan Industrial Location*. Lexington, Mass.: Lexington Books.

Tolley, George S., Philip E. Graves, and John L. Gardner, eds.  
1979. *Urban Growth Policy in a Market Economy*. New  
York: Academic Press.

Townroe, D. M. 1979. *Employment Decentralization: Policy  
Instruments for Large Cities in Less Developed Countries*.

Vol. 10, pt. 2 in D. R. Diamond and J. McLoughlin, eds.,  
*Progress in Planning*. Elmsford, N.Y.: Pergamon.

World Bank. 1979. *World Development Report 1979*. New  
York: Oxford University Press.

## *Industrial Deconcentration Policy: Venezuela*

Benjamin Reif

In 1974 the Venezuelan government, conscious of problems caused by the concentration of activities in the north central area around the cities of Caracas, Maracay, and Valencia and the limited economic activity in lagging areas, decided to intervene in the spatial distribution of the country's manufacturing industry. A series of presidential decrees provided a legal framework for an industrial deconcentration policy. Evaluations of the results show little agreement: a report by the Ministry of the Environment (Ministerio del Ambiente 1979a, p. 81) claims that the behavior of the entrepreneur is changing, but a document prepared by the National Planning Office (Cordiplan 1979, pp. 7-9) indicates that the policy merely accentuated existing tendencies toward concentration of manufacturing activity in metropolitan Caracas and its surroundings. This study attempts to throw light on the debate.

Venezuela's industrial deconcentration policy had three aims: to prohibit the location of new manufacturing in the metropolitan region (Caracas and its surroundings); to induce industries that are considered hazardous to move to safer places and to encourage nonhazardous firms already located in the congested metropolitan region to move to designated development areas; and to attract new manufacturing plants to the designated development areas. The main policy instruments for carrying out industrial deconcentration have been direct incentives (including financial and fiscal incentives such as grants, loans, and tax concessions), indirect forms of assistance, such as the provision of industrial parks and the improvement of transport and communication facilities, and negative incentives, such as control over location. Incentives vary among zones in accordance with the national goal of reducing interre-

gional disparities in wealth and well-being. Venezuela's five areas of industrial deconcentration are shown in map 9-1.

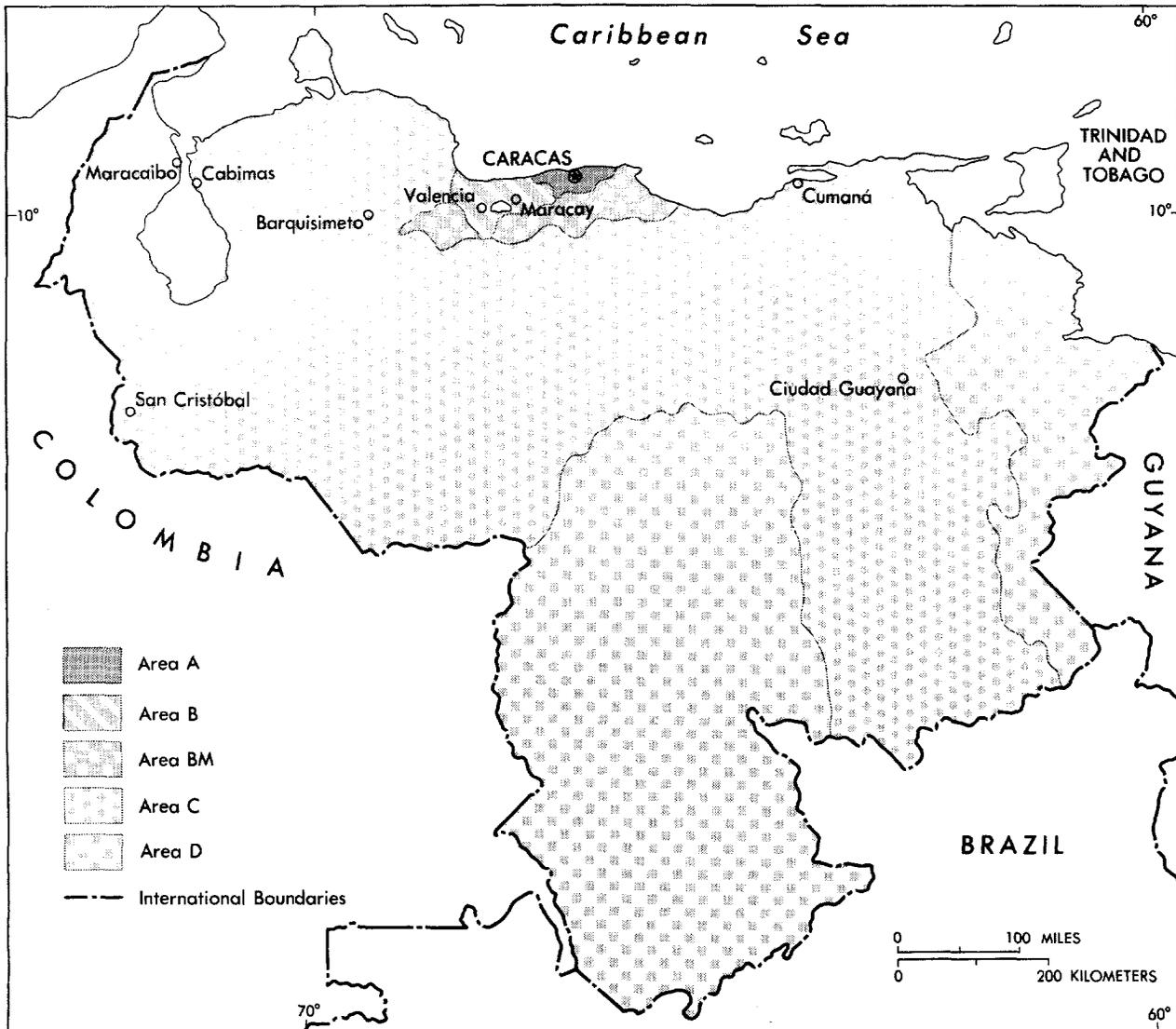
The main question is whether government instruments, including incentives and restrictions, have contributed to the deconcentration of industry. This question brings up more specific questions. Has deconcentration of industry taken place? Is deconcentration a result of the policy instruments? What other factors have influenced industrial location, and what has been their relative importance? The following sections attempt to answer these questions.

### **Has Deconcentration of Industry Taken Place?**

The effects of industrial deconcentration policy are expected to be felt at the regional level. The firm is the basic statistical unit, and in this study it is defined to be identical to an establishment. Much of the data, such as the number of employees and the production value, is tabulated by firm. On the basis of aggregate data for administrative regions, used to identify changes in the distribution of establishments (table 9-1), it can hardly be concluded that the spatial pattern of industry by administrative region has changed substantially. At most, it can be said that between 1974 and 1976 the concentration of manufacturing in the Capital region increased by 5 percent at the expense of the other regions.

It may be argued that the results, when presented as a whole, may fail to illuminate the fact that individual industrial subgroups may behave differently. Several

Map 9-1. Industrial Deconcentration Areas in Venezuela



manufacturing sectors have therefore been selected for analysis. (A study of all the sectors is beyond the scope of this research.)

Of the twenty-seven three-digit isic (International Standard Industrial Classification) manufacturing codes, sectors having between 140 and 300 establishments—the middle range—were chosen for study. Eleven sectors were initially selected, after excluding (1) industrial sectors with many establishments that serve local markets and are little affected by regional policy (for example, bakeries) and (2) sectors that have such a small number of plants (probably serving national or international markets, as is the case with oil refineries) that studying them would yield only limited insights into locational behavior. Selected industries also had to experience a high rate of growth. The industries chosen had to have strong linkages with other manufacturing

industries and with local input and product markets, as indicated by the input-output table of the Venezuelan economy, to include the possible effect of spatial policies on the location of transport-oriented industries. The levels of effective protection for the industries had to be diverse, as did their intensities of capital and labor. (The mean electricity consumption value, in kilowatt-hours per thousand hours, was used as a proxy to reflect any tendency toward deconcentration of labor-intensive industries.) Six three-digit isic codes were chosen for further study: textiles (321), leather products (323), plastic products (356), nonelectrical machinery and equipment (382), electrical machinery and equipment (383), and transport equipment (384).

Venezuelan manufacturing is concentrated in an axis that extends over parts of the Capital and Central regions. One way to ascertain whether deconcentration of

**Table 9-1. Shares of Manufacturing Establishments, by Administrative Region**  
(percent)

Region	1971	1974	1976	1978
Capital	51	50	55	54
Central	16	17	16	15
Centro-Occidental	8	7	7	7
Zuliana	8	9	7	7
Andes	7	8	7	8
Sur	1	0	0	0
Nor-Oriental	7	6	5	6
Guayana	2	3	3	3

*Memorandum:*

Number of establishments	6,401	7,554	9,538	10,478
--------------------------	-------	-------	-------	--------

Sources: Cordiplan (1973); Oficina Central de Estadística e Informática (1974, 1976, 1978).

**Table 9-2. Shares of Manufacturing Establishments in the Capital and Central Regions Combined**  
(percent)

Industry	1971	1974	1976	1978
Textiles	84	88	88	88
Leather	76	71	77	75
Plastics	97	90	90	89
Nonelectrical machinery	69	71	77	75
Electrical machinery	91	88	90	87
Transport equipment	71	72	75	72

Sources: Same as for table 9-1.

manufacturing has taken place is to combine the shares of the Capital and Central regions, as in table 9-2. No major changes took place during the period of the study except in the plastics and nonelectrical machinery industries, which show a relatively higher share for 1971 and 1976, respectively; that is, no significant evidence of concentration or deconcentration can be found.

Because the results are not given for the officially defined areas of industrial deconcentration, a proper evaluation of the effects of the policy instruments is difficult. Furthermore, aggregate figures like those used here do not permit assessment of the spatial changes in the manufacturing industry. The data are for the net number of establishments that existed in a particular year and do not show the actual numbers of new firms established in an area, of firms that went out of business, or of plants that changed locations. Attempts were made to find data that overcome these limitations.

By comparing the 1974 and 1976 industrial directories, changes that occurred in 1975 and 1976 were identified. Similarly, by comparing the 1976 and 1978 directories, changes during 1977 and 1978 were identified. Once the directories were matched, factors could be examined for each period—for example, which firms were *new* (founded during the period); which *closed down* (went out of business); which were *stationary* (remained in the same location); and which *moved* to another location or from another region. The results of the matching (table 9-3) show the share of new establishments in the Venezuelan industrial axis (areas A, B, and BM on map 9-1).

Each of the six industries shows a similar trend toward deconcentration. Some of this may be a consequence of the big increases in income and demand in 1974–76 as a result of oil price increases. The effects of these increases were perceived mainly in Caracas and led to an unusually high concentration of new firms. Nevertheless, as table 9-3 shows, for leather, nonelectrical machinery, and electrical machinery the share of the industrial axis in the national total decreased by more than 33 percent by the second period. It thus appears that between periods these three industrial subgroups experienced deconcentration. Declines in the shares of the other manufacturing groups—textiles, plastics, and transport equipment—are mainly a result of the initial distortions created by the big increase in demand and therefore are not necessarily evidence of decentralization.

The absence of disaggregated data before 1974 places limits on the conclusions. When an industrial subsector shows a tendency to deconcentrate, the tendency may have started in 1974 or may have its origins in an earlier period. This uncertainty makes it difficult to relate the

**Table 9-3. Share of New Manufacturing Establishments in the Industrial Axis**  
(percent)

Industry	1974–76	1976–78	Change
Textiles	89 (82)	81 (26)	– 8
Leather	81 (52)	47 (17)	– 34
Plastics	92 (98)	74 (27)	– 18
Nonelectrical machinery	87 (87)	51 (33)	– 36
Electrical machinery	88 (61)	33 (12)	– 55
Transport equipment	76 (74)	59 (34)	– 17

Note: Number of new firms is in parentheses.

Source: Author's estimates on the basis of the matched directories and Oficina Central de Estadística e Informática (1978).

pattern of industrial deconcentration to government policies. But the important issue is whether deconcentration of manufacturing took place over the two periods for which there is evidence for three of the six subsectors analyzed.

**Is Deconcentration a Result of Policy?**

Owing to difficulties in measuring the effects of regional policies (variables may be affected by migration and long-term structural changes as well as by regional policy), it is not possible to employ conventional statistical methods for rejecting or accepting hypotheses. Instead, we will assess alternative hypotheses on the basis of indirect evidence. Additional evidence from econometric models is discussed in subsequent sections.

**Financial Incentives**

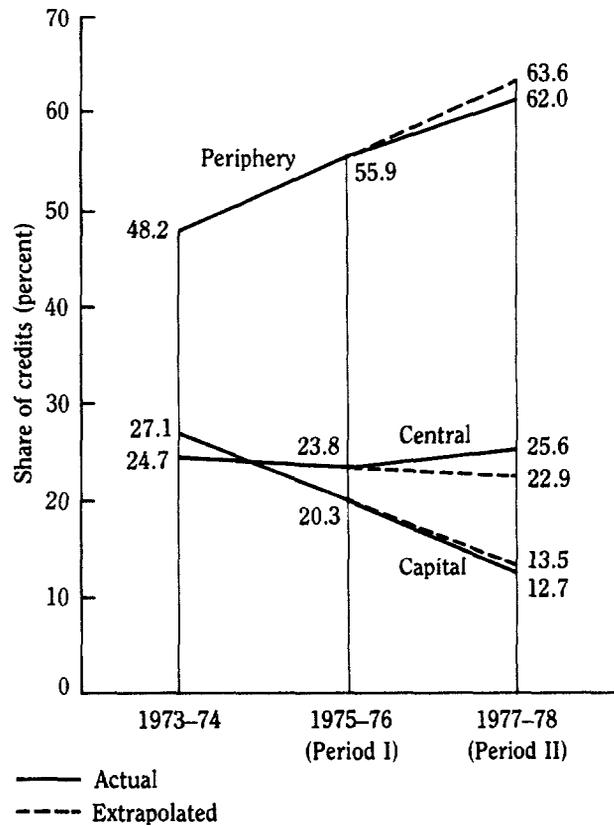
If industrialists took advantage of government-sponsored financial incentives in developing areas, this should be reflected in the number of credit applications in these areas. Available data are for number of credits awarded, not for number of applications; it is assumed that the proportion of applications approved does not vary by region. The number of credits granted to a region was thus used as a proxy for the number of applications. Similarly, the data did not differentiate between credits for new firms and for expansions; it was assumed that the proportion of expansions to new firms did not vary by region.

Figure 9-1 shows the actual and extrapolated shares of credits for the Capital, Central, and peripheral regions. The results suggest that the regional trends existed before the introduction of financial incentives in March 1976 and continued over the next two years, and that the introduction of the incentives may have had little additional effect on the spatial distribution of manufacturing.

**Fiscal Incentives**

A study by the Ministry of the Environment (Ministerio del Ambiente 1979b, p. 33) estimates that the savings from tax holidays represent only 1.5 to 2 percent of sales value. World Bank data indicate that—subject to caveats regarding statistical reliability—industrial profit rates are very high in Venezuela. When petroleum refining and the state-enterprise-dominated basic metal industries are removed from the total the average industrial profit rate, net of depreciation and taxes, in the 1970s rises from 19 to 28 percent, a figure that more accurately reflects what the private sector, of both national and

Figure 9-1. Actual and Extrapolated Share of Credits, Three Regions



foreign ownership, is able to earn through industrial investment. In relation to the average profit margins of Venezuelan manufacturing industry, the saving owing to fiscal incentives may be insignificant in many cases.

**Negative Incentives**

As early as January 1975 a decree was issued that prohibited new establishments from locating in the metropolitan area (Caracas and its surroundings) unless they qualified under special conditions. Furthermore, the government drew up three lists of industries, identified here as groups 1, 2, and 3. A firm's right to continue to operate or to locate in area A depended on the list to which it was assigned. It is assumed that the lists, although approved in 1976, had actually been in effect since January 1975, when the prohibition on locating in area A first took effect. The lag constituted an informal grace period to reduce hardships on firms whose location was already planned.

**New Firms.** New firms in group 1 are the only ones allowed to locate in the metropolitan area. All others are

considered "transgressors." The table shows that there was a lower percentage of transgressors in 1975-76 than in 1977-78, although the absolute number of transgressors was higher.<sup>1</sup>

Period	Number of new firms in area A	Number of transgressors	Percentage of transgressors
I (1975-76)	290	127	43.8
II (1977-78)	64	33	51.6

An explanation for the lower percentage in 1975-76 may be that government officials were stricter during that period. There are no disaggregated data before 1975, and it is not possible to know what was happening before the location controls were imposed. It may be that the percentage of firms identified with groups 1 and 2 that located in the metropolitan area was higher before the controls were imposed. Or, the controls may not have been implemented effectively in 1977-78.

*Moves of Existing Establishments.* Figure 9-2 shows an origin-destination matrix of interarea moves for each

Figure 9-2. *Origin-Destination Matrix of Interarea Movers, 1975-76 and 1977-78*

		1975-76					Moved out	Moved in	Net change
		To area							
From area		A	B	BM	C	D			
A			7	4			11	12	+1
B				1			1	7	+6
BM		12					12	5	-7
C							0	0	0

		1977-78					Moved out	Moved in	Net change
		To area							
From area		A	B	BM	C	D			
A			2	2			4	5	+1
B		2					2	2	0
BM		3					3	2	-1
D							0	0	0

Note: The framed cells represent the expected impact of the policy.  
Source: Author's estimates on the basis of the matched directories and on Oficina Central de Estadística e Informática (1968).

period. The framed cells represent the expected impact of deconcentration policy. In spite of the reduced number of cases, the general trend seems to be one of concentration. The evidence indicates that negative incentives were not applied effectively during the study period.

*Preliminary Conclusions.* There seems to be little evidence that changes in the spatial pattern of industry are related to the application of government instruments. The nearly stable trend in regional shares of credits does not support the hypothesis that financial incentives influenced the location of new firms. Information on the actual application of fiscal incentives is generally not available, but the impact of such incentives is likely to be insignificant compared with the profits of the manufacturing industry, and it is probable that they have had little effect on location. Negative incentives do not seem to have been applied effectively during the study period.

## Other Factors

The evidence suggests that some manufacturing industries had a tendency to deconcentrate and that government policies had little effect on industrial deconcentration. This leaves deconcentration unexplained and leads toward investigation of other factors that influence industrial location.

Changes in the Venezuelan economy during 1974, when the national budget increased threefold, probably had two significant effects on manufacturing. First, many new manufacturing firms appeared that were primarily concerned with satisfying the sudden increase in demand and that benefited from the high profits. During this period most entrepreneurs had a short-term commercial bent rather than a long-term interest in manufacturing. As a consequence, most of the firms established during 1975-76 chose sites without much concern for cost or other locational factors. Second, the steep increase in the number of manufacturing establishments aggravated the diseconomies of the physical and social infrastructure. This started in Caracas and slowly spread to many of the surrounding cities. New industries that located in places where the level of services was inadequate suffered along with plants already there.

By the second period (1977-78) the economy had reached equilibrium in the supply of and demand for national manufactured products. By then most entrepreneurs were not only aware of the problems of locating in the congested industrial axis but were also aware that the bonanza, or high-profit, period was over. These entrepreneurs were therefore more concerned with lo-

cational factors than were entrepreneurs during period I, and they probably planned more carefully the selection of their plant sites, looking for places with lower wages, a satisfactory water supply, a good social infrastructure, and low transport costs. The search for those characteristics may have led some industrialists to consider settlements outside the industrial axis, and consequently deconcentration occurred. Deconcentration probably began before the economic boom of 1974, but the sudden change in the economy at that time may have temporarily slowed the process.

These arguments may explain why some manufacturing sectors deconcentrated, but what about those which did not? A clue may be found in World Bank research which indicates that the salient feature of the Venezuelan industrial environment in the 1970s was the high degree of protection from import competition accorded by government policies. This suggests that highly protected industries were attracted to Caracas. For example, small, highly protected industries would probably consider closeness to the government's decisionmaking center to be crucial; exonerations, licenses, and other benefits received on a case-by-case basis might outweigh the disadvantages of locating in the industrial axis. Murray (1982) states that "access to government is one input in many production processes." Renaud (chapter 5, above), in a discussion of the unintended biases of national economic policies, stresses that the management practices of the central government and its regulation of economic activities require location of these activities close to the capital and contribute to the urban vortex. A complementary explanation would be that the savings that could have been obtained through deconcentration were low compared with the large profits made possible by the high level of protection.

If this argument holds, we would expect to find supporting evidence, for example:

- Concentration of highly protected manufacturing in the industrial axis and deconcentration of less protected manufacturing firms
- Greater concern with location on the part of less protected industries; a higher sensitivity of second-period industrialists to locational factors, owing both to their awareness of the physical limitations in cities of the industrial axis and to a return to more realistic profit levels
- Relatively little weight attached to government incentives as a factor in location.

The rest of this section will address these issues.

### *Concentration and Protection*

If the benefits of being close to government agencies that grant exonerations and other benefits outweigh the

diseconomies of being in the industrial axis, we might expect highly protected industries to concentrate and less protected industries to deconcentrate. The six manufacturing groups considered here behaved as expected: textiles, plastics, and transport equipment are highly protected and are concentrated; leather, non-electrical machinery, and electrical machinery face a low level of protection and are deconcentrated.<sup>2</sup>

One could argue that industries which concentrated did so not because of government protection but for other reasons and that the pattern of association described above is spurious. This is possible, and we do not want to present the results shown here as conclusive. Attempts to find other causes were made. For example, we tested capital or labor intensity, the product cycle theory, and other factors, but none were associated with concentration.

The next tasks were to determine whether less protected industries were more concerned with factors of location, to identify these factors, and to evaluate the extent to which government financial incentives have been considered as a locational factor by industrialists. For this purpose a discrete choice logit model was used that represents an entrepreneur who considers the attributes of both the firm and the city in deciding where to locate a manufacturing plant. The model relates the variables of the cities to the locational choice of the entrepreneur. The choice is made from a known set of alternative cities. Modelers are faced with the task of explicitly defining this location choice set. In one respect this is virtually a hopeless task: Venezuela consists of many settlements of varying size and quality, and the subset actively considered by any given entrepreneur is probably known to the industrialist alone. Central place theory, however, argues that a natural hierarchy, or system of settlements, exists within a country. If such a hierarchy exists, it is possible to identify the most important settlements for manufacturing activity. (These settlements presumably constitute a large sample of possible locations for those manufacturing entrepreneurs who desire a regional or national market for their products.) For such a hierarchy the modeling technique employed here is useful. The logit model assumes one underlying utility function and looks at relative choice frequencies as sample draws.

Because of computational limitations—it is unlikely that existing computer programs could adequately handle more than twenty potential choices of location—a limited number of alternative locations was selected. For this purpose, several criteria were applied.

- The choice set should be large enough to adequately represent the variables to be analyzed, that is, to have a wide range of values for the independent variables.

- The choice set should include a mix of core and peripheral cities as well as settlements with and without manufacturing. The manufacturing need not be new.
- The choice set should represent all areas of deconcentration as well as the special growth centers identified by the policy of industrial deconcentration.
- The choice set should have geographic balance; it should include cities located in the east, west, and south of the country as well as in the traditional north-central area.

With the use of this selection procedure, twenty cities were chosen.

### *Specification of the Logit Model*

Many independent variables may affect locational choice. To reduce them to a manageable set, we examined the statistical correlation among a large number of variables for which we had data. High correlations between groups of variables were found—for example, between accessibility to local inputs and accessibility to final consumers. After examination of the variables in terms of statistical correlation, existing empirical work (see, for example, Carlton 1979), reliability of the source, and, especially, the economic theory of location, the independent variables were selected. They are as follows:

<i>UI</i>	1 in the Caracas alternative, 0 otherwise
<i>FI</i>	Percentage savings in present value terms owing to financial incentives (log)
<i>TWAGE</i>	Wage times number of employees in the firm (log)
<i>LIP</i>	Agglomeration potential associated with local inputs (log)
<i>TEC</i>	Proportion of the city's professionals, technicians, and related occupations to its population (log)
<i>STR</i>	Average number of working hours lost per thousand workers per year because of unexpected stoppages (log)
<i>PHN</i>	Number of telephone lines (installed capacity) per thousand inhabitants (log)
<i>DOC</i>	Number of inhabitants per doctor (log)
<i>FIP</i>	Agglomeration potential associated with foreign inputs (log)
<i>WPR</i>	Percentage of the city's population serviced with water (log)

The variable *UI* is a constant term in the utility function that measures the so-called pure alternative effect,

that is, the net effect of all attributes of an alternative which is not measured by the other variables. A dummy variable was assigned to Caracas because it is the location chosen by about 66 percent of new manufacturing industries.

The variable *FI* represents the financial incentives of the industrial deconcentration policy. It incorporates a number of elements, including interest rates, grace and loan periods, and uses of money, whether for land and construction, machinery and equipment, or working capital.

Total wage (*TWAGE*) represents the interaction between the prevailing wages of the city and the number of employees in the firm—that is, the total payroll that each firm would have in each of the twenty cities. The interaction reflects the hypothesis that firms with large payrolls are proportionately much more sensitive to labor costs than are smaller ones. Furthermore, it may be hypothesized that firms tend to locate where wages are lower and hence a negative coefficient is expected.

The agglomeration potential associated with local inputs, *LIP*, was designed to reflect a firm's accessibility to both local inputs and product markets in each city. It is a composite variable that reflects not only industrial linkages but also transport costs. It is assumed that as the distance to a particular source of inputs increases, the attractiveness of that source decreases.

The variable *TEC* reflects the hypothesis that entrepreneurs look for places where technicians and related professionals are available. This should be especially true for firms with a high level and proportion of skilled workers. A positive coefficient is expected.

The variable *STR* reflects the hypothesis that entrepreneurs tend to locate their plants away from areas of high labor unrest. It is measured in terms of average working hours lost per thousand workers per year because of unexpected stoppages. A negative coefficient is expected.

The variable *PHN* reflects the effect of telephone service on an industry's desire to locate in cities with telecommunication facilities. A positive coefficient is expected.

The variable *DOC* measures the number of inhabitants per doctor. It may be hypothesized that entrepreneurs are more attracted to cities with better health facilities, and it is assumed that a lower number of inhabitants per doctor implies better health services. A negative coefficient is expected.

The agglomeration potential associated with foreign inputs, *FIP*, was designed to reflect the accessibility to foreign inputs in each city. It is a composite variable that reflects dependence on foreign raw materials and components and the transport costs to international ports. The hypothesis is that the higher the proportion of

Table 9-4. Parameter Estimates for Five Model Specifications for 1975-76

Variable	Model specification				
	1	2	3	4	5
<i>UI</i>	0.35 (0.87)	-0.95 (1.50)	-2.85 (2.71)	-3.04 (2.88)	-1.06 (1.75)
<i>FI</i>	—	—	—	—	—
<i>TWAGE</i>	-1.32 (2.90)	-0.91 (1.93)	-1.42 (3.16)	-1.45 (3.21)	-1.47 (3.18)
<i>LIP</i>	1.12 (8.18)	1.16 (8.06)	1.36 (7.48)	1.29 (6.91)	1.09 (7.19)
<i>TEC</i>	1.90 (3.01)	3.20 (4.20)	3.26 (4.25)	3.41 (4.48)	2.71 (3.90)
<i>STR</i>	-0.19 (1.99)	-0.16 (1.52)	-0.49 (3.69)	-0.51 (3.68)	-0.31 (2.77)
<i>PHN</i>	—	1.24 (3.21)	—	0.76 (1.94)	1.15 (3.16)
<i>DOC</i>	—	—	-2.37 (3.52)	-0.82 (2.46)	—
<i>FIP</i>	—	-0.50 (3.74)	—	—	—
<i>WPR</i>	1.39 (1.83)	—	4.12 (3.28)	3.78 (2.83)	1.59 (1.73)
Sum of squared residuals	12,350	19,890	43,820	40,660	15,300
Degrees of freedom	7,157	7,156	7,156	7,155	7,156
Percentage correctly predicted	70.29	70.29	70.29	70.29	70.29
Likelihood ratio index	0.5637	0.5742	0.5696	0.5714	0.5686
Likelihood ratio statistic	1,273	1,297	1,287	1,291	1,284

— Not applicable.

Note: Number of firms in the sample, 377. Numbers in parentheses are *t*-statistics.

imported components for local assembly, the closer the assembly plant will be to international ports. A positive coefficient is expected.

The variable *WPR* reflects the hypothesis that industries tend to locate where water is available. This variable may be considered a proxy for the availability of infrastructure services.

#### Selection of the Model Specification

For any applied model of locational choice, the selection of an appropriate specification must include experimentation with the data. It is, of course, important to avoid mining the data. To guard against this danger, all experimentation was done with the subsample of data for the period before the introduction of the financial incentives package. As will be seen, this model fits the second subsample equally well.

Several distinct logit model specifications were estimated with data for the first period. These models correspond to different combinations of the variables selected earlier. The details of five of the model specifications estimated are given in table 9-4. Of the five models, we reject model 2 because of the unexpected sign on the

variable *FIP*. Models 3 and 4 are dropped because of high variation in the parameter values of the variable *DOC*. We disregard model 5 because it has one more variable than model 1 and adds no additional explanatory power. (Both models have a likelihood ratio index of 0.56.) Furthermore, the unexplained residuals are higher in model 5. (The sum of squared residuals is 15,300 in model 5 and 12,350 in model 1.) Hence, of the five specifications we select the first; indeed, specification 1 was selected out of all those analyzed as the model to be used in this research. The likelihood ratio index of 0.56 for this model indicates that its overall goodness of fit is satisfactory.

To analyze the data for the second period (1977-78), variable *FI* (financial incentives) was added to specification 1, and the parameters were estimated. The results appear in table 9-5, which allows us to compare the results from the model with data from the two periods 1975-76 and 1977-78. The model seems to be structurally stable over time. The signs of the coefficients (except *UI*) are the same for both periods. Furthermore, the values of the coefficients are similar except for *WPR* (availability of water or infrastructure services), which shows a significant increase in the second period.

Table 9-5. *Parameter Estimates for the Sample of Six Industry Subgroups, 1975-76 and 1977-78*

<i>Variable</i>	1975-76	1977-78
<i>UI</i>	0.35 (0.87)	-0.06 (0.08)
<i>FI</i>	—	0.01 (0.31)
<i>TWAGE</i>	-1.32 (2.90)	-1.69 (3.86)
<i>LIP</i>	1.12 (8.18)	0.96 (5.26)
<i>TEC</i>	1.90 (3.01)	1.41 (1.50)
<i>STR</i>	-0.19 (1.99)	-0.21 (1.91)
<i>WPR</i>	1.39 (1.83)	4.79 (3.45)
Sum of squared residuals	12,350	1,914
Degrees of freedom	7,157	2,235
Percentage correctly predicted	70.29	52.54
Likelihood ratio index	0.5637	0.3683
Likelihood ratio statistic	1,273	260.4
<i>Memorandum: Number of firms in the sample</i>	377	118

— Not applicable.

Note: Numbers in parentheses are *t*-statistics.

### *Analysis of the Coefficients*

The dummy coefficient for Caracas, *UI*, is subject to a wide confidence interval in both periods, as is shown in table 9-5. The net effect of all attributes of Caracas which are not measured by the other variables may be the most important factor in attracting new industry, but the data simply do not allow us to determine this effect with much precision. Indeed, the opposite could be the case; Caracas may exert a negative effect. (Some evidence of this is seen in table 9-4.) In those cases in which we have controlled for more variables, a negative coefficient for Caracas is obtained.

The coefficient of the financial incentives variable (*FI*) is very small and statistically insignificant. The failure of financial incentives to show up as an important influence on location is consistent with our previous finding. Wage levels, especially when they interact with the number of employees, *TWAGE*, exert a large influence on the location of industry. The coefficients are statistically significant in both periods. The accessibility to sources of local inputs and to product markets, *LIP*, enters significantly in both periods; the coefficients are statistically significant. Having a pool of technical expertise in the city, *TEC*, as measured by the proportion of the city's professionals, technicians, and related occupations to its population, seems to be an important factor in location. The coefficient of the variable measuring

work stoppages, *STR*, is not large but is statistically significant in both periods, which implies that entrepreneurs are concerned with locating their plants away from cities with high labor unrest. The coefficient for the variable measuring water availability, *WPR*, increases more than three times in the second period and becomes even more statistically significant. A close look at variable *WPR* in table 9-5 indicates two issues. First, manufacturing firms seem to be attracted to places where water (or infrastructure services) is available. Second, it is likely that second-period entrepreneurs were more sensitive to availability of water (or infrastructure services).

On the basis of these results it seems reasonable to conclude the following.

- Wages are an important factor in explaining the location of industry.
- The evidence provides little support for the proposition that financial incentives have exerted an effect on the location of manufacturing.
- Accessibility to both local inputs and product markets exerts a large influence on the locational decision of a manufacturing firm.
- The presence of technical expertise is probably an important locational factor.
- The more labor unrest a city has, the less likely it is to attract new manufacturing.
- The availability of water (or infrastructure services) exerts a major influence on the location of new manufacturing. It is likely that this variable was more important during the second period than during the first period.

### *Results, by Industry Subgroups*

So far the results have dealt with all six industrial subgroups together. The next question is whether the results apply to each subgroup. One might expect the importance of the attributes to vary among manufacturing groups.

The number of establishments (by industry subgroup and by year of foundation) to be used in our econometric analysis is less than the total number of new establishments because firms that locate outside the choice set (the twenty cities selected) are not considered. Because electrical machinery had only nine establishments in the second period, and it is unlikely that we can infer any meaningful conclusions from the analysis of such a small number of observations, results for electrical machinery for the second period are omitted.

The results, stratified by manufacturing groups, are presented in tables 9-6 to 9-8. For plastics, the maximum likelihood routine had difficulty converging when

Table 9-6. *Parameter Estimates for the Textile and Transport Equipment Industries, 1975-76 and 1977-78*

Variable	Textiles		Transport equipment	
	1975-76	1977-78	1975-76	1977-78
<i>UI</i>	0.05 (0.02)	-0.09 (0.02)	-2.38 (1.78)	-0.33 (0.16)
<i>FI</i>	—	0.04 (0.38)	—	0.04 (0.25)
<i>TWAGE</i>	-2.66 (0.88)	-2.54 (0.79)	-3.31 (1.78)	0.14 (0.13)
<i>LIP</i>	1.42 (1.79)	1.36 (1.54)	2.00 (3.31)	1.23 (2.36)
<i>TEC</i>	4.69 (1.78)	2.21 (0.51)	3.54 (1.85)	-0.16 (0.07)
<i>STR</i>	-0.22 (0.80)	-0.09 (0.31)	-0.55 (2.12)	-0.47 (1.68)
<i>WPR</i>	9.73 (1.37)	20.93 (1.66)	1.22 (0.72)	6.88 (1.52)
Sum of squared residuals	795.7	267.7	929.4	465.1
Degrees of freedom	1,286	430	1,153	468
Percentage correctly predicted	86.76	78.26	44.26	52.0
Likelihood ratio index	0.7685	0.67	0.3478	0.3793
Likelihood ratio statistic	313.1	92.32	127.1	56.81
<i>Memorandum:</i> Number of firms in the sample	68	23	61	25

— Not applicable.

Note: Numbers in parentheses are *t*-statistics.

the *WPR* variable was included in the model for 1977-78. Hence, in table 9-8 we present the results for plastics for period I only. To investigate the possible effects of the independent variables on locational decisions in the plastics industry, several other specifications without *WPR* were analyzed, and the results are presented in table 9-9.

*Wages (TWAGE).* We indicated earlier that, for the unstratified sample of industries, wages seem to exert a large influence on location. Under stratification the wage coefficient for all six industries is always subject to a wide confidence interval (see tables 9-6 to 9-8). Except for the plastics industry, the expected negative sign predominates. Wages seem to be very important for the leather, nonelectrical machinery, and textiles industries. Furthermore, for leather and nonelectrical machinery the coefficients in period II are larger, which may imply an increased sensitivity to this variable.

The hypothesis proposed here is that less protected industries are more concerned with locational factors and that the economic bonanza of the first period may have masked the effects of locational factors. We expect coefficient values for the less protected industries to be

larger during period II; that is, it is hypothesized that during period II entrepreneurs of less protected industries searched more carefully for places that offered reduced costs and increased revenues.

The pattern of association between the protection level and the expected behavior of wages shows that leather and nonelectrical machinery, with low levels of protection, have larger wage coefficients (*TWAGE*) during period II, and textiles and transport equipment, which are highly protected, do not, so they behave as expected. (Plastics and electrical machinery are not considered because of the limitations indicated earlier.) Thus, less protected industries are more sensitive to wages during period II, in accordance with our hypothesis. Even if the plastics and electrical machinery industries are considered, these conclusions hold. This finding lends additional support to the hypothesis that protection for manufacturing may suppress the entrepreneur's interest in locating in low-wage areas.

*Access to Local Inputs and Product Markets (LIP).* The coefficients for the variable which measures access to sources of local inputs and product markets (*LIP*) are usually important for all six manufacturing

Table 9-7. *Parameter Estimates for the Leather and Nonelectrical Machinery Industries, 1975-76 and 1977-78*

Variable	Leather		Nonelectrical machinery	
	1975-76	1977-78	1975-76	1977-78
<i>UI</i>	5.26 (2.02)	-2.26 (0.67)	1.26 (1.36)	-0.20 (0.13)
<i>FI</i>	—	-0.16 (0.99)	—	0.15 (0.97)
<i>TWAGE</i>	-2.66 (0.73)	-12.23 (1.40)	-0.86 (0.49)	-2.26 (0.94)
<i>LIP</i>	0.66 (0.95)	1.15 (1.10)	1.03 (2.89)	0.96 (2.36)
<i>TEC</i>	1.01 (0.30)	7.77 (1.32)	0.90 (0.43)	2.82 (0.76)
<i>STR</i>	-0.46 (1.57)	-0.48 (0.54)	-0.26 (0.80)	0.10 (0.26)
<i>WPR</i>	10.36 (2.32)	9.71 (1.76)	1.15 (0.61)	4.09 (1.22)
Sum of squared residuals	803.1	190.6	1,364	345.2
Degrees of freedom	887	259	1,343	449
Percentage correctly predicted	87.23	57.14	73.24	29.17
Likelihood ratio index	0.7891	0.47	0.5831	0.2118
Likelihood ratio statistic	222.2	39.50	248.0	30.46
<i>Memorandum:</i> Number of firms in the sample	47	14	71	24

— Not applicable.

Note: Numbers in parentheses are *t*-statistics.Table 9-8. *Parameter Estimates for the Plastics and Electrical Machinery Industries, 1975-76*

Variable	Plastics	Electrical machinery
<i>UI</i>	1.87 (2.22)	-0.59 (0.63)
<i>FI</i>	—	—
<i>TWAGE</i>	1.09 (0.74)	-0.42 (0.17)
<i>LIP</i>	0.45 (1.69)	1.40 (3.61)
<i>TEC</i>	1.51 (1.15)	1.20 (0.54)
<i>STR</i>	-0.11 (0.51)	-0.27 (0.72)
<i>WPR</i>	-0.93 (0.60)	0.42 (0.21)
Sum of squared residuals	1,727	764.8
Degrees of freedom	1,495	963
Percentage correctly predicted	74.68	52.94
Likelihood ratio index	0.5889	0.4428
Likelihood ratio statistic	278.7	135.3
<i>Memorandum:</i> Number of firms in the sample	79	51

— Not applicable.

Note: Numbers in parentheses are *t*-statistics.

groups (see tables 9-6 to 9-8). Coefficients for transport equipment, nonelectrical machinery, and electrical machinery are statistically significant, as is also true for plastics, especially in the four model specifications that appear in table 9-9.

*Technical Expertise (TEC).* When results are stratified by industry subgroup, technical expertise exerts influence on each of the six manufacturing sectors. The coefficients have wide confidence intervals (primarily during period I) except for textiles and transport equipment. Table 9-6 shows that textiles and transport equipment behave unexpectedly by lowering their coefficients during period II. Table 9-7 indicates that leather and nonelectrical machinery have larger coefficients in period II, in accordance with the expected behavior.

The pattern of association between the level of protection and the expected behavior for technical expertise provides additional evidence in support of the hypothesis that during period II less protected industries were more concerned with locational factors—that is, the values for the coefficients of the less protected industries during the second period are larger than the values for period I. Leather and nonelectrical machinery, which have low levels of protection, have larger *TEC* coefficients; textiles

Table 9-9. *Parameter Estimates for Four Model Specifications for the Plastics Industry, 1977-78*

Variable	Model specification			
	A	B	C	D
<i>UI</i>	67.32 (2.74)	69.26 (2.65)	70.71 (2.65)	67.08 (2.66)
<i>FI</i>	7.30 (2.78)	8.21 (2.76)	7.11 (2.80)	7.52 (2.78)
<i>TWAGE</i>	1.57 (0.93)	3.07 (1.45)	-1.32 (0.29)	0.03 (0.00)
<i>LIP</i>	2.70 (2.52)	2.57 (2.27)	3.18 (2.31)	3.40 (1.78)
<i>TEC</i>	3.70 (1.07)	9.94 (2.04)	2.63 (0.68)	7.81 (1.55)
<i>STR</i>	-1.33 (1.81)	-1.57 (1.90)	-1.32 (1.64)	-1.73 (1.51)
<i>PHN</i>	—	4.75 (1.39)	—	6.38 (1.15)
<i>DOC</i>	—	—	5.70 (0.77)	7.70 (1.06)
Sum of squared residuals	168.1	95.92	159.8	87.13
Degrees of freedom	431	430	430	429
Percentage correctly predicted	60.87	52.17	60.87	52.17
Likelihood ratio index	0.5736	0.6055	0.5793	0.6150
Likelihood ratio statistic	79.04	83.44	79.83	84.75
<i>Memorandum:</i> Number of firms in the sample	23	23	23	23

— Not applicable.

Note: Numbers in parentheses are *t*-statistics.

and transport equipment, which are highly protected, do not, so they behave as expected. An explanation as to why some subgroups show a lower coefficient for technical expertise in period II may be that it is cheaper for the firm to pay whatever amount is required to relocate its technicians with it than to locate in places where technical expertise is available but where other locational factors are lacking. The additional cost is more easily absorbed by high-profit firms, which usually are highly protected.

*Labor Unrest (STR).* The coefficient for labor unrest is usually small and is subject to a wide confidence interval, except in transport equipment, where it is statistically significant. An explanation for the high significance (although low value) in this sector may be that it includes motor vehicle assembly and the specialized manufacture of motor vehicle parts. Because workers in these activities are usually well organized, the transport equipment sector is prone to conflicts. Locating plants away from areas of high labor unrest may be a measure to reduce the probability of unexpected stoppages.

*Water Availability (WPR).* Plastics and electrical machinery seem to be little affected by water availability, but the coefficient for this variable is quite large for the

other four manufacturing industries. Furthermore, with the exception of the leather industry, for which the coefficient in period II remains stable, the parameters for the other three subgroups increase dramatically during 1977-78. Again we may argue that this increase may be a result of heightened awareness among second-period entrepreneurs. The reason that the coefficient remains unchanged during period II for leather manufacturing can be found in the ISIC 323 code, leather and leather products, which includes tanning and leather finishing as well as fur dressing and dyeing. According to Fair and Geyer (1954, p. 856), high-quality water is crucial to production for this subgroup, and it is unlikely that first-period entrepreneurs would have overlooked its importance. Hence, water availability was seen as equally important during both periods by leather industrialists.

*Caracas Dummy (UI).* The net effect of all attributes of Caracas not measured by the other variables is reflected in the dummy variable, *UI*. Among the attributes not captured are diseconomies related to congestion, land cost, and location control. Other factors, however, such as access to the bureaucracy, work in favor of Caracas.

The contribution of Caracas to the utility functions of

less protected firms is expected to decline with time because of increasing diseconomies. For highly protected industries the disadvantages of Caracas also increase with time, but this effect may be offset by the rising fear that protection will be removed (a likelihood for textiles today), or by the search for additional benefits, such as those related to local content regulation. Both factors make Caracas attractive. Hence, large negative effects in period II cannot be expected for highly protected industries, since nearness to government is an advantage that counteracts the disadvantages of locating in Caracas. The pattern of association between the level of governmental protection and the expected behavior for the Caracas variable shows that leather and non-electrical machinery, with low levels of protection, have a lower negative coefficient ( $UI$ ); textiles and transport equipment, which are highly protected, do not, so they behave as expected.

It may be argued that Caracas was not always perceived by entrepreneurs as experiencing diseconomies. The larger negative effects observed in period II probably represent a return to conditions similar to those that existed before the distortions of the sudden economic changes of 1974. Unfortunately, the effects of location control cannot be separated from those of other diseconomies such as congestion and land cost.

*Financial Incentives (FI).* In tables 9-6 and 9-7 the coefficients on the variable for financial incentives are sometimes negative, always small, and usually statistically insignificant. Until adequate data are available conclusions cannot be drawn for the electrical machinery and plastics industries. Only nine new electrical firms located within the twenty cities, too small a number for statistical tests. For plastics we recall that the maximum likelihood routine had difficulty converging with the *WPR* variable in the model for 1977-78. Several other specifications without *WPR* were analyzed for the plastics industry, however, and the results for four of them are presented in table 9-9. An analysis of the coefficients for the variable *FI* shows that for plastics the financial incentive variable enters in all four specifications with a large and significant coefficient. Thus, for this sector financial incentives exerted a substantial effect on location. One might argue that these results are a product of sampling error. On the assumption that they are not, we offer the following explanation, based on interviews with entrepreneurs. ISIC code 356 includes the molding, extruding, and fabricating of plastic articles. A typical factory has few employees, requires relatively little floor space, and has one or two machines that process raw material. The basic component of the machines is the mold, which determines the end product (such as a fork, plate, or syringe) and accounts for a high proportion of

machine cost. Usually, a different mold is needed for each product.

The relative ease of setting up a plastics factory contributed to the growth of this sector. During period I, seventy-nine new establishments were started, fifty-nine of them in Caracas. By period II private financial institutions had apparently become more aware of the risks involved in extending loans to this sector because of the proliferation of plastics factories. The low flexibility of plastics machinery may have contributed to this perception: in case of bankruptcy, a bank would have to take over the plant and sell it, and a prospective buyer would have to be interested in producing products similar to the existing line or would have to make additional costly investments in molds. That the earlier owner went bankrupt because of high competition lent the operation an added sense of risk. Consequently, during period II private financial organizations were probably reluctant to give loans for manufacture of plastic products. Under these conditions an industrialist in plastics might well find it advantageous to move to the designated areas to become eligible for government loans. Some industrialists therefore reacted favorably to the financial incentives. It is likely that some of these entrepreneurs selected locations in the area of deconcentration closest to Caracas (BM). By doing so, they remained close to the industrial axis and did not contribute to deconcentration.

## Conclusions

The hypothesis that less protected industries are more concerned with locational factors is supported by the evidence. More generally, this model attempts to identify the factors that affect the locational decisions of firms. Alternative explanatory factors were sought when evidence was found that government incentives do not affect a firm's location. In particular, it was found that entrepreneurs attach an insignificant weight to government financial incentives.

The model does not claim to represent completely the complex set of locational factors that enter into the decision process; it only attempts to point out some factors that decisionmakers seem to consider when choosing a location. Some of these may be subject to control by policymakers.

The findings should be interpreted in the context of two limitations. First, only a short period has elapsed since 1974, when the first presidential decree concerning industrial deconcentration was issued. Second, the data base of this study cannot be accepted without the standard caveats, and the results must be considered tentative. This study reaches cautious conclusions on

the basis of available statistics and indicates a methodological approach for future work.

## Notes

1. Data are author's estimates based on the matched directories and Oficina Central de Estadística e Informática (1978).

2. World Bank data; Oficina Central de Estadística e Informática (1978); matched directories for 1974 and 1976. Throughout this chapter, an industry group is classified as having a low level of protection if its effective protection is below 80 percent and as having a high level if effective protection is above 130 percent.

## Bibliography

Carlton, Dennis. 1979. "Why Do New Firms Locate Where They Do: An Economic Model." In William Wheaton, ed., *Interregional Movements and Regional Growth*, pp. 13-50. Washington, D.C.: Urban Institute.

Cordiplan. 1973. *Encuesta Industrial 1971*. Caracas.

———. 1979. *Evaluación de la Política de Desconcentración Industrial*. Caracas.

Fair, Gordon, and John Geyer. 1954. *Water and Wastewater Disposal*. New York: John Wiley & Sons.

Ministerio del Ambiente. 1979a. *Esquema de Ordenamiento de la Región Centro Norte Costera*. Vol. diagnóstico, tomo de síntesis. Caracas.

———. 1979b. *Esquema de Ordenamiento de la Región Centro Norte Costera*. vol. 7: *Industria*. Caracas.

Murray, Michael P. 1982. "Here, There, Where? A Strategy for Evaluating Industrial Relocation Policies in Korea." Urban Development Discussion Paper 6. World Bank, Washington, D.C.

Oficina Central de Estadística e Informática. 1974. *Encuesta Industrial*. Caracas.

———. 1976. *Encuesta Industrial*. Caracas.

———. 1978. *Directorio Industrial*. Caracas.



# Part IV

---

## *Managing the City*

Part IV is devoted to a detailed discussion of specific urban problems. We have traced the roots of urbanization concerns and reviewed the limits of aggregative concentration and decentralization policies; now we turn to more sectoral policy issues. Bahl and Linn in chapter 10 address the fiscal problems of city management. Problems of housing, transport, and the environment are covered by Ingram, Pachon, and Thomas in chapters 11, 12, and 13, respectively. Evaluation of urban projects is discussed by Keare in chapter 14. The discussions directly address policy issues to reveal the scope for ameliorating urban problems.

An important aspect of urban finances is their implications for alleviation of urban poverty. In principle, a progressive tax policy can favor the poor, although a greater impact on poverty can probably be obtained from expenditure measures in such areas as housing, transport, and social services. There are, however, extensive examples of failed initiatives in these areas—slum removal, the banning of street vendors, and expensive investment programs. Urban investments, if they are to contribute to the welfare of the poor, must take into account the needs of the poor and their ability to afford the services.

### Toward Greater Fiscal Efficiency

Chapter 10 provides a comprehensive statement on urban revenue and expenditure for developing countries. The growing fiscal problems in cities in developing countries can be attributed to unprecedented urban growth, for which many countries are ill-prepared, and to the demands made on local services by particular new investments for economic development. The demand for local services is sensitive to increasing population, espe-

cially to growing numbers of poor people. Positive income elasticities for public services and demonstration effects from the developed world also influence the demand for local services. Factors that affect the costs of local services include the increasing costs of loans for large urban infrastructure investments, wages, labor unions, rising land and energy prices, and inflation.

The analytical and empirical findings presented in chapter 10 are relevant to most developing countries. Efficient means of raising taxes and controlling expenditure are explored. In many urban contexts a greater use of property and motor vehicle taxes, to be adopted carefully by local authorities, could be effective. Greater reliance on user charges has the merit of being directly linked to services provided. The chapter also reviews controversial proposals and discusses political considerations. Reform proposals are commonplace; concrete actions have been rarer, for a variety of reasons mentioned in the chapter. It could be inferred from developing-country experience that gradual adjustments might be more realistically expected, and perhaps more sustainable, than radical and sweeping reforms.

### Urban Housing

Overcrowding and lack of adequate housing are widely observed urban phenomena. In response, many governments attach high priority to the provision of urban shelter to low-income groups. Most analysts now recognize that housing includes not only shelter structures but also the land and the services provided, including water and sewerage. Urban policy toward the provision of such public services has a great impact on the quality of life, comparable to the provision of shelter itself.

The importance of housing is borne out by the high proportion of household budgets that is spent on it. World Bank estimates indicate that average housing expenditures in developing countries range from 15 to 25 percent of income. Furthermore, according to U.N. data an average 4 percent of the GDP in developing countries is spent on new residential construction. Of course, there is considerable variation in these estimates among countries, and the quality of data varies widely.

The differences between urban and rural areas bring out the origins of the urban housing problem. Housing expenditures are systematically higher in urban centers than in rural areas. Thomas (1978) shows that the average household's nominal expenditures in Lima may be three times higher than in the rural areas of Peru. Similar findings are contained in a study on Brazil by the same author (Thomas 1982). In general, much of the variation in urban-rural expenditures seems to be explained by the disproportionately larger outlays in urban areas on nontradables, the most important of which is housing.

The quality and quantity of housing in urban areas are an important aspect of the urban problem. To improve the housing supply, policymakers must take into account in greater detail the behavior of market demand and the cost of providing housing. The demand for housing reflects willingness to pay for a set of housing services. It is important to measure adequately the demand for housing in the context of the growth of cities, income expansion, and changes in relative prices.

Chapter 11 deals with one aspect of housing demand and provides an illustration of the equation for estimation of housing demand. The estimates are based on household interview data from Bogotá and Cali, Colombia. A comparison of parameter values with those obtained from North American data sets shows that demand elasticities in Colombia are generally comparable in magnitude with those in the United States. The approach employed to represent variation in housing price in the demand equations uses a theoretically attractive and computationally straightforward procedure that is based on residential location theory. A simple exercise illustrates the magnitude of bias of the income elasticity of demand that can result from incorrect data aggregation techniques. Moreover, correctly aggregated data produce income elasticity estimates that are similar to those obtained from disaggregated or micro data.

## Urban Transport

Transport is both a factor that drives city growth and a source of urban problems. Transport technology and

costs contribute to the configuration and growth of cities, and transport often turns out to be a critical constraint on further expansion and on the smooth operation of the city. Government policy on urban transport has important effects on the quantity, quality, and price of this service. Policy issues include government intervention regarding licensing, public versus private service, fares and subsidies, routing decisions, and investments.

Careful estimation of supply and demand for transport is necessary for carrying out investments to improve service. Transport demand and costs vary, roughly, directly with city size, which implies that transport policy is increasingly important as urbanization proceeds. The impact of poor transport is particularly serious for the poor who commute to work. The study of Colombia in chapter 12 draws attention to the need to focus on cost-effectiveness and affordability in the formulation of transport policy. Of particular interest are the implications regarding efficiency and equity in several areas: public versus private transport, old versus new vehicles, large versus small buses, and government interventions and subsidies that affect these choices.

## Urban Services and the Environment

This volume does not elaborate on several of the important services that contribute to the quality of urban life: education, health care, nutrition programs, family planning, and so on. Many difficulties in these areas stem from the basic problem of urban poverty, and policies that stimulate development would go a long way toward addressing them. Direct programs also have a role. It is important, however, that, as in the case of housing and transport, emphasis is placed on observed demand patterns, affordability, and the cost-effectiveness of policy instruments.

A variety of amenities contribute to urban well-being. Chapter 13 deals with the impact of environmental pollution and the means of improving environmental quality. Air, water, and noise pollution and solid wastes are commonplace problems in the developing-country metropolis. Damages from these sources affect health and property. The chapter provides evidence of the ill effects of air pollution on human health, indicating one type of benefit from pollution abatement.

The control of pollution is costly, however, and cost-efficient policy tools must be devised. Cost estimates for certain types of industries and locations are provided, illustrating criteria for differentiating among sources in antipollution policy. Generally, it pays to differentiate between locations and pollution sources and to reduce

pollution where it is relatively less costly to do so and where the benefits of so doing are relatively large.

### Improving the Efficiency of Shelter Projects

Careful project evaluation has high payoffs in making government interventions, where they are justified, efficient. Chapter 14 summarizes the experiences of four World Bank loan programs designed to improve shelter; the lessons apply to other areas as well. Successful shelter improvement projects have striven for efficient resource use through decentralized decisionmaking. The chapter evaluates project performance and stresses market solutions where they are feasible, on the presumption that project participants are the best judges of their own self-interest.

In general, projects should provide participants with suitable locations, secure tenure, and adequate credit but beyond these should leave most decisions to the participants. The advantages and disadvantages of construction projects, self-help requirements, housing standards, rentals, and restricted credit policies are discussed. Also addressed are the costs of delayed occupancy and inadequate maintenance and the importance of project cost recovery. The findings underscore the importance of carrying out evaluations.

### Bibliography

- Churchill, Anthony A. 1972. *Road User Charges in Central America*. Baltimore, Md.: Johns Hopkins University Press.
- Dwyer, D. J. 1975. *People and Housing in Third World Cities*. London: Longman.
- Linn, Johannes F. 1983. *Cities in the Developing World: Policies for Their Equitable and Efficient Growth*. New York: Oxford University Press.
- Payne, Geoffrey K. 1977. *Urban Housing in the Third World*. Boston: Routledge and Kegan Paul.
- Pfeffermann, Guy Pierre, and Richard C. Webb. 1977. *The Distribution of Income in Brazil*. World Bank Staff Working Paper 356. Washington, D.C.
- Thomas, Vinod. 1978. *The Measurement of Spatial Differences in Poverty: The Case of Peru*. World Bank Staff Working Paper 273. Washington, D.C. Also published in *Journal of Development Economics*, vol. 7 (1980), pp. 85–98.
- . 1982. *Differences in Income, Nutrition and Poverty within Brazil*. World Bank Staff Working Paper 505. Washington, D.C.
- Walters, Alan. 1979. "The Benefits of Minibuses." *Journal of Transport Economics and Policy*, vol. 13, no. 3 (September), pp. 320–34.

## *Intergovernmental Fiscal Relations in Developing Countries*

Roy Bahl and Johannes Linn

The financing and management problems of many of the largest cities in developing countries have begun to attract the serious attention of economists. Several factors explain why these issues have been so long ignored: the traditional concerns of development economists have been macro growth models and the agricultural sector; the urban fiscal problem was small in relation to the financial problems of the central government; aid donors dealt with central governments; the fiscal data of local governments were scanty; and troublesome local government issues seemed better left to administrative specialists. Things have changed, primarily because the fiscal problems of cities have become national concerns and because donors have recognized that the success of capital projects in urban areas is closely tied to the ability of local governments to meet recurrent cost obligations.

This chapter addresses an increasingly important aspect of urbanization in developing countries—the problems and practices of urban government finances. It identifies and analyzes the most important pressures on local budgets, suggests major options for reforms, and sets out constraints on improvements. An important limitation to this effort is the inadequacy of comparable data, as reflected by the paucity of empirical support for these arguments.<sup>1</sup>

### Urban Government Finance as a National Problem

That the financial health of cities is an important national issue for developing countries is borne out by at least five considerations. First, economic growth at the national level is often led by urbanization and the in-

creased productivity associated with it. It follows that if cities are unable to finance services, maintain their infrastructure, or accommodate population growth, national economic growth will be slowed. Second, there is increasing concern about the problems of the urban poor and the need to provide them with improved living conditions. Water, sewerage, primary education, and refuse collection are locally provided services to which the urban poor often do not have adequate access. Important related problems concern issues of how to allocate more resources to the provision of such services, how to distribute a fair share to the urban poor, and whether and how to charge recipients for their use of these services.

Third, the resource mobilization issue has assumed growing importance. In most developing countries the share of GNP that is mobilized for total public sector activities is thought to be too low, and local governments may contribute measurably to increasing it. Urbanization generates increases in taxable capacity that can sometimes be more easily reached by local than by central governments. Increases in property values can be captured by property taxes and growing business activities by business taxes. Some expansions in formal employment can be more easily tapped by local than by central income taxes, and automobile-related charges might be effectively levied by local governments. With respect to the provision of certain services, such as water supply, sewerage, and transport, there is a substantial opportunity for local governments to recapture costs through user charges. It would seem reasonable to assume that increasing the mobilization of local government resources could have a significant effect on national tax effort.

Fourth, external lenders have begun to recognize the

importance of strengthening local government finances. Capital projects have annual operating and maintenance costs, and in some cases the supporting services must be provided. For example, a sites and services housing project must have a water and sewer system, adequate refuse collection, street lighting, road maintenance, and access to primary education services. In many cases the provision of these services lies with the local government, and therefore the success of the entire project depends on the ability of local government to meet the necessary recurrent costs.

Finally, questions have been raised about the optimal distribution of city size. Some have argued that cities have become too large, that the resulting fiscal and management problems are insurmountable, and that centralization of population and economic activity is somehow bad for economic development. Others have countered by noting that productivity advantages are associated with city growth and that effective urban management is possible with the correct institutional framework. Nevertheless, a decentralization movement is under way in many developing countries to increase the fiscal autonomy of local governments and to bring government decisionmaking closer to the people.

### The Urban Fiscal Gap

Urban governments in developing and developed countries alike complain that their resources are inadequate for providing sufficient urban services. The gap between perceived service needs and financial resources (the "fiscal gap") has been interpreted in several ways. The most common explanation begins with the observation that urban populations in developing countries have expanded rapidly in recent years and are likely to continue expanding in the foreseeable future (World Bank 1979). Urbanization has led to rapid increases in expenditure requirements, but revenues have not increased commensurately because local governments are often restricted to income-inelastic revenue sources.

A better understanding of the nature of the problem and the alternative reform possibilities requires some understanding of the components and determinants of the fiscal gap. A useful starting point may be to cast the problem in terms of a set of identities that define the expenditure requirements and revenue constraints of urban authorities. Expenditure needs or requirements for the  $i$ th public service in a particular city may be defined as:

$$(10-1) \quad \hat{E}_i = \frac{\hat{E}_i}{\hat{Q}_i} \frac{\hat{Q}_i}{P} P = \hat{e}_i \hat{q}_i P$$

where  $\hat{E}_i$  is the required expenditure for service  $i$ ,  $\hat{Q}_i$

is the quantity of service  $i$  required,  $P$  is population,  $\hat{e}_i$  is  $\hat{E}_i/\hat{Q}_i$ , the unit cost of the required service, and  $\hat{q}_i$  is  $\hat{Q}_i/P$ , quantity of service  $i$  required for each inhabitant.

The actual level of expenditures for public service,  $E_i$ , may be defined as

$$(10-2) \quad E_i = \frac{E_i}{Q_i} \frac{Q_i}{P} P = e_i q_i P$$

where  $Q_i$  is the quantity of service  $i$  actually provided.

Local government revenues,  $R$ , may be defined as

$$(10-3) \quad R = T + C + G$$

where  $T$  is taxes,  $C$  is user charges and other current revenues, and  $G$  is externally raised revenues.

The fiscal gap,  $\hat{D}$ , in a city may then be defined as

$$(10-4) \quad \hat{D} = \sum_i \hat{E}_i - R = \sum_i (\hat{e}_i \hat{q}_i P) - R.$$

By contrast, the actual budgetary deficit,  $D$ , is

$$(10-5) \quad D = \sum_i E_i - R = \sum_i (e_i q_i P) - R.$$

This formulation clarifies the distinction between the fiscal gap, which reflects the shortage of revenue available to provide required services, and the budget deficit, which reflects the actual shortfall of revenues. Budget deficits do not always occur, but fiscal gaps are commonplace. Equations 10-1 through 10-5 are also useful in organizing the discussion of urban fiscal problems by focusing separately on the expenditure and revenue sides of the local government budget.

### Expenditure Pressures

The expenditure side of the local budget is subject to two sets of pressures: demand or needs which raise  $Q_i$ , and cost factors which raise  $e_i$ . Much has been written about how these factors affect expenditures in developed countries (Bahl, Johnson, and Wasylenko 1980), but less attention has been paid to expenditure determinants in developing countries. One might ask whether the decisionmaking model used to explain urban fiscal outcomes in advanced countries fits the developing-country experience.

The advanced-country model of expenditure determination holds that the fiscal choices of politicians are influenced by the preferences of the median voter, the relative prices of public goods and services, the income level of the community, and the availability of external resources. The developing-country case often differs because voters have less chance to express their preferences; local councils are as often appointed as elected, and the chief administrator of the city may be a central government employee with substantial autonomy. In addition, local government financial autonomy is quite

restricted. For instance, it is common for the central government to place tight controls on local government tax rate changes and borrowing practices and to impose constraining mandates for service levels. Although local fiscal choices in developing countries are more constrained, however, they are by no means nonexistent. Many local councils and mayors are elected; the composition of centrally appointed local councils often reflects local political considerations; and appointed city managers do attempt to take local preferences into account. Even in the most centrally planned and centrally controlled developing countries, the public protests high bus fares or water rates, sometimes effectively resists tax rate increases, and often demands increased public services.

All of this implies that a more constrained version of the traditional maximization model would be relevant for developing countries. To this end it seems necessary to explain the determinants of urban government spending levels in terms of demand-related or need-related factors, cost factors, and the capacity to finance.

#### *Demand or Need Factors*

The demand for local public goods is determined by relative prices, incomes, needs, and taste factors. Tastes, or preferences, affect the rate at which consumers are willing to substitute private for public goods. For example, changing preferences may reflect the demand for better educational services by families whose income has risen above subsistence levels, changing societal values such as substitution of welfare and housing services for the extended family structure, the demand for more redistributive actions to prevent unrest,<sup>2</sup> and willingness to pay more taxes in return for governmental action to offset negative externalities that result from the growing underprovision of urban public services.

*Needs.* An important consequence of urbanization is that public service requirements change and decision-makers may have to interfere with or override individual preferences in providing them. The continuing increase in the numbers of the urban poor calls for increased social and economic services and perhaps for a different package of public services—for example, serviced sites rather than permanent housing, small health clinics rather than hospital additions, more standposts rather than water main extensions, and the like.

The growth in the need for public services is most often associated with increasing population. Some would argue that expenditures must increase at least in proportion to population to maintain even a constant per capita level of service (equation 10-1). For example, water system expansions may involve increasing mar-

ginal costs because of the greater depth required for tubewells or the greater distance to a catchment area. Urban population growth rates in developing countries tend to lie considerably above national population growth rates (table 10-1). Moreover, in some developing countries the rates of growth of the largest cities tend to be even higher than the growth of the total urban population.

*Income Effects.* The positive and strong relation between urbanization and per capita income in developing countries has been well established in two respects: the

**Table 10-1. Urbanization in Selected Developing Countries, 1960–80**

Country	Urban population as percent of total population		Percent of urban population in largest city	
	1960	1980	1960	1980
Low income <sup>a</sup>	15	17	11	13
India	18	22	7	6
Sri Lanka	18	27	28	16
Pakistan	22	28	20	21
Tanzania	5	12	34	50
Zaire	16	34	14	28
Indonesia	15	20	20	23
Sudan	10	25	30	31
Middle income <sup>b</sup>	37	50	28	29
Kenya	7	14	40	57
Ghana	23	36	25	35
Egypt	38	45	38	39
Zambia	23	38	n.a.	35
Thailand	13	14	65	69
Philippines	30	36	27	30
Nigeria	13	20	13	17
Peru	46	67	38	39
Colombia	48	70	17	26
Côte d'Ivoire	19	38	27	34
Tunisia	36	52	40	30
Jamaica	34	50	77	66
Malaysia	25	29	19	27
Korea, Rep. of	28	55	35	41
Algeria	30	44	27	12
Mexico	51	67	28	32
Chile	68	80	38	44
Brazil	46	65	14	16
Iran	34	50	26	28
Argentina	74	82	46	45
Venezuela	67	83	26	26

n.a. Not available.

Note: Countries are listed in ascending order of per capita GNP.

a. Weighted average for all low-income countries given in World Bank (1981).

b. Weighted average for all middle-income countries given in World Bank (1981).

Source: World Bank (1981).

more urbanized developing countries tend to have higher per capita incomes (Beier and others 1975; Renaud 1981; Smith 1974), and per capita income in the largest cities tends to be the highest in the country (Linn 1982). There is less objective evidence on the relation between the increase in urban population and the increase in per capita income in urban areas.

Increasing per capita income tends to increase the per capita demand for services ( $q$ ); the magnitude of the increase is dependent on the income elasticity of demand for locally provided services. Positive income elasticities for urban services have been observed for water supply, electricity, telephone service, and solid waste disposal services (Linn 1982). These higher levels of consumption of utilities may be largely attributable to increased ownership of appliances that use water, electricity, or both (washing machines, radios, televisions, and the like), while the increased need for solid waste disposal may be associated with generally higher consumption levels and reduced recycling in the home.

The demand for motor vehicles is highly income-elastic, which implies that the demand for urban highway infrastructure is positively related to per capita income.<sup>3</sup> The demand for schooling is strongly correlated with household income, since the lower the income the more likely it is that children will be forced to drop out of school to seek employment (Beier and others 1975), and the less able are households to bear out-of-pocket expenditures for education (Meerman 1979). Similarly, the demand for health care increases with incomes, with better education, and with rising familiarity with modern health care techniques.

Per capita income increases also result in a demand for a higher quality of urban services. Higher quality may mean individual rather than communal water supply and sanitary facilities,<sup>4</sup> reduced risk of electric outages, more rapid communication and transport, and better health care, education, and fire and police protection.

*Expectations and Demonstration Effects.* Changing expectations regarding appropriate service levels and quality also increase the demand for urban services. This is especially true for water- and electricity-using appliances, motor vehicles, and educational achievement and therefore for the derived demand for such related urban services as public utilities, road construction, and education. The demand for public services may also be heavily influenced by a demonstration effect from more developed countries. As a result, developing-country governments have frequently raised standards rapidly to attempt to attain the levels of quality and technology found elsewhere. Examples are the often unchecked growth of private automobile ownership, high use of and

high quality standards for water supply, water-borne sewerage technologies to replace such traditional disposal techniques as night-soil collection, and construction of incinerators or composting plants for solid waste disposal to replace conventional recycling techniques. All these can result in significant increases in expenditure requirements (Linn 1982).

*Migration and Poverty.* Increased local government expenditures also result from the locational decisions of poor migrants, who typically swell the populations of cities in developing countries. Often these migrants squat on or purchase at cheap prices land that is difficult to service owing to topography (mountainsides, swamps, flood areas, and so on). As these settlements become relatively well established, the need to service them increasingly becomes a political and humanitarian necessity for urban governments, and the result is substantial expenditure requirements.

### Cost Factors

Local government expenditures may also rise because the cost of providing a given quantity of public services rises—that is, because of an increase in  $e_i$ . Increasing factor costs, particularly for labor, are a prime reason for increases in local government expenditures. Conversely, local government unit costs may fall as population increases if there are economies of scale in the provision of urban public services.

*Unit Cost of Inputs.* Probably the principal factor responsible for unit cost increases is inflation. To the extent that developing countries are plagued with higher levels of price inflation than are industrial countries, local government costs in developing countries will increase more rapidly than those in industrial countries. General inflation can lead to increasing per capita expenditures but does not explain an increasing share of local government expenditure in income. There are, however, factors that tend to raise urban public service input costs more rapidly than the general inflation rate.

First, the provision of public utility services frequently requires investment outlays on a large scale. Urban governments therefore have to rely on foreign credits from international agencies, from the international capital market, or from suppliers. As is well-known, the supply curve for funds is upward sloping, and large cities are likely to run into increasing costs of capital unless the central government is the primary borrower and passes on the loan funds at subsidized rates to urban governments. The extent and terms of borrowing by urban governments are almost always controlled by the central government, and thus the cost of capital as it is

passed on to local authorities is effectively a policy instrument of the central government. The rising cost of capital is reflected ultimately in the public sector as a financial cost, whether at the national or at the local level, and thus the large capital-intensive infrastructure works required by rapid urbanization may well involve higher-than-average interest rates.

Second, the average cost of public employees may also be affected by urbanization, although the direction of this effect is not completely clear. If workers become more productive as city size increases, and if private sector wages respond to productivity increases, there is some possibility for a wage rollout to the public sector. This suggests increasing public sector wages during the urbanization process. Yet the extent to which the competitive wage thesis holds depends on the rate of immigration and, other things being equal, a greater rate of immigration will dampen the wage rollout effect. The local public sector, as part of the general services sector, is labor-intensive and does not require a skilled work force. Newly arrived migrants swell the numbers of unskilled workers available and hold down the wage rate in the services sector.

It may also be argued that wage rates increase more rapidly than does the general price level, owing to institutional factors that determine the wages and salaries of local government employees. Labor unions for local civil servants exist in some countries and can be extremely vocal in pressing for higher wages.<sup>5</sup> In other cases local government salaries and wages are determined by the central government. Since it frequently does not bear the brunt of local government salary readjustments, the central government may be quite willing to raise local civil service salaries more rapidly than the general price level.<sup>6</sup>

Third, land prices are also likely to rise more rapidly than the general price level in cities as a result of rapid urban population growth, increased density, and the resulting scarcity of serviced urban land. Rising land prices tend to have particularly strong impacts on the unit costs of services that are relatively land-intensive, such as the transport sector—which requires substantial amounts of urban land for streets and sidewalks—parks and recreation, schools, and solid waste disposal.

Finally, the unit cost of energy has risen more rapidly than the general price level and may well continue to do so. This can burden larger local governments that tend to rely more heavily on the use of motor vehicles for provision of services and on the use of electricity—for instance for street lighting and for the pumping and treatment of potable water and of sewage.

Confirmation of the hypothesis that input prices tend to increase with the population size of cities is found in frequent observations of price levels that are higher in

larger cities than in rural areas, although these differences are rarely measured accurately. Thomas (1978) found that in Peru the average cost of living in Lima exceeded that in rural areas by a substantial margin. This may be largely explained by differentially higher input prices in Lima, compared with the rest of the country. A similar conclusion for Brazil is presented by Thomas (1982).

*Technological Progress and Economies of Scale.* Technological progress should reduce input requirements per unit of public output and thereby reduce unit cost ( $e_i$ ). It is generally recognized, however, that technological progress in most public services is slow (Baumol 1967) and that some technological advances may lead developing countries to adopt modern technologies which are inefficient and excessively costly. Examples are composting facilities for solid waste disposal, waterborne sewerage systems and treatment plants, limited-access rapid speed highways, subways, and possibly even premature computerization.

Technological economies of scale may also imply declining unit cost during urbanization. A detailed study of engineering costs for certain urban services in India appears to support this hypothesis (Stanford Research Institute 1968). Similarly, a study of water supply costs for small and intermediate-size cities in Colombia has shown declining unit costs (Insfopal 1975), which are probably largely attributable to economies of scale. These economies of scale, however, are likely to be limited to public utilities, and it is not even clear whether governments in rapidly growing urban areas are always in a position to benefit from them. Moreover, economies of scale may be offset by diseconomies of agglomeration, particularly in larger, denser urban areas. Diseconomies result from congestion, which tends to increase with city size and is especially problematic in transport. Diseconomies may also result from limits in the carrying capacity of the natural environment (examples are air and water pollution) and from the increasing scarcity of natural resources, especially water and energy. Other examples are the increased need for disease control and fire and police protection which are associated with the large scale and high density of urban living. In all cases some inputs (space, natural resources, and so on) grow scarce as urbanization proceeds, and larger amounts of other inputs (labor, capital, and intermediate inputs, in particular) have to be applied to maintain service levels and provide clean water and air, good health care, and a safe environment.

*Public Service Employment.* Cost increases for urban services may result if local government is viewed as an employer of last resort. Overstaffing of local gov-

ernment offices is typical, and as a result the financial costs of urban services are inflated. Whether this is desirable as a matter of policy depends on the tradeoff between the gains from additional employment, often at low economic costs (because of widespread unemployment or underemployment of skilled and semiskilled labor), and the financial impact of such a policy.

### Revenue Constraints and Opportunities

Urbanization most often puts pressure on local government budgets by driving up expenditures; at the same time revenues may rise by less than a commensurate amount. Following equation 10-4, total revenues available to finance the growing expenditure requirements of cities may be separated into tax revenues ( $T$ ), other current revenues, including user charges ( $C$ ), and external funding ( $G$ ).

#### Taxes

Total tax revenues of an urban government are determined by a set of factors which may be summarized in the definitional identity

$$(10-6) \quad T = \sum_j \left( \frac{T_j}{L_j} \frac{L_j}{B_j} \frac{B_j}{Y} \frac{Y}{P} \right) = \sum_j r_j t_j b_j y P$$

where

- $T$  = total tax revenue
- $T_j$  = tax revenue of tax  $j$
- $L_j$  = legal tax liability of tax  $j$  for given tax statutes
- $B_j$  = base of tax  $j$
- $Y$  = total personal income
- $r_j$  = collection rate
- $t_j$  = legal tax rate
- $b_j$  = base to income ratio
- $y$  = per capita income
- $P$  = population

Considering the rapid growth in population and income and the relatively high per capita income levels associated with urban growth, the process of urbanization should and does provide a relatively strong taxable capacity for urban governments. Not only do increased population, income, and motorization represent a strong economic base (in relation to the rest of the country), but larger cities also have the potential for more efficient collection. Urban local governments have a better professional expertise than other local governments, and urbanization brings growing formal employment, more cars, and increased property values—all readily identifiable tax handles. Finally, the statutory tax

rate ceiling is usually higher for larger cities. Thus, urbanization may give a substantial boost to local government tax revenues through increases in the values of  $Y$ ,  $P$ ,  $r_j$ , and  $t_j$ .

This favorable impact on local government revenues can, however, be offset by important constraints. Local governments may have little opportunity to capture the increases in taxable capacity associated with urbanization. For example, the central government fixes  $B_j$  by specifying the tax bases available to local governments;  $t_j$ , the tax rate, is commonly fixed by the central government; and in some cases even tax collection,  $r_j$ , is largely a central government matter. In sum, there is no correspondence between the expected increases in expenditures and in revenues in response to a given increase in urbanization. The resulting gap or deficit will vary from place to place and will depend on the public service responsibilities and discretionary fiscal powers of local governments.

#### User Charges

The revenues generated by user charges ( $C$ ) may be represented as

$$(10-7) \quad C = \sum_i \left( \frac{C_i}{Q_i} \frac{Q_i}{P} \right) = \sum_i (c_i q_i P)$$

where  $C_i$  is user charges collected for urban service  $i$ ,  $Q_i$  is the quantity of service  $i$  consumed,  $c_i$  is the unit charge for service  $i$ , and  $q_i$  is the quantity of service  $i$  provided per capita.

In contrast to urban taxes, user charges for services show a direct link between the quantity of services provided and the revenues generated to finance those services. The extent to which user charges cover the cost of provision of services depends, however, on how the average price charged,  $c_i$ , compares with the average cost of providing service,  $e_i$ .

The evidence on the relations between city size and the  $c_i/e_i$  ratio is mixed. On the one hand, autonomous public utility agencies in large cities seem able to charge high enough rates to cover increasing marginal costs and sometimes to generate a surplus. The same result seems to hold for special assessments on urban landowners, for example, in Korea (Doebale 1979) and Colombia (Doebale, Grimes, and Linn 1979). Urbanization thus seems to create a demand for these services and a capacity to pay full costs. On the other hand, some services (notably transport and housing) do not generate enough revenue from charges to cover their full costs. The problem here is that urbanization also generates a great many social costs, such as congestion and pollution, and poverty problems that may dictate holding  $c_i$  below  $e_i$ .

### *External Funds*

Grants and loans ( $G$ ) are not under the control of local authorities but depend on the decisions of higher-level authorities. Whether urbanization brings more grant assistance to local governments depends on how the revenue sharing system operates. For example, if grants are distributed according to population or on the basis of local tax collections, as in a shared tax, urbanization may generate an increased inflow of external resources. The same may be true if grants are made on a cost reimbursement basis.

In the functional relation  $G = G(P, Y, Q_i)$ , one would usually expect to find partial derivatives such that  $dG/dP \geq 0$ ,  $dG/dY \geq 0$ , and  $dG/dQ_i \geq 0$ .

In other words, grants tend to vary directly with city population size and with the amount of services provided under a system of per capita or cost reimbursement grants. External resource flows may increase or decrease in response to increases in per capita income in the city, depending on the grant system. A tax sharing scheme will channel more funds to cities as urbanization proceeds, whereas a formula that equalizes grants across jurisdictions may have the opposite effect.

### *Policy Options for Urban Fiscal Reform*

It may be concluded with some certainty that public expenditure requirements increase with urbanization in absolute terms and probably also in per capita terms. Most often, urbanization also enhances the revenue capacity of urban governments, but revenue growth in most cities has been hampered by the limited revenue authority granted to local governments and by their poor revenue efforts. As a result, revenues have not kept pace with expenditure needs, and service deficits have resulted. There is little reason to expect that this situation will change significantly in the future.

This conclusion raises the question of the constraints on reform. In principle, an urban fiscal gap can be closed in four ways: increased local revenue effort with unchanged revenue authority, increased local revenue authority, increased transfers from higher-level government, or reduced responsibility for expenditures. With these alternatives in mind, the reform question may be addressed by focusing separately on the expenditure and revenue sides of the budget and by noting important interrelations between the two sides.

### *The Reassignment of Functions*

A common response of higher-level governments to city fiscal crises has been to assume responsibility for

certain urban services such as public utilities, roads, education, and health. In addition to arguments of relieving financial pressures, efficiency and distributive concerns are often cited as justifications for such reassignment. Because of the difficulties of measuring the efficiency and equity gains, however, it is difficult to justify such proposals. In any case, two important considerations weigh heavily against this approach to resolving urban fiscal problems. First, the transfer of specific functions to higher-level government reduces the local authority's potential for responding to urban policy issues and problems. This is undesirable because many urban development activities are interrelated and require an integrated approach to planning and implementation, and local authorities are often better equipped to provide such planning than are national ministries or special-purpose agencies. Second, national governments often assume only the responsibility for making capital investments and leave it to the local authorities to operate and maintain the facilities. This turnkey approach tends to burden local authorities with facilities that are often beyond their financial and technical capacity to operate and maintain and that may not reflect local preferences.

### *Increased Local Taxes*

The judgment about which taxes are most appropriately allocated to local authorities depends in part on the perspective of the decisionmaker. From the central government's perspective the main goals are to (1) limit local competition for the important national tax bases (broad-based wealth, income, and expenditure taxes); (2) limit the local use of taxes that are mainly exported to other jurisdictions; (3) provide local authorities with a reasonably buoyant revenue base; (4) avoid local reliance on regressive taxes; (5) encourage the use of taxes that are most easily administered at the local level; and (6) encourage the use of taxes which are closely linked to urban infrastructure and congestion costs, to internalize some of the externalities in the urban economy.

From the local perspective, criteria 3, 4, 5, and 6 are likely to be equally relevant, although they may vary in strength. For example, local governments are likely to place a greater emphasis than higher-level governments on buoyancy and administrative ease and perhaps a different emphasis on equity and efficiency. As regards criteria 1 and 2—competition with national tax bases and nonexport of revenues—local authorities are likely to have priorities exactly opposite to those of higher-level government. Since the broad-based taxes tend to be the more buoyant, and the most easily tapped, local governments desire access to them. Reliance on taxes which can be shifted to taxpayers outside the jurisdic-

tion will naturally also be politically attractive to local governments.

Given these sometimes contradictory goals, it appears that the property and motor vehicle taxes are, on balance, the most desirable and least objectionable of the major tax instruments that could be delegated to local jurisdictions. From the central perspective, they do not compete substantially with national taxes, and tax exporting is likely to be limited, particularly for large cities.<sup>7</sup> From the local perspective, too, these taxes are largely appropriate. Local access to broad-based consumption, income, and wealth taxes is generally not granted by national governments—a reflection of the overriding influence of central government objectives. Exportable taxes, such as selective excise taxes, octroi (a tax on goods brought into a town), and tourism and hotel taxes, are popular among local authorities but tend to receive only mixed blessings from the central authorities or to be prohibited. They are sometimes tolerated mainly because they reduce local governments' claims on national tax resources.

The empirical evidence on the actual use of taxes by local governments in the cities of developing countries is fully compatible with the gist of this discussion (Bahl, Holland, and Linn 1982). The practical difficulty is that the scope for transferring additional taxing authority to local governments may be severely circumscribed by national economic development policies. Rather than expect the allocation of new revenue authority, it may be more realistic to argue for reduced central government limitations on the use of taxes already collected by local governments. Such policies, together with local measures for the improvement of local tax effort, will permit local government to better reach the growth in taxable capacity that comes with increased urbanization.

### *User Charges*

From the national perspective, an effectively administered set of user charges would seem to be the most appropriate source of local government revenue. User charges do not compete with central government revenue bases; they are largely nonexportable; they can have desirable revenue, efficiency, and equity characteristics; and they are administratively feasible at the local level. It may therefore come as a surprise that national governments have sometimes counteracted local authorities' intentions to mobilize more resources through increases in user charges. There are two reasons for these interventions: national governments are concerned about inflation and want to limit the impact of rising local public service charges, and they fear the political repercussions of price increases for urban services, since urban consumers are often quite

emphatic in their opposition to increases and at times even endanger the political stability of the country through riots and the like. Local authorities also have mixed attitudes about raising user fees, partly for the same political reasons that concern the national government and partly because they may have come to rely on central government transfers to finance significant portions of their public service investments.

On balance, however, it appears that local service charges have become increasingly important sources of additional revenue, particularly in cities in which local authorities provide public utility services. In many cases user charges have been utilized effectively by local authorities in lieu of alternative revenue sources and have contributed to the general financing of the local government. It is in the area of user charges that judicious support by central authorities for local revenue raising efforts, in place of the frequently practiced obstructionism, may be most productive in closing the urban fiscal gap.

### *Grants*

An increase in fiscal transfers from the central to local governments is another means of coming to grips with the urban fiscal gap. There are a number of good reasons for developing a system of revenue sharing between different levels of government: the greater administrative efficiency of collecting taxes centrally, the need for the central government to work to equalize the revenues of local governments, and—the traditional justification—the need to provide incentives for efficient local government decisions on resource allocation. For large cities, however, it is unrealistic and inappropriate to expect that transfers will reliably and permanently fill the fiscal gap. The pressures on the central fisc tend to be such that transfers to local authorities are the first to be cut when national austerity programs are implemented.

## **The Politics and Prospects of Urban Fiscal Reform**

Reform proposals to alleviate the fiscal problems of urban governments are commonplace. Although the nature of these proposals has varied from place to place, in line with local conditions and with the makeup of the study team, it is clear from a review of the evidence on urban public finance reform that the proposals have far outnumbered the reforms.<sup>8</sup> Major local government fiscal reforms in the last twenty years have taken place mainly in the industrial countries. Examples are Germany (consolidation of communes and reform of revenue sharing arrangements), Sweden (consolidation of

communes), and the United States (reform of revenue sharing arrangements). Among the developing countries the rule would seem to be that changes in urban finance arrangements come slowly and that it may take decades for fundamental changes to take place, if they occur at all.

Typical for the developing countries are minor and slow adjustments in urban finance practices, such as the creation of special districts for capital cities with special expenditure responsibilities and revenue authority (for example, Manila and Seoul), enlargement of metropolitan jurisdictions by annexation of adjacent municipalities (as in Bogotá), gradual development of new revenue sources (for example, betterment levies in Colombian cities, land readjustment schemes in Korea, and vehicle taxation in Jakarta), gradual reform of existing revenue sources (property taxation in Jakarta and Manila), reassignments of expenditure functions (Kenya and Zambia), and similar gradual and ad hoc responses to urban fiscal pressures. Major reform proposals have often been shelved or taken up only in very minor respects—examples are local government reform proposals in India and property tax reform in Jakarta. Where major readjustments in the fiscal structure of urban governments have occurred, it was where higher-level governments took over important revenue sources previously allocated to local authorities (Kenya and Iran), where sweeping political changes resulted in major shifts of national policy priorities (Nigeria, Tanzania, and Uganda), or where the fiscal problems became so unmanageable that some drastic reform was unavoidable (removal of most important expenditure responsibilities from rural and small-town councils in Kenya). The lessons from the history of urban fiscal reform proposals and implementation are that major proposals rarely have a chance of adoption and implementation and that gradual, stepwise adjustments of the existing structure toward a more desirable state are all that can be hoped for. Gradual adjustments may, in fact, have a better chance of eventual implementation.

There are three major reasons for the inertia typically found in the face of the need for urban public finance reform. First, policymakers and citizens share an antipathy to the uncertain effects of untested large-scale changes in the economic environment. Second, most major reforms are associated with substantial windfall losses to relatively few among the urban population—mostly among the elites—whereas windfall gains are likely to be spread over a much larger number of people—mostly the less well-off. Third, although there has been a growing concern in developing countries about how to strengthen the capability of urban governments to come to grips with their tasks, progress has tended to become bogged down in a three-way debate over the fiscal decentralization issue that typically involves the

ministry of finance, the ministry of local government, and the city governments.

The ministry of finance, which usually is the strongest party, tends to argue in favor of the status quo. It generally refuses to relinquish control over major tax sources or borrowing and argues that decentralization would compromise the central government's important fiscal and tax policy programs. On the expenditure side, the ministry of finance would rather emphasize central government projects and priorities and is often suspicious of the ability of the ministry of local government to regulate the fiscal operations of local governments. It will agree to a grant system but would prefer the grant pool to be decided on a year-by-year basis rather than take the form of a shared tax. In general, the finance ministry looks on local governments as junior partners in the fiscal process, is doubtful about their management and tax collection abilities, and feels that they can get much more out of their existing revenue authorities without recourse to new sources of revenue.

The ministry of local government—in some countries the ministry of the interior—is usually less influential than the finance ministry and is often less well staffed. It usually argues for an extensive grant system and for other regulatory mechanisms which allow a greater measure of control over local government finances. It would prefer that the total grant pool be determined as a fixed share of some national tax and that the distribution of some or all of these grants be at its discretion. Such a scheme would at the same time limit the susceptibility of the grant system to changing finance ministry budget priorities and maximize the extent of the control of the ministry of local government over local authorities.

Administrations of large cities would prefer more independent taxing power and less central regulation of their finances. If there is a grant system, a shared tax based on origin of collection would be the preferred form and a grant pool determined by the finance ministry and allocated by the ministry of local government would be less preferable.

The sense of competition, the suspicion, and the lack of mutual confidence that frequently characterize the debate between national and local government authorities in developing countries have constantly compromised the potential for success of virtually any far-reaching urban government reform. If reforms are to succeed, a mutually supportive system of local and central government relations must be established.

## Notes

1. Perhaps the best comparable fiscal data available for developing countries are those reported in International Monetary Fund, *Government Financial Statistics*. Even this year-

book is more complete for the central government than for local governments; in many cases, local government activity is not reported at all.

2. It might be argued that utility functions become more strongly interdependent during the development process, at least partly because there is more to protect from the dangers of civil unrest. For the basic model see Hochman and Rodgers (1969).

3. See Smith (1974) for evidence that the density of automobile ownership is much higher in the major cities of developing countries than in each country as a whole. Furthermore, Smith and Kim (1979) show that the density of automobile ownership is higher in Seoul and Pusan than in the intermediate-size cities of Korea. Data in World Bank (1975) also indicate a positive association between automobile ownership and per capita income for a cross-section of cities in developing countries (see Linn 1983, ch. 4).

4. Note, for instance, that in the higher-income Latin American countries communal facilities are generally rejected by the population, whereas they have found acceptance in lower-income Asian and African countries.

5. In Colombia, for instance, local teachers and health service personnel were extremely active in the 1970s in pushing for higher wages; they resorted to strikes, protest marches, and sit-ins at the municipal offices.

6. A good example is Korea, where the central government raised local government wages and salaries by 20 percent in 1975 and by a further 30 percent in 1976, both times significantly above the general inflation level (Smith and Kim, 1979). Other examples of centrally decreed salary increases for local officials have occurred, for example, in Kenya, Turkey, India, and Pakistan. In Pakistan, however, the central government offset the increased expenditure requirements by providing a special grant to local authorities. Also, there are documented examples in which local salaries have increased less rapidly than the general price level, as in the case of municipal teachers in Bogotá, where salaries, adjusted for increases in the price level, declined between 1971 and 1973.

7. Note, however, that opposition from groups that are influential in national government circles has been known to hamper the development of effective property and motor vehicle taxation because of its progressive impact, which tends to be borne by precisely these groups.

8. Walsh (1969); Robson and Regan (1972).

## Bibliography

- Bahl, Roy W., Daniel Holland, and Johannes Linn. 1982. "Urban Growth and Local Taxes in Less Developed Countries." In Occasional Paper 57, Metropolitan Studies Program, Maxwell School, Syracuse University. Syracuse, N.Y.
- Bahl, Roy W., Marvin Johnson, and Michael Wasylenko. 1980. "State and Local Government Expenditure Determinants: The Traditional View and a New Approach." In Roy Bahl, Jesse Burkhead, and Bernard Jump, Jr., eds., *Public Employment and State and Local Government Finances*. Cambridge, Mass.: Ballinger.
- Baumol, William. 1967. "The Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis." *American Economic Review*, vol. 57 (June), pp. 415-26.
- Beier, George, Anthony Churchill, Michael Cohen, and Bertrand Renaud. 1975. *The Task Ahead for the Cities of Developing Countries*. World Bank Staff Working Paper 209. Washington, D.C. Also in *World Development*, vol. 4, no. 5 (May 1976), pp. 363-409.
- Doebele, William. 1979. "Land Readjustment as an Alternative to Taxation for the Recovery of Betterment: The Case of South Korea." In Roy Bahl, ed., *The Taxation of Urban Property in Less Developed Countries*, pp. 163-90. Madison: University of Wisconsin Press.
- Doebele, William, Orville Grimes, and Johannes Linn. 1979. "Participation of Beneficiaries in Financing Urban Services: Valorization Charges in Bogotá, Colombia." *Land Economics*, vol. 55, no. 1 (February), pp. 73-92.
- Hochman, H. M., and J. D. Rodgers. 1969. "Pareto Optimal Redistribution." *American Economic Review*, vol. 49 (September), pp. 542-57.
- Infopop. 1975. "Metodología para la formulación del Plan Nacional de Acueductos y Alcantarillados." Bogotá: Instituto de Fomento Municipal.
- Linn, Johannes F. 1980. "Property Taxation in Bogotá, Colombia: An Analysis of Poor Revenue Performance." *Public Finance Quarterly*, vol. 8, no. 4 (October), pp. 457-76.
- . 1982. "The Costs of Urbanization in Developing Countries." *Economic Development and Cultural Change*, vol. 30, no. 3 (April), pp. 625-48.
- . 1983. *Cities in the Developing World: Policies for Their Equitable and Efficient Growth*. New York: Oxford University Press.
- Meerman, Jacob. 1979. *Public Expenditure in Malaysia: Who Benefits and Why?* New York: Oxford University Press.
- Renaud, Bertrand M. 1981. *National Urbanization Policy in Developing Countries*. New York: Oxford University Press.
- Robson, William A., and D. E. Regan. 1972. *Great Cities of the World: Their Government, Politics and Planning*. 3rd ed. London: George Allen & Unwin.
- Smith, Roger S. 1974. "Financing Cities in Developing Countries." *IMF Staff Papers*, vol. 21, pp. 329-88. Washington, D.C.: International Monetary Fund.
- Smith, Roger S., and C.-I. Kim. 1979. "Local Finances in Non-Metropolitan Cities of Korea." Urban and Regional Report 79-1. World Bank, Washington, D.C. Processed.
- Stanford Research Institute. 1968. "Cost of Urban Infrastructure for Industry as Related to City Size in Developing Countries: India Case Study." Stanford, Calif. Processed.
- Thomas, Vinod. 1978. *The Measurement of Spatial Differences in Poverty: The Case of Peru*. World Bank Staff Working Paper 273. Washington, D.C. Also published in *Journal of Development Economics*, vol. 7 (1980), pp. 85-98.
- . 1982. *Differences in Income, Nutrition, and Poverty within Brazil*. World Bank Staff Working Paper 505. Washington, D.C.
- United Nations. 1980. "Human Settlements Finance and Management." Theme paper for the third session of the United

- Nations Commission on Human Settlements, Mexico City, May 6–14. New York. Processed.
- Walsh, A. H. 1969. *The Urban Challenge to Government*. New York: Praeger.
- World Bank. 1975. *Urban Transport*. Sector Policy Paper. Washington, D.C.
- . 1979. *World Development Report 1979*. New York: Oxford University Press.
- . 1981. *World Development Report 1981*. New York: Oxford University Press.

## *Housing Demand in the Developing-Country Metropolis*

Gregory K. Ingram

A policymaker who realistically hopes to cope with urban housing problems should understand housing demand and supply conditions in the city. Effective policy requires knowledge of how markets allocate resources to housing, how homeowners and renters bid for dwellings, how developers and contractors respond to housing demand, and how government regulations and actions stimulate or constrain housing market activity. Our knowledge about many of these phenomena is modest, but our current understanding of housing markets in developing-country cities supports several policy generalizations.

The urban poor in developing and developed countries alike are the most likely to suffer from misallocated housing resources. Slum removal worsens housing conditions for the poor unless affordable low-cost dwellings are provided as substitutes for destroyed units. Often, new units must satisfy quality standards for space, structure, and services, which substantially add to construction costs. Unless cities underwrite large housing subsidies—which are beyond the fiscal means of most city budgets in developing countries—most new construction that meets high quality standards will be affordable only to middle- and high-income families. City managers must recognize that housing goals that are set too high translate into no housing for those who have the lowest incomes.

*Note:* The author has benefited from comments and discussions with Michael Hartley, Stephen Mayo, Janet Pack, Joseph de Salvo, Peter Schmidt, and participants in seminars at Princeton University, Michigan State, MIT, the World Bank, and the Corporación Centro Regional de Población. He thanks Sungyong Kang for research assistance, María Elena Edwards for manuscript preparation, and the staff of the Departamento Administrativo Nacional Estadística, Colombia, for aid with the data.

Developing countries can take several steps to more effectively meet the housing needs of their cities. Most important is the identification of those tasks best accomplished by municipal governments and those best left to private developers and service suppliers in the housing market.

Some services are most efficiently provided by the municipalities. Electricity, water, and sewerage utilities have substantial scale economies that public corporations may be best suited to provide. In addition, public health can be effectively protected only if everyone participates in proper waste disposal and disease prevention programs. Government services provided to all in such cases promote the public good. Local authorities, by establishing and enforcing laws, can also define property rights and facilitate the transfer of property. In developing countries customs regarding definition of property rights may hamper commercial and residential land use. In some developing countries unequal treatment under the law may bar groups from participating in urban development.

Many well-intentioned municipal policies should be avoided because of their perverse effects. Building codes inform tenants of housing quality, but they can be harmful if requirements are set too high. Zoning effectively bars developers from providing low-income housing in many sections of cities. Rent controls discourage prospective developers by reducing future return on housing investment, and existing structures fall into disrepair if the returns from controlled rents do not allow for maintenance expenditures. Similarly, usury laws that prevent banks from charging rates that cover their costs for small loans can effectively bar poor people from formal housing finance and leave them with the options of progressive construction or informal finance

for housing. Municipalities can, however, sponsor credit programs when scale economies are involved and when private financial intermediation is primitive. Risk sharing can extend credit opportunities to those for whom such options previously did not exist.

Private developers should be encouraged to provide for the housing needs of all urban dwellers. With higher incomes from urban employment and with expanded credit opportunities, the urban poor can finance their own housing when standards are set appropriately. Some cities may decide to subsidize urban housing, which would enhance the resources at the command of the poor. Fundamentally, though, sound urban housing policy should be rooted in practices that are self-supporting if housing is not to be a drag on urban development.

The scope for analysis of housing issues is obviously broad, and this chapter addresses only one of many issues. It looks exclusively at urban housing demand and provides quantitative estimates of the factors that appear to determine housing demand. Such quantitative estimates are required to design policies that permit the supply of housing which is affordable for the urban poor.

Three sets of results related to the estimation of housing demand equations are reported. First, estimates of housing demand parameters based on household interview data from Bogotá and Cali, Colombia, are given. A comparison of parameter values with those obtained from North American data sets shows that the Colombian demand elasticities are generally comparable in magnitude to those from the United States. Second, an approach, derived from residential location theory, that stratifies people by place of work is employed to represent housing price variation in the demand equations. Finally, a simple exercise illustrates the magnitude of bias of the income elasticity of demand that can result from incorrect data aggregation techniques. Correctly aggregated data produce income elasticity estimates that are similar to those obtained from disaggregate or micro data.

### The Price Term in Housing Demand Equations

Estimating demand equations for housing from cross-sectional data presents many challenges, but measuring the variation in the unit price of housing is probably one of the greatest difficulties. Data sets typically report the total expenditure on housing rather than a unit price and the quantity of housing. Hence the unit price must be inferred by relating variations in expenditure to variations in quantity. Moreover, housing is inherently multidimensional; it includes attributes of size, dwelling

quality, location, public services, and neighborhood amenities that are obtained in a single, tied purchase. Since there is no widespread agreement as to how to measure the quantity of housing, one analyst's price variation may be another analyst's quantity variation. Finally, even if we can agree that housing prices may vary, it is not obvious that all price variation should be included in a housing demand equation. For example, if a metropolitan area's housing prices vary with the location of housing but households can locate anywhere, we cannot simply put the price actually paid by the household into the demand equation, because the household faces the whole schedule of prices. Simple inclusion of price indexes in a demand equation requires that households be in different market segments.

Numerous approaches have been employed to deal with one or more of these difficulties. Some examples follow.

1. Assume that intrametropolitan price variation does not exist, so that all variations in expenditures reflect variations in quantities; use expenditures in demand analysis as an index number to measure quantities (Muth 1969).

2. Allow intrametropolitan prices to vary across neighborhoods; estimate neighborhood-based price indices; then estimate demand equations on the assumption that residents of each neighborhood face only the prices in their own neighborhood (King 1975).

3. Allow intrametropolitan prices to vary by individual dwelling unit; estimate a dwelling unit price index with the use of a production function for housing and varying input prices; estimate demand equations on the assumption that occupants of each dwelling unit face only the price of their own dwelling unit (Polinsky and Elwood 1979).

4. Allow the marginal cost of attributes to differ within a metropolitan area; estimate a nonlinear hedonic price index and use the first derivative of the index with respect to specific attributes as the price term in a demand equation for the attribute (Witte, Sumka, and Erekson 1979).

Each approach has potential shortcomings. Omission of price variation, as in (1), can bias other demand equation parameters if the omitted price term is correlated with included variables. The assumption that households face only their neighborhood or dwelling unit prices, as in (2) and (3), may lead to a fundamental misstatement of the price variation in the sample if households are not limited in their choices to specific neighborhoods or dwelling units. If all purchasers face all prices, the price "chosen" may reflect the impact of other household characteristics. Price variations based on neighborhood or dwelling unit require that market

segmentation based on those dimensions be justified. Estimation of demand equations for specific attributes of housing, as in (4), may not be relevant if we are really interested in the demand for housing as a composite good.

A relatively simple application of residential location theory suggests an alternative way of incorporating price variation into a demand equation for housing as a composite good. Simple models of residential location theory are essentially based on the precepts of cost minimization. A worker surveys the housing market from his workplace,  $j$ . He typically observes that housing prices,  $R$ , decline with distance,  $d$ , from his workplace in at least one direction, but that travel costs,  $t$ , increase with distance from his workplace. For any given amount of housing,  $H$ , he faces a total expenditure on housing,  $Z$ , composed of a housing expenditure plus a transport expenditure,

$$(11-1) \quad Z_j = R_j(d)H + t_j(d).$$

For quantity  $H_0$  the worker can solve for the least-cost distance by taking derivatives

$$(11-2) \quad Z_j' = R_j'(d)H_0 + t_j'(d) = 0$$

and solving the expression for  $d_j^*$ , the optimal distance or location for quantity  $H_0$  and workplace  $j$ . This least-cost distance can be substituted back into equation 11-1 to calculate the minimum total expenditure for quantity  $H_0$ ,

$$(11-3) \quad Z_j^* = R_j(d_j^*)H_0 + t_j(d_j^*).$$

Consider carrying out this exercise for different workplaces in a metropolitan area. The decline of housing prices with distance,  $R_j(d)$ , differs systematically across workplaces and probably shows steep rates of decline with distance for centrally located workplaces and gradual rates of decline for peripheral workplaces. Travel costs per unit distance may also differ by workplace, but in ways that may be difficult to generalize. For example, transit speeds may be higher but transit headways longer at peripheral locations as compared with a central location. As the workplace varies, however, so does the optimal housing and travel expenditure required for housing quantity  $H_0$ . This variation in expense by workplace for a given quantity of housing is used as a measure of price variation in the housing demand equations estimated here. A price index is estimated for each workplace zone. Households whose heads work at a particular workplace zone face the same housing price index, households with heads at another workplace face the price index at their workplace, and so forth. Prices will vary by workplaces.

If housing prices vary by workplace, it is worth asking why all workers do not try to obtain jobs at the workplace

that has the lowest housing price index. Urban economists have long argued that a metropolitan area with multiple workplaces and a price gradient for housing must have differential wage levels across workplaces for households to be in equilibrium (Moses 1972). Accordingly, workplaces can have different housing prices, but they then must also have compensating differentials in wages to keep households in equilibrium. The existence of wage gradients across workplaces thus becomes a necessary condition for the workplace-based housing price variation approach taken here.<sup>1</sup>

## Housing Demand and Workplace-Based Price Variation

In developing a workplace-based price index for housing, there are two possible formulations for the demand system, each of which uses a different definition of the price of housing. Differing definitions alter the specification of the demand equations. In one formulation the price of housing is based only on the housing expenditure and does not include the travel expenditure. In this case the budget constraint is written

$$(11-4) \quad Y = P_H H + P_v V + t(d)$$

where  $Y$  is income,  $P_H$  the price of housing, and  $P_v$  the price of composite commodity  $V$ . In this formulation the travel expenditure,  $t(d)$ , is included in the income constraint, and the derived demand equation is

$$(11-5) \quad H = f\{P_H, [Y - t(d)]\}.$$

That is, travel costs have to be subtracted from income in the demand equation. If travel costs are an unknown function of distance,  $d$ , then  $d$  is included in the demand equation as a separate variable.

In the second possible formulation the price of housing is the so-called gross price and is based on the housing expenditure plus the travel expenditure. In this case the budget constraint is

$$(11-6) \quad Y = Z_H H + P_v V$$

where  $Z_H$  is the gross price term. The travel cost does not enter separately into the budget constraint, and the distance term will not appear in the demand equations. To implement this second approach, however, one must be able to specify a priori the travel cost function, which is a combination of out-of-pocket cost and the opportunity cost of travel time. Since insufficient information is available for Bogotá and Cali to permit specification of the travel cost function with confidence, the first approach has been implemented here. The estimated demand equations therefore include distance to the

workplace, as in equation 11-5, and the workplace-based price term is based on housing expenditure only.

The relevant housing expenditure that is used to define a price index for a given workplace is the efficient, or optimal, expenditure implicit in the solution of equations 11-2 and 11-3. Corresponding to each quantity of housing,  $H$ , is an optimal location or optimal distance,  $d^*$ , and an optimal expenditure,  $R(d^*)H$ . If households employ the kind of locational calculus embodied in residential location theory, the choice made by households with a head employed at a particular workplace is at or near the optimal location for that workplace, and their housing expenditure approximates the optimal expenditure for their workplace and housing quantity.<sup>2</sup> The relation between housing expenditures and housing quantity for a given workplace can be captured by regressing the observed housing expenditure on measures of housing quantity for households whose heads work in the same work zone. The relation between housing expenditure and housing quantity can then be used to formulate a price index for the given workplace. This procedure can be repeated to yield a price index for each workplace, and these workplace-specific price indexes can then be used as a price term in demand equations for housing as a composite good.

The specific procedure that has been implemented in this chapter can be summarized as follows. The sample contains  $M$  households whose household heads have jobs located at one of  $J$  work zones, and there are  $N_j$  households associated with workplace  $j$ . For household  $i$  ( $i = 1$  to  $N_j$ ) at workplace  $j$ , the monthly expenditure on housing (or the dwelling unit value),  $R_{ij}$ , and a set of  $K$  dwelling unit characteristics,  $X_{ijk}$ , are known. For each of the  $J$  work zones an estimation is made of the equation

$$(11-7) \quad R_{ij} = \sum_{k=1}^K B_{jk} X_{ijk}$$

by regressing housing expenditure on the measure of dwelling characteristics, and  $J$  sets of parameters which indicate how the cost of housing attributes varies by workplace are obtained. A representative dwelling unit in the housing market is then defined as the unit that has the samplewide average amount of each dwelling unit characteristic, where the average quantity is

$$(11-8) \quad \bar{X}_k = \frac{1}{M} \sum_{j=1}^J \sum_{i=1}^{N_j} X_{ijk}$$

The dwelling with attributes  $\bar{X}_k$  then becomes the standard unit—the equivalent for housing of the standardized market basket. For each workplace the estimated parameters in equation 11-7 are used to calculate the cost of the standard unit,

$$(11-9) \quad \bar{R}_j = \sum_k B_{jk} \bar{X}_k$$

The cost of a standardized unit is used to formulate a workplace price index by choosing workplace 1 as a numeraire and calculating a price index

$$(11-10) \quad \Pi_j = \frac{\bar{R}_j}{R_1}$$

The households in the sample also have associated with them  $C$  household characteristics,  $HC_c$ , that affect household demand for housing. These characteristics of the households and the distance from home to work,  $d_{ij}$ , are used in a demand equation whose dependent variable is housing expenditure divided by the price index in equation 11-10, or a quantity index of housing. The demand equation is of the form

$$(11-11) \quad \frac{R_{ij}}{\Pi_j} = f(\Pi_j, HC_{ijc}, d_{ij})$$

and is estimated over the sample of all  $M$  households as a single pooled demand function. In equation 11-11 both linear and double log specifications are used.

## The Setting and the Data

The household interview data used to implement the housing demand procedure just outlined are from Bogotá and Cali, Colombia. The principal data set used was collected in 1978 and covers both owners and renters, for whom equations are estimated separately, in Bogotá and in Cali. A second data set is available for Bogotá in 1972, but data for renters only can be used to estimate the demand for housing in 1972.

In 1978 Bogotá had a population of roughly 3.5 million and Cali a population of roughly 1.1 million. Both cities have experienced rapid rates of population growth in the past (Bogotá's population in 1972 was 2.8 million), but current population growth rates are moderating in both cities. Per capita income in 1978 was about \$800 a year in both places. The cities differ significantly in climate because of their different altitudes: Bogotá is 8,000 feet above sea level and has temperate weather with cool nights; Cali, at 3,000 feet above sea level, is semitropical and warmer than Bogotá. Differences in size and climate may explain some of the differences in housing demand in the two cities.

To implement the workplace-derived price indexes it was necessary to divide the two cities into a number of work zones the boundaries of which are arbitrary but are based on considerations that include compactness, respect for significant internal divisions, and a requirement that there be an adequate number of observations in each work zone. The same work zone system was used in Bogotá in 1972 and 1978 and for renters and owners in each city. Tabulations of residence and workplace by

annular ring and radial sector indicate a high degree of association between place of work of the household head and place of residence of the household. Empirical analyses indicate that although the workplace of secondary workers may have a slight influence on a household's residential location, the workplace of the household head is clearly a dominant determinant (Pineda 1981). Average commute lengths in kilometers for each work zone and tenure type in both cities differ by up to a factor of 3. In both cities commute lengths are long for centrally located workplaces and for workplaces located along the mountains.

### The Hedonic Price Equations

Separate hedonic equations were estimated for each work zone and tenure type in Bogotá and Cali in 1978. For 1972 in Bogotá a hedonic equation was estimated only for renters because no data were available for the value of owner-occupied units in the 1972 sample. For renters the dependent variable is the monthly rent; for owners the dependent variable is the value of the dwelling unit in thousands of pesos.

Because the 1978 data were all collected in the same survey with the same questionnaire, it is possible to use the same specification for the four sets of 1978 equations. In the 1978 equations the independent variables included the dwelling unit area in square meters, *DUAREA*, the lot area in square meters, *LOTAREA*, the number of blocks to the nearest bus line, *BLKTOBUS*, a dummy variable equal to 1 if the residence had a private or public phone, *DPHONACSS*, a dummy variable equal to 1 if the dwelling unit had its own nonshared kitchen and bathroom facilities, *DEXCLUSE*, and a dummy variable equal to 1 if the dwelling unit had its garbage picked up by municipal authorities, *DGARBCOL*. It is interesting to note the similarities and differences between tenure classes and cities in the independent variables. Renters in Bogotá and Cali, for example, have similar-size units on similar-size lots, but Bogotá renters have more phones, while Cali renters have better garbage collection. Bogotá owners have larger, more expensive homes on larger lots than do Cali owners. The most striking differences between renters and owners are in the average area of the unit and the proportion of units having exclusive bath and kitchen facilities; owners are better housed than renters. Finally, there is more variability in the average dependent variable across work zones than there seems to be in the average independent variables, which indicates that prices do vary.

The independent variables used in the 1972 equations differ from those used in 1978 because the questionnaire was quite different. The variables are building age, num-

ber of rooms, presence of garbage collection, distance to nearest bus line, type of structure, condition of unit, presence of nonresidential use, and presence of a public or private phone. Current prices are used in both time periods. The consumer price index had approximately tripled, from 47 in 1972 to 150 in 1978.

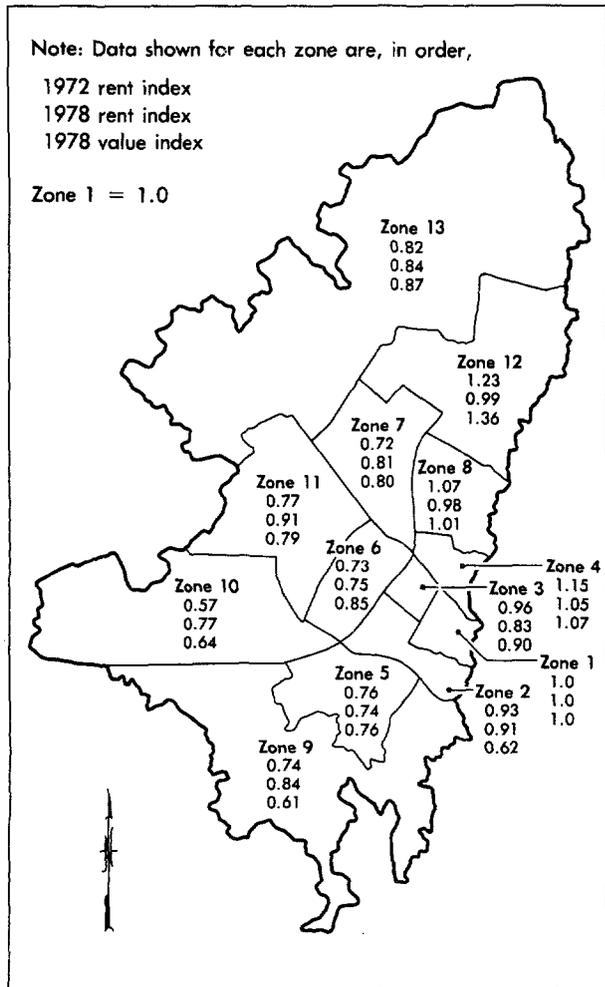
The coefficients from the hedonic price equations are most comparable for 1978. In 1978 there are equations for thirteen Bogotá and eight Cali work zones and for two tenure types, a total of forty-two equations. The only variable that always has the correct sign in all forty-two equations is dwelling unit area. Access to a phone, exclusive bath and kitchen facilities, and garbage collection also perform well; they have the expected sign thirty-six, thirty-seven, and thirty-two times, respectively. The number of blocks to a bus is positive only half of the time; it is possible that there is some disamenity associated with being too close to a bus route. Lot area does not perform well in the hedonic equations, and it does very poorly in Cali, where owners, in particular, do not seem to value larger lot size. The hedonic equations for the 1972 Bogotá renters are similar to those for 1978 in that the measure of interior space, the number of rooms, has a positive effect on rent.<sup>3</sup>

A measure of the explanatory power of the hedonic price equations is shown in table 11-1, which summarizes the explanatory power of the regression equations and the workplace stratification in an analysis of variance framework. Overall, the analysis explains from 45 to 69 percent of the variation in housing prices; the equations have much more explanatory power than does the workplace stratification. Interestingly, the workplace stratification has more explanatory power for owner-occupied units than for renter-occupied units. This is consistent with the empirical regularity that owner-occupied units have steeper price gradients in urban areas than do renter-occupied units. Hence, workplace location matters more in the owner market than in the renter market.

Table 11-1. Analysis of Variance: Hedonic Price Equations

Data	Percent of variation explained by		
	Work zone stratification	Equations	Total
Bogotá			
1972 renters	4.7	49.3	54.0
1978 renters	2.5	47.6	50.1
1978 owners	8.7	36.4	45.1
Cali			
1978 renters	1.9	64.3	66.2
1978 owners	8.0	60.9	68.9

Map 11-1. Bogota: Workplace Price Indexes, Thirteen Work Zones, 1972 and 1976



The standardized rents and values were obtained by plugging the average renter and owner unit characteristics for Bogotá and Cali into their respective workplace hedonic equations. For use in the demand equations these rents and values are transformed into spatial price indexes by dividing through by the relevant rent or value for work zone 1, the central business district. The resulting normalized price indexes are displayed for Bogotá in map 11-1 and for Cali in map 11-2. There obviously is variation across work zones in these price indexes. In both Bogotá and Cali there is more variation in the price index for owners (the range covers a factor of 2) than for renters.

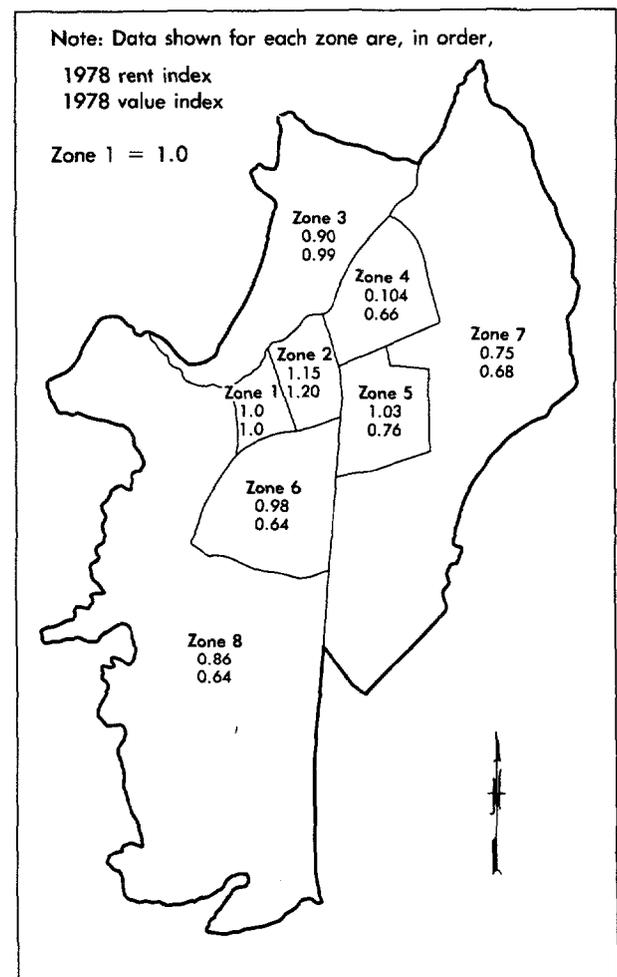
### The Housing Demand Equations

The dependent variable in the demand equations is the monthly rent or value divided by the workplace-specific price index, as shown in equation 11-11. The

independent variables are monthly household income (a measure of current income) in pesos, the price index described above, and the air distance from home to work, in meters. Three additional household characteristics are included in the demand equations: a dummy variable for the sex of the household head (1 = male); family size (the number of persons in the household); and the age of the household head. The hypothesis is that these three characteristics capture differences in taste (sex of household head), in the need for housing (family size), and in assets or wealth (age of the household head). Two functional forms, double log and linear, are estimated. In the linear specifications squared terms for family size and the age of the head are entered to capture nonlinearity in the effects of those variables.

Five sets of equations are estimated; each set is a different combination of year, tenure choice, and city. A comparison of mean values across the five samples shows that renters have younger household heads, smaller families, and lower incomes than owners. Differ-

Map 11-2. Cali: Workplace Price Indexes, Eight Work Zones, 1978



ences between Bogotá and Cali are slight except for income; Bogotá owners have much higher average incomes than do Cali owners, and Bogotá renters have average incomes similar to those of Cali renters. 1978 Bogotá renters had smaller families and younger household heads than did 1972 Bogotá renters.

The demand equations perform well with  $R^2$  statistics that range from 0.25 to 0.6. Income is by far the most important explanatory variable. Age of the head and family size are usually significant; sex of the head is usually not significant, but it always has a negative sign. The housing price index is significant in two of the five samples, and it always has the correct sign. Distance from home to work is significant in four of the five samples and has the correct sign in nine of the ten equations.

A summary of results from the fully specified demand equations, in the form of elasticities for each independent variable, is displayed in table 11-2 for renters and table 11-3 for owners. The elasticities are calculated in the linear equations with the use of the mean value of each independent variable except income. The linear elasticities are shown for the first, second, and third quartiles of each sample's income distribution. In each case the sample mean and the seventy-fifth percentile of the income distribution are essentially identical.

The magnitudes of the various elasticities obviously

vary among the samples and specifications shown, but they display a consistent and stable pattern for most of the variables. All income elasticities are less than one, and at the sample mean they lie in a narrow range of 0.6 to 0.8 except for the equations for Cali renters. The elasticity of the sex of the head is always negative and small (in the interval 0 to  $-0.2$ ). Family size elasticities show an interesting pattern; they are negative for owners and are usually positive for renters. Since renter-occupied units are usually smaller than owner-occupied units, it appears that space is a binding constraint for renters, and larger renter families obtain more housing. Owner-occupants seem to be able to reduce the quantity of housing demand as family size increases because they have larger units on the average and the quantity of housing is not constrained by family size. Age of the household head has a consistently positive demand elasticity when evaluated at the sample mean. By using the linear demand equation with the squared term for age of the head, it is possible to calculate the age at which housing demand is at a maximum. This is consistently ages 50–57 except for Bogotá owners, for whom it increases throughout the relevant range. The price elasticity of demand lies consistently in the interval 0 to  $-1.0$  in absolute value and becomes absolutely quite small for some of the linear specifications. Finally, the distance elasticity is almost consistently negative and quite small.

Table 11-2. Demand Elasticities at Various Income Levels, Renters

Income percentile and level <sup>a</sup>	Income	Sex of head	Family size	Age of head	Price	Home-to-work distance
1972, linear, Bogotá						
25 (1,000)	0.32	-0.16	0.30	0.23	-0.91	-0.05
50 (1,700)	0.45	-0.13	0.25	0.19	-0.75	-0.04
75 (3,079) <sup>b</sup>	0.59	-0.09	0.18	0.14	-0.55	-0.04
1972, log-log, Bogotá						
All	0.77	-0.14	0.14	0.12	-0.70	-0.06
1978, linear, Bogotá						
25 (3,500)	0.55	-0.03	-0.24	0.95	-0.17	-0.23
50 (7,100)	0.71	-0.02	-0.16	0.61	-0.11	-0.15
75 (11,260) <sup>b</sup>	0.80	-0.003	-0.11	0.43	-0.08	-0.10
1978, log-log, Bogotá						
All	0.72	-0.07	0.10	0.07	-0.28	-0.06
1978, linear, Cali						
25 (3,500)	0.05	-0.01	0.48	0.62	-0.34	-0.16
50 (7,300)	0.10	-0.01	0.46	0.59	-0.32	-0.15
75 (12,829) <sup>b</sup>	0.16	-0.01	0.42	0.55	-0.30	-0.14
1978, log-log, Cali						
All	0.47	-0.20	0.36	0.43	-0.48	-0.03

a. Income levels are in parentheses.

b. Sample mean.

Table 11-3. Demand Elasticities at Various Income Levels, Owners

Income percentile and level <sup>a</sup>	Income	Sex of head	Family size	Age of head	Price	Home-to-work distance
1978, linear, Bogotá						
25 (6,000)	0.33	-0.03	-0.34	0.66	-0.31	0.02
50 (10,900)	0.47	-0.02	-0.27	0.52	-0.24	0.01
75 (17,942) <sup>b</sup>	0.60	-0.02	-0.21	0.40	-0.19	0.01
1978, log-log, Bogotá						
All	0.78	-0.09	-0.25	0.25	-0.44	-0.02
1978, linear, Cali						
25 (5,000)	0.39	-0.06	-0.57	0.53	-0.27	-0.06
50 (8,800)	0.53	-0.05	-0.44	0.18	-0.21	-0.05
75 (13,841)	0.64	-0.04	-0.34	0.13	-0.16	-0.04
1978, log-log, Cali						
All	0.76	-0.06	-0.30	0.08	-0.33	-0.02

a. Income levels are in parentheses.

b. Sample mean.

Table 11-4. Range of Housing Demand Elasticities in Various Countries (from Household Observations)

Country	Elasticity of housing demand with respect to				
	Current income	Price	Family size	Age of head	Sex of head (1 = male)
Renters					
Colombia	0.2 to 0.8	-0.1 to -0.7	-0.1 to 0.4	0.1 to 0.6	-0.01 to -0.2
United States	0.1 to 0.4	-0.2 to -0.7	?	?	Consistently negative
Korea, Rep. of	0.12	-0.06 to 0.03	0.15 to 0.25	—	—
Owners					
Colombia	0.6 to 0.8	-0.15 to -0.40	-0.2 to -0.35	0.1 to 0.4	-0.02 to -0.1
United States	0.2 to 0.5	-0.5 to -0.6	?	?	Negative
Korea, Rep. of	0.21	-0.05 to 0.07	-0.02 to 0.15	—	—

? Not conclusive.

— Not reported.

Sources: United States: Mayo (1981); Korea: Follain, Lim, and Renaud (1980).

Table 11-4 summarizes the range of demand elasticities obtained in Cali and Bogotá and compares them with estimates obtained from household surveys in the United States and Korea. The general pattern of results is quite similar for the United States and Colombia; both countries differ somewhat from Korea. The Colombian income elasticities are somewhat higher than those obtained in the United States, whereas the Colombian price elasticities may be lower. The elasticities of housing demand with respect to family size and age of the head in Colombia cannot be compared with data from the United States but are somewhat similar to the Korean estimates. Finally, the effect of the sex of the household head, although usually statistically insignificant in Colombia, is also always negative, as in the United

States. There are three possible explanations for this result. First, female-headed households may have stronger preferences for housing than male-headed households. Second, female-headed households may be discriminated against and face higher prices, which could produce larger expenditures on housing, and those larger expenditures could show up as a preference for larger quantities in the demand equations for renters. The discrimination hypothesis is, however, unconvincing for owner-occupants. Third, female household heads have shorter commuting distances than male household heads and may therefore systematically pay higher prices for housing, since rents decrease with distance from the center of the city. The demand equations used should account for this difference, however,

because distance is included. Accordingly, the first explanation, based on preference differences, may be the most plausible.

To investigate the effect of distance on the coefficient for the sex of the head and to see how sensitive the other parameters are to both the price and distance terms, the housing demand equations were estimated without those terms. Omission of the price and distance terms tends to reduce the income coefficient very slightly, often only in the third significant digit. The family size effects are also affected only minimally by the omission of the two terms. The coefficients for sex and age of head do change considerably in percentage terms, however; this seems to be largely a consequence of the omission of the distance term. Female-headed households live closer to the head's workplace than do male-headed households, as do households with older heads compared with households with younger heads. In general, however, the parameter estimates for the included variables are stable with respect to the omission of the price and distance terms.

These exercises suggest that neither the housing prices as specified in these demand questions nor the distance from home to work are collinear with household income. Indeed, in Bogotá and Cali, as in many other cities, the use of micro data dramatically reduces problems of multicollinearity in the estimation of housing demand equations.

### Aggregate Estimates of Income Elasticities

All of the parameter estimates that have been presented so far have been obtained from computer-based multivariate regressions that use individual households as observations. In many situations it may not be possible to gain access to individual household records because of confidentiality restrictions, while in other situations lack of time or of adequate computer facilities may make parameter estimation with micro data impossible. This section briefly investigates the adequacy of the parameter estimates that could be made from published aggregated data. The focus is on the estimation of the income elasticity of the demand for housing because that parameter is often of interest in both the design and the evaluation of housing programs, policies, and projects.

Each of the five samples analyzed above is now summarized in a matrix dimensioned by rent or value and by income. Eight income categories were defined for the 1978 data and nine for the 1972 data. The average rent or value was calculated for each income category; this average was then regressed on the midpoints of the income categories in a log-log specification using a

hand-held calculator. The equations yielded by this exercise are shown in table 11-5, and the resulting income elasticities are compared with those from the disaggregated, fully specified equations in table 11-6. The aggregate estimates each differ by less than 20 percent from the disaggregated log-log estimates, and in four of five cases the aggregate log-log estimates lie between the linear and log-log disaggregated estimates. It is obvious that aggregate-based estimates of the income elasticities of the expenditure for housing could be a good approximation for the income elasticity of demand for housing in the samples used here.

It is important to note that the aggregate estimates obtained are sensitive to the way in which the underlying micro data are aggregated. An experiment that illustrates this was performed with the 1972 sample of renters. The sample was aggregated to the level of sixty-three zones for the city of Bogotá, and average rents and incomes were calculated for each zone. A hand-held calculator was then used to calculate a log-log regres-

Table 11-5. *Housing Demand Equations from Aggregate Data*

Sample	$B_0$	$B_1$	$R^2$
1972 phase II renter	2.92	0.71	0.99
1978 Bogotá renter	1.54	0.79	0.99
1978 Cali renter	12.38	0.55	0.97
1978 Bogotá owner	9.11	0.67	0.99
1978 Cali owner	7.81	0.66	0.97

Note: The equation is of the form  $R = B_0 \cdot Y^{B_1}$ , where  $R$  is rent and  $Y$  is income.

Table 11-6. *Aggregate and Disaggregated Income Elasticities of Housing Demand*

Sample and specification	Aggregate	Disaggregated
1972 Bogotá renter		
Log-log	0.71	0.77
Linear	—	0.59
1978 Bogotá renter		
Log-log	0.79	0.72
Linear	—	0.80
1978 Cali renter		
Log-log	0.55	0.47
Linear	—	0.16
1978 Bogotá owner		
Log-log	0.67	0.78
Linear	—	0.60
1978 Cali owner		
Log-log	0.66	0.76
Linear	—	0.64

— Not applicable.

sion of average zonal rent on average zonal income, using all sixty-three observations. The resulting income elasticity, 0.95, was substantially higher than the 0.71 estimate that was obtained with nine observations from the correctly aggregated sample. A second experiment was then run on the 1972 Bogotá data. For this experiment the data in the rent-income matrix were *incorrectly* aggregated by calculating the average income for each rent category and regressing the rent category midpoints on the mean incomes. This rent-stratified approach yielded an income elasticity estimate of 1.36, nearly twice the 0.71 obtained with the use of an income-stratified aggregation procedure. It is obvious that the aggregation bias in estimates of income elasticities can be very large, but that correctly aggregated data can give useful results.

## Conclusion

This paper has described and implemented a two-step estimation procedure for incorporating price variation in the estimation of demand equations for housing, using household survey data from Bogotá and Cali, Colombia. The demand equations estimated with this procedure give significant results for the income elasticity of the demand for housing; estimates of the income elasticity generally lie in the upper end of the 0.2–0.8 range. Although the price term in the demand equations gave less significant results, the price elasticity of demand appears to be between 0 and  $-1$ . There is, however, greater uncertainty about the magnitude of the price elasticity than about the magnitude of the income elasticity. Other household characteristics involved in the demand equations have low demand elasticities, typically less than 0.5 in absolute magnitude. The age of the head has a positive elasticity over most of its range, while family size usually has a positive elasticity for renters and a negative elasticity for owners. The demand equations suggest that female-headed households consume more housing than male-headed households, but this result is rarely statistically significant. Distance from home to work is entered into the demand equations as an adjustment to income, but it undoubtedly also represents price variation within the workplace strata, which are used as the main representation of price variation. The distance elasticity is small—less than  $-0.2$ —and is almost always negative.

Comparisons of elasticity estimates with those obtained from U.S. data sets indicate that the range of the Colombian estimates generally overlaps the range of the U.S. estimates. This similarity of values may seem surprising at first but is much less so on reflection. Housing is a nontraded good, and its price is endoge-

nous to the local economy and reflects, among other things, local income levels. Perhaps we should be more surprised at the similarities between Bogotá and Cali, which have markedly different climates.

Simple experiments involving the aggregation of the household survey data used to obtain micro data estimates suggest that income elasticity estimates on the basis of correctly aggregated data can be good proxies for estimates on the basis of fully specified models that use household observations. Estimates on the basis of incorrectly aggregated micro data, however, can produce estimates of the income elasticity of demand that are badly biased.

## Notes

1. Preliminary empirical work indicates that a wage gradient with a peak in the central business district does exist in Bogotá.
2. Equation 11-3 can be solved for the expansion path of expenditures as the quantity of housing increases.
3. Parameter estimates are available from the author.

## Bibliography

- Follain, J., G. C. Lim, and B. Renaud. 1980. "The Demand for Housing in Developing Countries: The Case of Korea." *Journal of Urban Economics*, vol. 7, no. 3 (May), pp. 315–36.
- King, Thomas. 1975. "The Demand for Housing: Integrating the Roles of Journey-to-Work, Neighborhood Quality, and Prices." In N. Terleckyj, ed., *Household Production and Consumption*, pp. 451–83. New York: Columbia University Press.
- Mayo, Stephen K. 1981. "Theory and Estimation in the Economics of Housing Demand." *Journal of Urban Economics*, vol. 10, no. 1 (July), pp. 95–116.
- Moses, Leon N. 1972. "Toward a Theory of Intra-Urban Wage Differentials and Their Influence on Travel Patterns." *Papers and Proceedings of the Regional Science Association*.
- Muth, Richard F. 1969. *Cities and Housing*. Chicago: University of Chicago Press.
- Pineda, José Fernando. 1981. "Residential Location Decisions of Multiple Worker Households in Bogotá, Colombia." Presented at annual meetings of the Eastern Economic Association, Philadelphia, Pa., April. Processed.
- Polinsky, A. Mitchell, and David M. Elwood. 1979. "An Empirical Reconciliation of Micro and Grouped Estimates of the Demand for Housing." *Review of Economics and Statistics*, vol. 61, no. 2 (May), pp. 199–205.
- Witte, Ann D., Howard J. Sumka, and O. Homer Erekson. 1979. "An Estimate of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets." *Econometrica*, vol. 47, no. 5 (September), pp. 1151–73.

## *Urban Transport Policy: Colombia*

Alvaro Pachón

Our knowledge of urban transport in Colombia comes in great part from transport and urban finance studies made during the past thirteen years in the big cities and especially in Bogotá. Discussions have centered on two main themes: public versus private transport, and choices in public transport.

### Operations

The less efficient use that private vehicles make of road space and fuel per passenger has been identified as one of the causes of the transport problem (see World Bank 1975, pp. 84–85; Urrutia 1981, pp. 12–13). To improve resource allocation, two kinds of measures have been suggested: taxes and subsidies to shape the transport market, and quantitative restrictions to limit use of the roads.

A proposal that relies on market mechanisms is the use of a fuel tax to control urban congestion (Republic of Colombia 1973, pp. 99–111). McClure (1974) has proposed a parking tax. Although the government has not accepted either proposal, it has adopted some measures that clearly favor public over private transport, including sales and import tax exemptions for public transport vehicles, bus fares that are set below average cost, and lower fuel prices for commercial vehicles. The main restriction designed to favor collective transport is the prohibition of private vehicles on Carrera 10, one of Bogotá's principal streets. (In contrast, the local transport authority in Cali has set restrictions on bus access to the traditional center.)

The considerable amount of investigation that has been devoted to the subject of collective transport is

explained by the high percentage of trips made on public transport and the low rate of private motorization (see Pachón 1981c; Westin 1980). The main elements of what is known about collective transport in Colombia are summarized below.

### *State Intervention in Collective Transport*

The state has recognized collective transport as a public service and a basic need. Inability to charge directly for road use leads to excessive use of private vehicles and inefficient use of road space, and additional inefficiencies are created in a system of free competition when drivers do not consider the negative effects of their behavior on other vehicles on a congested road. Collective transport can reduce some of these problems. As for transport as a basic need, under free competition, low-income consumers may not be served when the cost of providing service is above what they can afford, and state intervention may be warranted. Intervention may be direct, through production or pricing of the basic need, or indirect, through income transfers. As high-income classes seldom use buses, direct intervention is likely to use less resources to meet the basic need objective and hence is preferable to indirect intervention.

### *Organization of Collective Transport*

There are two types of ownership of collective transport firms: ownership of the vehicles by the municipal government—the predominant arrangement in the United States—and ownership by private organizations, as in Argentina, Brazil, Chile, and Colombia. Colombia has a basic system of affiliating companies which coexist

with state transport companies and cooperatives. In Bogotá's system of affiliating companies the state assigns the routes and gives some rights to the private company. In some cases the company owns some of the buses, but generally bus owners affiliate with a company which distributes the routes and charges its associates a rolling charge for use of its routes. The managers of this type of organization are the strongest group in the transport system because they maintain a close relation with the official sector, through which they obtain operating licenses and route authorizations.

In the state companies, the state administers and owns the transport equipment, as in the case of the *Empresa Distrital de Buses de Bogotá* (District Bus Company of Bogotá). This company's importance has decreased; in 1980 it accounted for less than 1 percent of public transport buses and *busetas* (minibuses). Cooperatives, which own their transport equipment, with the associated drivers having shares, predominate in some cities, such as Cali, and occupy a second place in others, including Bogotá.

In 1980 Bogotá had thirty-eight public transport companies, of which twelve were cooperatives, one was state-owned, and twenty-five were limited liability companies (affiliating companies). In Cali in 1979 there were 1,426 buses distributed among sixteen private companies with forty-two routes and 181 *busetas* distributed among three companies with seven assigned routes. There has been considerable debate about the organization of public transport. Defenders of a public collective transport system have pointed out the inefficiencies in resource allocation associated with the existing system in Colombia, while the defenders of the present system emphasize the low costs of private companies. In Leibenstein's terminology (1966), the defenders of a state system focus on efficient resource allocation, while the defenders of the present system focus on the inherent X-efficiency of private property. Interconsult Ltda. (1970) and Urrutia (1981) have documented the existing system's low operating costs, the good maintenance level of the vehicles, and the system's ability to adapt to changes in the demand for transport. A 1980 study by the Administrative Department of District Planning shows the high levels of X-inefficiency of Bogotá's public company. Walters and Feibel (1980) and Urrutia (1981) also show that in other parts of the world private companies have lower cost levels and are more X-efficient than public companies. As an exception, Nelson (1972) shows that in the United States in 1960 and 1968 public companies had costs 10 percent lower than those of private companies.

The inefficiency of resource allocation, especially as concerns congestion, has been documented by, among others, Interconsult Ltda. (1970) and Departamento

Nacional de Planeación—Instituto Nacional de Transporte (INTRA) (1974). By means of a route simulation model, these studies bring out the savings, in operating costs, of a centralized collective transport system.

Economic evaluation of the two types of organization requires a simultaneous analysis of the two types of efficiency that takes into account the variations state ownership introduces in the operating costs and considers the benefits and costs associated with trip duration and waiting and transfer times. In addition, political considerations are relevant. Which type of organization better promotes the satisfaction of basic needs? Which type is more reliable and less affected by strikes? (Although strikes are prohibited in government services, slowdowns are often used for the same purposes.)

### *The Route System*

The route systems in Colombia and especially in Bogotá have been frequently studied. Cifuentes (1978) notes that authority to assign routes has been delegated by INTRA to the mayoral offices in the principal cities. The resulting separation of functions has made control of the public transport system more difficult.

To assign new routes, or to extend existing routes, the Administrative Department of Transit and Transportation (DATT) studies the potential demand and opens a tender to assign the new route. The tender specifies the number of vehicles required and the frequency of service during peak and nonpeak hours. The assignment process has been much criticized. It has been argued that the system does not respond to the real needs of the population because the decisions respond to pressure by the *Juntas de Acción Comunal* (community action boards) of the districts and do not take into consideration the needs of the whole population. It has also been argued that the prevailing system of routes and route assignments produces high levels of congestion in the central area because all the companies prefer to serve areas where demand levels are high. The existence of parallel routes and the information problem caused by the great number of routes have also been criticized. To remedy some of these problems the Interconsult Ltda. and INTRA studies proposed the consolidation of routes on the basis of the results of a simulation model.

It has, however, also been established that the existing route assignment system in Colombia produces a network with wide coverage that has adapted to changes in trip demand patterns. The efficiency with which owners, drivers, and users process the existing information to obtain a satisfactory allocation has been noted.

Discussion of the route system in Bogotá leads to some important conclusions. First, a series of incentives is bringing about a satisfactory solution. The urban

transport market is processing a great deal of information at a low cost. If there were no externalities, the result produced by the market would be the same as would be obtained in a model for the whole transport system. Second, given the existence of such externalities as congestion and buses' stopping on demand, the present system of decentralized decisionmaking may be faulty. Third, since there is no charge for road use in the central area, the income from scarce road space is transferred to the private sector, probably to the affiliating companies. Fourth, although the existing data allow the use of simulation models developed for Colombia, it is probable that more effort is needed to collect and analyze data and to refine the models if their results are needed to assign a route system. Fifth, to ensure good service to areas with difficult topographic conditions and high operating costs, it is necessary to specify differential fares or subsidies for routes that serve those areas. Sixth, if the state established a fee for the right to operate routes that serve the central areas, resource allocation could be improved and funds could be generated to provide subsidized service to the low-income districts that have unfavorable cost conditions.

#### *Public Service Vehicles*

The initial investigations by Interconsult Ltda. and INTRA considered it desirable to define the technical characteristics that would assure more efficient and comfortable service. It was thought that metropolitan buses like those in the United States should be used and that the school-bus type of vehicle should be taken out of circulation. The success that minibuses and busetas were having at the end of the 1960s was acknowledged, however, and was attributed to their greater speed, privacy, and comfort.

In Colombia as in other developing countries it was soon recognized that although the metropolitan bus had favorable technical attributes, the system of school buses, minibuses, minibuses, and collective taxis offered many more advantages (see Owen 1973; Walters 1979). First, a small vehicle allows more intensive use of the most abundant factor in developing countries, labor, and because the level of investment is lower, vehicles are easier to obtain, more drivers can own their own vehicles, and there are greater incentives to maintain vehicles in good condition. Second, conditions on the demand side make the use of low-capacity vehicles attractive: because the waiting and travel times are shorter and there are fewer stops, this type of service is more like that of a private vehicle and is more attractive to middle- and high-income groups. All these arguments have contributed to legitimize the persistence of diverse vehicle sizes.

To balance supply and demand for urban transport, the state has enunciated policies which are sometimes contradictory. For example, it has established a financial intermediary, the Financial Transportation Corporation (CFT), with credit lines for financing body work and chassis for buses and busetas. This credit has been subsidized to stimulate investment in public service vehicles and to facilitate the purchase of vehicles by the drivers. The subsidy, which amounts to nearly a fifth of the vehicle's value, has undoubtedly made investment in public vehicles more attractive.

The state also wants to maintain an adequate transport supply and assure an adequate income yield for new vehicles. Differential fares were established during the late 1960s according to the age of the vehicle, with higher fares for newer vehicles. Since this system did not work, a differential subsidy by age was tried. The existing subsidy system does not fulfill the initial objective of favoring investment in new buses.

Moreover, the government has established a license—known in Bogotá as a *nota opción*—for new public service vehicles. This entry restriction has artificially limited the size of the vehicle stock and has created some income for those who already own public service vehicles. Also, because in some cities the procedure for granting licenses is stricter for small vehicles, the mix of vehicles in the stock is artificially distorted.

#### *Remuneration of Drivers*

The public companies pay drivers a monthly salary for a normal eight-hour day; the private companies pay a commission according to the number of passengers transported. The private companies' system has been criticized because it may lead drivers to compete to pick up passengers without regard to traffic laws. It has also been pointed out that the driver is being exploited and has long working days and unfavorable working conditions. The public companies have also been criticized: the fixed salary, which does not take into account the number of trips or passengers, destroys the incentive to pick up passengers and to maintain the vehicle in optimum condition.

The private system appears to decrease the information costs required for efficient operation. That is, since the driver has an incentive to pick up as many passengers as possible, less planning and monitoring are required to ensure that bus schedules and operations are meeting passengers' needs. This does not mean that a fixed salary scheme cannot establish information and control systems for efficient operation.

The incidence of the subsidy is often misunderstood. The user obtains an economic benefit from the lower fare or better service even though the subsidy is given to

the bus owners. The short-term effect is also sometimes confused with the long-term effect. In a system with no restrictions on the entry of new buses, an excessive subsidy can lead to large temporary profits for the vehicle owners. Free entry ensures, however, that in the long term automotive stock will increase, the service level will improve, and profits will become normal. Finally, transfers are often confused with economic costs. A transport subsidy is a transfer. Even though a subsidy decreases consumer surplus—because resources are being used at a higher value than what consumers are willing to pay—a public transport subsidy tends to correct the inefficient resource allocation caused by the absence of a price system for use of road space.

The absence of empirical research on the economic effects of government intervention in the transport sector is beginning to be remedied. Pachón (1981b) has found that the effect of a transport subsidy is essentially redistributive; the benefits reach mainly lower-income households. The maintenance of low fuel prices has also contributed to a better income distribution by decreasing the buses' operating costs. A preferential subsidy for bus users has fewer leakages than a general subsidy for all public transport users. Further research is needed to quantify the effects of transport subsidies.

### *Fares and Subsidies*

While the managers press the government for increases in fares whenever they consider that their investment yield has reached unsustainably low levels, workers and students protest when bus fares are increased. In an effort to lower costs to the public, the government establishes a fare below the cost of service, maintains low prices for gasoline and diesel fuel, controls the prices of such inputs as tires, and establishes subsidized credit lines through the CFT. To guarantee an adequate profit, it specifies a subsidy for bus owners and sets unsubsidized fares for means of public transport other than buses.

Some operational parameters have been defined, and an index of input costs is maintained that allows the government and the bus owners to recognize when periodic adjustments to public transport fares are needed to reflect increases in costs. Thus, adjustments were made in response to increases in world fuel prices. There is great confusion, however, regarding the economic effects of such measures. Transport subsidies are sometimes blamed for an exaggerated urban dispersion, but in a flat fare system, as in Bogotá, the marginal cost of a kilometer traveled is zero no matter what the fare, and since decisions on residential location depend on the marginal cost of transport, a transit subsidy has no effect on housing location. Thus, the cause of urban

dispersion is the flat fare, not the subsidy, and dispersion can be reduced by making the fare vary according to distance traveled.

### *New versus Old Buses*

Preferential treatment for more recent models was initially provided through differential fares by age of vehicle and, more recently, through differential subsidies by model. In October 1980 monthly subsidies were 42,812 pesos for 1974-and-later models, 30,443 pesos for 1970–73 models, 26,687 pesos for 1965–69 models, 25,811 pesos for 1960–64 models, and 24,917 pesos for 1959-and-earlier models. The CFT also favors the financing of new vehicles in granting credit.

### *The Rationale for Differential Treatment*

Differential fares and subsidies by age are defended as a means of stimulating the demand for new domestically produced vehicles. Furthermore, if it is desirable to augment the capital stock rapidly, preferential treatment for new vehicles is helpful. The operating costs of vehicles of different ages can also be relevant: to the extent that the operating cost of a new vehicle is less than that of an old vehicle, the former will be used more intensively.

There are also reasons for giving equal or similar treatment to old and new vehicles. Only 7.5 percent of the stock of vehicles in Bogotá is less than two years old. The high proportion of old vehicles with lower capital costs offers drivers the possibility of owning vehicles and thus broadens the base of ownership. A balanced treatment for used buses could be defended, since it promotes capital formation, widens to some degree ownership of the means of production, and allows people whose opportunity cost of time is lower to maintain older vehicles.

### *Operational Costs, by Age of Vehicle*

The Interconsult Ltda. study found no statistical difference in operating costs between new and old vehicles, but since such variables as routes served, distance traveled, and drivers were not isolated, more analysis is needed. Merewitz (1977) cites two studies which show that the operating costs of a company are independent of the average age of the fleet. As in the Interconsult Ltda. study, the estimation procedure and the data used have limitations.

A regression analysis of costs and revenue in relation to age has been carried out with the use of assumptions for the estimation of operating costs by ranges of bus

models and data on the monthly operating costs for different bus models as estimated for October 1980.<sup>1</sup> The following relations were calculated with the use of ordinary least squares.

$$(12-1) \quad \ln Y_a = \ln Y_0 + Ba$$

where  $Y$  is the variable whose behavior is to be analyzed,  $a$  is the vehicle's age, and  $\ln Y_0$  and  $B$  are the function's parameters.

The number of trips made, which varies inversely with age, influences the number of kilometers traveled and the monthly cost. To isolate the effect of trips made, the variable cost per kilometer ( $VCK$ ) was calculated and a regression was run, with the following results.<sup>2</sup>

$$(12-2) \quad \ln VCK_a = 27.925 + 0.0024 + a. \\ (4.807) \\ R^2 = 0.9001$$

Although the  $VCK$  changes very little (it decreases 1 percent for every four years of age), the hypothesis that it is independent of age cannot be accepted because the  $t$ -statistic value is greater than the critical value.

The results of monthly cost and revenue regressions were as follows, where  $VC$  is variable cost,  $FC$  is fixed cost, and  $ING$  is operating revenue. (The subsidy is not counted as revenue.)

$$(12-3) \quad \ln VC_a = 11.3088 - 0.01806 + a \\ (8.37) \\ R^2 = 0.9589$$

$$(12-4) \quad \ln FC_a = 8.8097 + 0.02771 + a \\ (14.13) \\ R^2 = 0.9852$$

$$(12-5) \quad \ln ING_a = 11.2157 - 0.02167 + a \\ (6.65) \\ R^2 = 0.9365$$

The results show that both monthly cost and revenue are negatively affected by age and that the coefficient of the age variable is statistically different from zero. According to these results, monthly income decreases more rapidly than monthly variable costs, and the intercept of the variable cost function is slightly greater than the corresponding value of the income functions. Thus, operational revenue does not cover the variable costs, and the gap between revenue and costs increases in percentage terms as the vehicle gets older because income decreases faster than costs.

Table 12-1 presents results for each model between 1959 and 1980. The adjusted values were calculated using equations 12-3, 12-4, and 12-5. To obtain capital return, the estimated income and the subsidy received were added and total costs were subtracted.

The evidence on operating costs suggests several con-

clusions. First, variations in the operating costs per kilometer of buses of different ages are relatively small. Second, because the monthly operating revenue by age decreases faster than the monthly variable costs by age, older buses tend to be less competitive. Third, the number of monthly trips decreases with age—new buses seem to be used more intensively. Fourth, additional research is needed on the cost and income structure by age of the vehicle, with the effect of such other variables as route conditions, driver characteristics, and the make of the vehicle isolated.

### Prices and Profitability

For used vehicles the present value of future income is basic to the price determination and in turn to evaluation of the effects of fare and subsidy policies. The price pattern of buses by age is satisfactorily approximated by an exponential functional form for cars in Colombia (Pachón 1981a). Data on the value of the vehicles have recently been assembled by the Center for Economic Development Studies (Centro de Estudios para Desarrollo Económico, CEDE) in Colombia, and these data are used to estimate the following form.

$$(12-6) \quad \ln P_a = \ln P_0 - b * a$$

where  $P_a$  is the price of a vehicle of age  $a$ . The results of the regression are

$$(12-7) \quad \ln P_a = 14.3961 - 0.06887a \\ (141.1) \quad (8.28) \\ R^2 = 0.7742 \\ 20 \text{ degrees of freedom}$$

The  $t$ -values are in parentheses.

The interpretation of the age variable coefficient in a regression of the price logarithm against the age of the vehicle is the yearly depreciation rate. According to the results given here, for each year of use the price of the bus decreases by slightly less than 7 percent. The value obtained in equation 12-7 is less than that used in the cost studies made by the transit companies, which show a useful life of seven years and a salvage value of 30 percent.<sup>3</sup> Thus the depreciation rates chosen in the studies on defining the fares tend to produce real profit rates higher than those agreed on between the government and the bus owners.

The economic profitability of used vehicles is affected by three sources of economic rent: depreciation (the loss in value owing to age), the opportunity cost (the monetary interest rate multiplied by the price of the good), and the appreciation of the good because of inflation and changes in relative prices. Given estimates, by age, of depreciation and the price of vehicles, a nominal profit-

Table 12-1. *Buses: Adjusted Costs, Estimated Income, and Capital Return, October 1980*  
(pesos)

<i>Model</i>	<i>Age</i>	<i>Estimated variable costs<sup>a</sup></i>	<i>Adjusted variable costs<sup>b</sup></i>	<i>Total costs</i>	<i>Estimated income<sup>c</sup></i>	<i>Subsidies</i>	<i>Capital return</i>
1959	21	55.800	3.743	59.543	47.126	24.917	12.400
1960	20	56.817	3.848	60.665	48.158	25.811	13.303
1961	19	57.852	3.957	61.809	49.213	25.811	13.215
1962	18	58.907	4.068	62.975	50.292	25.811	13.127
1963	17	59.980	4.182	64.163	51.393	25.811	13.041
1964	16	61.074	4.300	65.374	52.519	25.811	12.956
1965	15	62.187	4.420	66.608	53.670	26.687	13.749
1966	14	63.320	4.545	67.855	54.845	26.687	13.667
1967	13	64.474	4.672	69.147	56.047	26.687	13.586
1968	12	65.649	4.804	70.454	57.275	26.687	13.508
1969	11	66.846	4.939	71.785	58.530	26.687	13.431
1970	10	68.064	5.077	73.142	59.812	30.443	17.112
1971	9	69.305	5.220	74.525	61.122	30.443	17.039
1972	8	70.568	5.367	75.935	62.461	30.443	16.968
1973	7	71.855	5.517	77.372	63.829	30.443	16.900
1974	6	73.164	5.672	78.837	65.228	42.812	29.202
1975	5	74.498	5.832	80.330	66.657	42.812	29.138
1976	4	75.856	5.996	81.852	68.117	42.812	29.077
1977	3	77.238	6.164	83.403	69.609	42.812	29.018
1978	2	78.646	6.337	84.984	71.134	42.812	28.962
1979	1	80.079	6.515	86.595	72.693	42.812	28.909
1980	0	81.539	6.698	88.238	74.285	42.812	28.859

- a.  $VC_a = 81,539 [\exp(-0.0180^a)]$   $R^2 = 0.9589$   
 b.  $FC_a = 6,698 [\exp(-0.0216^a)]$   $R^2 = 0.9852$   
 c.  $INC_a = 74.285 [\exp(-0.0216^a)]$   $R^2 = 0.9365$

ability rate can be established once the rate of increase in bus prices is known. Estimated prices, capital return (current revenue minus current cost), and profitability yielded by the calculation are presented in table 12-2. The estimated values are based on regressions in which the only explanatory variable is the vehicle's age and which therefore may differ from real values.

The estimated profitability ( $i_a$ ) shown in table 12-2 is inversely related to age. This finding is quite unexpected. Under competitive conditions, profitability might be considered to be independent of age. A casual observer, however, might expect higher profitability for the new buses, since larger owners tend to own newer buses. Two possible explanations may be offered for the results in table 12-2. First, the difference could be a result of an underestimation of the operating costs of old models. The maintenance costs of old vehicles exclude the value of the time of the driver, who often owns and repairs the bus. Second, the difference could be a consequence of an underestimation of the price of older vehicles. Because the estimated prices are calculated on the basis of a regression whose only explanatory variable is the vehicle's age, these prices cannot capture the distortions caused by the subsidy and thereby tend to underestimate cases in which the subsidy is unneces-

sarily high. As this occurs for the older buses, the adjusted prices underestimate the prices of the older vehicles and consequently overstate their profitability.

To show that many of the distortions in prices and profitability are a result of the subsidies, data for busetas, which are not subsidized, were used. With the use of the buseta price data from the CEDE survey, an exponential function was estimated:

$$(12-8) \quad \ln P_a = 14.0933 - 0.06943 * a$$

$$(141.1) \quad (8.63) \quad R^2 = 0.8515$$

14 degrees of freedom

In table 12-3 the buseta's profitability is nearly the same for 1970-and-later models; for previous models profitability decreases with age. Thus, in the absence of a subsidy the estimated profitability is more in line with an expectation that profitability is independent of age.

### Buses or Busetas?

Government policy appears to have favored busetas, as is reflected in their rapid growth. Between 1971 and 1976 the number of buses is estimated to have grown by

**Table 12-2. Buses: Prices, Capital Return, and Profitability**  
(pesos)

	Model	Age	Price in 1980	Estimated price <sup>a</sup>	Capital return <sup>b</sup>	Profit- ability <sup>c</sup>
	1959	21	200,000	420,759	12,499	44.15
	1960	20	540,000	450,758	13,303	43.92
	1961	19	640,000	482,896	13,215	41.34
	1962	18	656,000	517,325	13,127	38.96
	1963	17	666,000	554,209	13,041	36.74
	1964	16	600,000	593,723	12,956	34.69
	1965	15	635,000	636,053	13,749	34.44
	1966	14	109,000	681,402	13,667	32.57
	1967	13	650,000	729,985	13,586	30.84
	1968	12	719,000	782,031	13,508	29.23
	1969	11	650,000	837,788	13,431	27.74
	1970	10	894,000	897,520	17,112	31.38
	1971	9	893,000	961,511	17,039	29.77
	1972	8	733,000	1,030,064	16,968	28.27
	1973	7	1,100,000	1,103,505	16,900	26.88
	1974	6	1,400,000	1,182,182	29,202	38.15
	1975	5	1,342,000	1,266,469	29,138	36.11
	1976	4	1,382,000	1,356,765	29,077	34.22
	1977	3	1,480,000	1,453,499	29,018	33.33
	1978	2	1,475,000	1,557,130	28,962	30.82
	1979	1	1,617,000	1,668,149	28,909	29.30
	1980	0	1,933,000	1,787,084	28,859	27.88

a.  $P_a = 1,787,084 [\exp(0.0688a)]$   
 $\ln P_a = \ln P_0 + 0.06887a$   
 (141.10) (8.28)  $R^2 = 0.7742$

where depreciation ( $\delta$ ) = 6.9 percent.

b. Capital return ( $RK$ ) = income - costs.

c. Assuming that  $RK_a = (\delta + i_a - \pi)P_a$ , then  $i_a = (RK_a/P_a) - \delta + \pi$ , where  $i_a$  is profitability,  $\delta = 6.9$  percent (see equation 12-8), and  $\pi = 15.41$  percent.

Source: 1980 prices: Jaramillo (1981).

2.6 percent annually; for busetas in urban areas the rate was 30 percent. Between 1971 and 1980 the average growth rates were 6.2 percent for the bus stock in Bogotá and 25 percent for the buseta stock; in Cali the growth rates were 2.4 percent for buses and 21.93 percent for busetas.

Walters (1979) has indicated that in developing countries minibuses are more efficient than buses with respect to costs, waiting time, and trip frequency. Pachón (1981c) concludes that the demand for busetas increases with income. Furthermore the income elasticities of demand for busetas based on time series are higher than those obtained from cross-sections (see Pachón 1981b). During the period considered, the number of microbuses, which are smaller than busetas, decreased.

#### Profitability of Buses and Busetas

For 1967-and-later models, the busetas' profitability is higher than that of buses of the same age.<sup>4</sup> This implies that policies regarding public transport fares and subsidies favor intermediate-size vehicles. This, together

with the preference for these vehicles on the demand side, explains their rapid growth in Colombia.

The difference in the profitability of each type of vehicle not only is consistent with the rapid growth of the buseta stock but also reflects restrictions on the addition of busetas to the stock. If there were no restrictions, the profitability of both types of vehicles could be expected to be equalized. As mentioned above, a license, or *nota opción*, is required before a vehicle can be affiliated to a transport company. To the extent that obtaining this license is more difficult for busetas than for buses, a higher profitability for busetas is ensured.

Data on the prices of used vehicles confirm the existence of greater restrictions on the entry of busetas. Automotive vehicles show rapid depreciation in the first year and a constant rate of depreciation for the remaining useful years (see Wykoff 1970, pp. 171-72). With such a pattern the adjusted value based on the regression would be less than the value charged by the distributor, since the owner of the vehicle has to pay a penalty to take the vehicle from the distributor. When, on the contrary, there are supply restrictions it is to be expected

Table 12-3. *Busetas: Capital Return and Profitability*  
(pesos)

	<i>Model</i>	<i>Age</i>	<i>Price in 1980</i>	<i>Estimated price<sup>a</sup></i>	<i>Capital return<sup>b</sup></i>	<i>Profitability<sup>c</sup></i>
	1959	21	n.a.	307,202	3,588	23.65
	1960	20	n.a.	329,289	4,326	25.40
	1961	19	n.a.	352,965	5,098	26.97
	1962	18	n.a.	378,343	5,906	28.37
	1963	17	n.a.	405,545	6,752	29.61
	1964	16	n.a.	434,703	7,636	30.72
	1965	15	566,000	465,958	8,560	31.68
	1966	14	624,000	499,459	9,526	32.52
	1967	13	475,000	535,370	10,535	33.25
	1968	12	550,000	573,862	11,580	33.87
	1969	11	495,000	615,122	12,688	34.39
	1970	10	678,000	659,349	13,836	34.82
	1971	9	605,000	706,755	15,034	35.16
	1972	8	789,000	757,570	16,283	35.43
	1973	7	650,000	812,038	17,585	35.62
	1974	6	900,000	870,423	18,944	35.75
	1975	5	965,000	933,005	20,359	35.82
	1976	4	994,000	1,000,087	21,834	35.83
	1977	3	n.a.	1,071,992	23,370	35.80
	1978	2	1,375,000	1,149,067	24,971	35.71
	1979	1	1,321,000	1,231,684	26,637	35.59
	1980	0	1,270,000	1,320,240	28,372	35.42

n.a. Not available.

a.  $P_a = 1,320,240 [\exp(-0.0694a)]$

$\ln P_a = 14.0933 - 0.06943a$   
(141.10) (8.63)  $R^2 = 0.8515$

where depreciation ( $\delta$ ) = 6.9 percent.

b. Capital return ( $RK$ ) = income - costs.

c. Assuming that  $RK_a = (\delta + i_a - \pi)P_a$ , then  $i_a = (RK_a/P_a) - \delta + \pi$ , where  $i_a$  is profitability,  $\delta$  = 6.9 percent (see equation 12-8), and  $\pi$  = 16.54 percent.

Source: 1980 prices: Jaramillo (1981).

that the owner will obtain a bonus for taking the vehicle from the distributor. In the first case the adjusted price (using the regression) will be less than the list price; in the second case the adjusted price will be greater than the list price. If the relation between the estimated price and the list price is defined as a premium factor, the degree of restriction in the market can be inferred because the bigger is the premium factor, the higher is the degree of restriction in the market.

For Bogotá in 1980 the bus premium factor,  $B_b$ , is 1.02; the premium factor for the buseta,  $B_{bt}$ , is 1.28. Both vehicles receive a bonus for leaving the distributor. This factor, for both buses and busetas, tells us that the government, through INTRA, still imposes severe restrictions, in relation to the demand for public transport, on the entry of these vehicles to the stock. In the case of busetas the restriction is even clearer, which is in accordance with their high profitability, and the supply of this mode should be even higher to obtain equilibrium between supply and demand.

If the rapid growth of the buseta stock is seen in perspective, it can be inferred that in the beginning there were more restrictions on the entry of busetas and that the gap between the profit rates for buses and busetas was larger. Kozel (1981) has shown that supply restrictions decreased during 1972-78. Essentially, busetas have been allowed to operate on the same routes as buses, and a greater adjustment of the buseta supply has been facilitated.

### The Load Factor

From the point of view of economic efficiency, where there are different types of transport vehicles the equipment should be assigned in order of efficiency, with the most efficient assigned first. The most efficient vehicles would be expected to work all day and the least efficient only during peak hours. Where there are two types of equipment with different fares, assignment would generally be in order of increasing break-even load fac-

tors: the equipment with the lowest break-even load factor would be assigned first and that with the highest break-even load factor last. Such an operating rule causes the vehicle with the lowest break-even load factor also to have the lowest observed load factor.

In Colombia, since the bus owner receives a monthly subsidy and the buseta owner receives no subsidy, the buseta's break-even load factor should be lower.<sup>5</sup> In consequence the buseta would tend to be used during both peak and nonpeak hours and to have a lower load factor.

With the use of assumptions for calculating the bus and buseta load factors in Bogotá, and cost data for some periods when fares were changed, the effect of these changes on the break-even load factors can be seen. For example, the bus break-even load factor increased from April 1978 to May 1980, which reflected the slower increase in fares than in variable costs. After the most recent fare increase studied, in October 1980, the break-even load factor decreased, since fares had increased more than costs. It can be inferred that the use of buses had been decreasing between April 1978 and October 1980, perhaps because the use of busetas increased.

The results show the dilemma between efficiency and equity which the government confronts in fixing fares. If the government wants to decrease the impact of the growth of operating costs by moderating the increase in fares and increasing the subsidy, the result may be stimulation of the use of vehicles that occupy more road space and that consume more fuel per passenger transported.

It can also be seen that although the load factor for the buseta increased between May and October 1980, there is a great difference between bus and buseta load factors. Whereas a bus must fill 78.5 percent of its capacity to cover its operating costs, the buseta needs to fill only 43.2 percent of its capacity. This great difference in load factors explains why a bus owner prefers to wait for peak hours to operate his bus. Only during those hours is there enough demand to guarantee that operating costs will be covered and a profit made. The buseta, in contrast, can operate on routes with low demand and during nonpeak hours because it needs to use only 43.2 percent of its capacity to cover its costs. This low factor suggests that the buseta's trip frequency will be greater than that of the bus because it does not have to wait for peak hours to operate profitably.

If we compare the observed load factor with the required load factor as in the table below, we find that the load factor for the bus is too low to cover operating costs, while that of the buseta is distinctly above the break-even load factor. This result tells us that the buses are incurring losses or that their operating costs have been overestimated.

Vehicle	Observed load factor (percent)	Break-even load factor (percent)
Bus	69	78.5
Buseta	66	43.2

The results obtained in this section show that the fare and subsidy policies encourage a more intensive use of busetas, which favors the buseta supply in both the short and the long run.

## Summary and Conclusions

In Colombia the state has tried to favor public over private transport. Given the absence of policies—other than the subsidy for buses—to remedy the externalities of private transport through the price system, it may be necessary to increase the use of quantitative restrictions. For example, some lanes may be dedicated to the exclusive use of public transport. Simultaneously with these quantitative restrictions, other policies that use the price system to discourage the use of private vehicles should be investigated. The proposals for a fuel tax and a parking tax to reduce congestion should be studied further. In addition, the possibility of establishing a license for access to the central zones, as in Singapore, should be examined.

The combined effect of the bus subsidy, the implicit fuel subsidy, and the vehicle tax has been to favor the lower-income classes. But since the private companies have much lower operating costs than does the district government company, it is probable that increases in allocational efficiency attained from a system with all government vehicles would increase operating costs. The route assignment process is flexible and has been able to adapt itself to changes in the location of activities.

Such problems of public transport as slow operating speed can be eliminated by such measures as establishing posted bus stops on all lines, reserving some roads for exclusive use by public transport, and charging, by tender, for the use of roads dedicated exclusively to public transport.

The credit established by the CRT has made investment in public service vehicles easier. The availability of credit, together with the fare and subsidy scheme, has assured an adequate profit rate for investment in different types of vehicles. The driver's remuneration system creates incentives for the efficient operation of the vehicles but can also contribute to traffic control problems. As has been proved in Cali, the authorities can make substantial improvements through campaigns directed at users.

The elaborate system for setting fares and subsidies has been described above. The adoption of a methodol-

ogy to calculate a transport cost index and the periodic collection of the required data have permitted relatively frequent adjustments that have maintained adequate profits for the investment in public transport vehicles despite inflation and changes in fuel prices. The flat fare scheme was identified as a possible cause of excessive urban dispersion. It was shown that the subsidy policies unnecessarily increase the price of old vehicles and artificially raise the prices of models at the lower end of the range to which the subsidy applies. It was further shown that fare and subsidy policies favor investment in buses and a more intensive use of these vehicles. The restrictions on the entry of buses into the stock in 1980 were greater than was the case for buses, and it was noted that this fact is compatible with a decrease in restrictions for bus entry during 1972–78.

The results of the research on the economic effects of the fare and subsidy policies lead to the conclusion that the present scheme is causing distortions in the supply of transport vehicles. Some simple reforms of the subsidy scheme could eliminate the distortions and at the same time solve other problems. A first proposal is to change the basis for the subsidy from model to age, to avoid the model effect and the price distortion of old vehicles. The same profit rate would be adopted for buses and buses. As in the present system, only buses would be subsidized.

In this analysis the values for buses and buses were calculated. In both cases a 40 percent annual profit rate in monetary terms was fixed. The bus fare covers operating costs, and the subsidy covers capital and fixed costs. Given a 40 percent profit rate, a uniform fare for a bus of any age is 3.50 pesos. The present subsidy is unnecessarily high for older buses. It is possible to fix a flat fare of 7.80 pesos for buses, a little higher than the present fare (7.50 pesos), to ensure a nominal profit rate of 40 percent, equal to that of buses.

A second reform would allow additional benefits. There is a desire in Colombia to maintain the price of fuel used by public transport below that of fuel used by private transport. The authorities have wanted to change all the public stock to diesel and to maintain low prices for that fuel, but this policy has not had much success. A possible measure is a subsidy that is related to the average fuel consumption for each type of vehicle used in public transport. If the proposal to adopt a subsidy scheme that would eliminate some of the distortions created by the present scheme is accepted, additional research on the vehicle's costs, income, and price by age will be needed. A better level of information would permit subsidies to be based on the costs of different ages of vehicles without incurring serious distortions.

Along with improvements in the definition of fares and subsidies, market forces should be permitted to

reduce the distortions from information problems or methodological limitations. For example, by eliminating restrictions on the entry of public transport vehicles, the same profitability would be ensured for all types of vehicles, and transport would be as profitable as other sectors of the economy. Similarly, some fares, such as night service charges, could be freed from controls.

Considerable knowledge exists concerning urban transport in Colombia, and the use of models to simulate routes is fairly well established. Government transport policies have been, in general, in the right direction. Policies have sought to give priority to public transport and have, by and large, favored lower-income classes.

## Notes

1. The information on costs may be overestimated because of the nature of the data and their relation to fares. The FEDESARROLLO (Urrutia 1981) study used the operating costs of its buses and obtained lower operating costs.
2. The  $t$ -values are in parentheses, there are three degrees of freedom, and the critical  $t$ -value for three degrees of freedom is 3.182.
3. The  $t$ -value to prove the hypothesis  $b = 0.1$  is 2.846. It is greater than the critical value 2.086 with a 95 percent confidence level and 20 degrees of freedom.
4. The profitability of buses can be less than that observed to the extent that there are delays in the payment of the subsidy by the CFT, the entity in charge of the payment.
5. The bus owner must cover both variable and fixed costs from fares, while the bus owner covers his costs out of the subsidy and the fares. The bus owner therefore needs a lower load factor to cover his variable costs.

## Bibliography

- Cifuentes, Jorge Ignacio. 1978. "Urban Transportation in Bogotá." Water Supply and Urban Development Department, World Bank, Washington, D.C. Processed.
- Corporación Financiera del Transporte. 1979. *Aspectos estadísticos del parque y las empresas de servicios de transporte colectivo urbano*. Bogotá.
- Departamento Nacional de Planeación—Instituto Nacional de Transporte (INTRA). 1974. *Estudio de organización y administración del transporte colectivo urbano de Bogotá* (OAT II). Bogotá.
- INTRA. 1980. *Informe final sobre el programa de índices de precios de los insumos del transporte*. Bogotá.
- Interconsult Ltda. 1970. "Estudio de organización administrativa del transporte colectivo de Bogotá" (ROT). Bogotá.
- Jaramillo, Samuel. 1981. "Situación de los conductores del transporte colectivo urbano en Bogotá." *Estudios laborales no. 1. Desarrollo y sociedad*. Bogotá.

- Kozel, Valerie. 1981. "Travel Demand Models for Developing Countries: The Case of Bogotá, Colombia." Urban and Regional Report 81-26. World Bank, Washington, D.C. Processed.
- Leibenstein, Harvey. 1966. "Allocative Efficiency vs. X Efficiency." *American Economic Review*, vol. 56 (June), pp. 392-415.
- McClure, Charles. 1974. "Incidencia de la tributación en Colombia." In Gilis Musgrave, ed., *Propuesta de una reforma fiscal para Colombia*, vol. 2, Bogotá: Banco de la República.
- Merewitz, Leonard. 1977. "On Measuring the Efficiency of Public Enterprises: Bus Operating Companies in the San Francisco Bay Area." *Transportation* (March), pp. 45-55.
- Nelson, G. 1972. "An Econometric Model of Urban Bus Transit Operations." In J. D. Wells and others, ed., *Economic Characteristics of the Urban Public Transportation Industry*. Washington, D.C.: Institute for Defense Analyses, for the U.S. Department of Transportation. Available as PB 265-325 from National Technical Information Service.
- Owen, Wilfred. 1973. *Automobiles and Cities: Strategies for Developing Countries*. World Bank Staff Working Paper 162. Washington, D.C.
- Pachón, Alvaro. 1981a. "El automóvil en dos metrópolis del tercer mundo." Documento de trabajo 17. Bogotá: Corporación Centro Regional de Población.
- . 1981b. "El impacto redistributivo de la intervención del gobierno en el transporte." Documento de trabajo 19. Bogotá: Corporación Centro Regional de Población.
- . 1981c. "Situación del transporte en Cali y Bogotá, década de los setenta." Documento de trabajo 18. Bogotá: Corporación Centro Regional de Población.
- Republic of Colombia. 1973. *Fiscal Studies*. Bogotá Urban Development Study, Phase II. Bogotá.
- Urrutia, Miguel. 1981. "Evaluación del sistema de transporte público en Bogotá." In *Buses y Busetas—Una evaluación del transporte urbano de Bogotá*. Bogotá: FEDESARROLLO.
- Walters, A. A. 1979. *Costs and Scale of Bus Services*. World Bank Staff Working Paper 325. Washington, D.C.
- Walters, A. A., and Charles Feibel. 1980. *Ownership and Efficiency of Urban Buses*. World Bank Staff Working Paper 371. Washington, D.C.
- Westin, B. Richard. 1980. "Travel Patterns in Bogotá: A Descriptive Analysis for 1972." Water Supply and Urban Development Department, World Bank, Washington, D.C. Processed.
- World Bank. 1975. "Transportes urbanos." *Documento de política sectorial*. Washington, D.C.
- Wykoff, Frank C. 1970. "Capital Depreciation in the Postwar Period: Automobiles." *Review of Economics and Statistics*, vol. 52 no. 2 (May), pp. 168-72.

## *Evaluating Pollution Control: The Case of São Paulo*

Vinod Thomas

**S**ão Paulo is the most industrialized area of Latin America, and after Mexico City and Shanghai it is the third largest metropolitan center in the world.<sup>1</sup> Today some of the most urgent problems of the urban environment are found here, particularly in the highly industrialized sections of the Greater São Paulo Metropolitan Area (GSP).

### **Pollution: Problems and Policies**

Until the early 1970s Brazilian policymakers tended to opt for environmental damage in the pursuit of economic growth and competitiveness in international markets. As pollution grew, however, the virtues of accepting such a tradeoff—even if it existed—became questionable. Under growing pressure from localities national antipollution policies, supported on the state and metropolitan levels, began to evolve in the mid-1970s. The enactment of laws is of more recent origin, and actual measures have been mostly limited to the GSP. As policies are now beginning to be implemented, however, old anxieties have re-emerged about their possible effects on competitiveness in world markets, on inflation, on efficient use of energy, and on the government budget. The shifts in opinion about environmental regulations call attention to the need for objective analysis and estimates of the effects of policies. This chapter uses a cost-benefit framework to analyze the welfare effects of industrial location and pollution control, with the focus on industrial air emissions.

*Note:* This chapter has been published in a somewhat different form as "Evaluating Pollution Control," *Journal of Development Economics*, vol. 19 (1985), pp. 133–46.

### *The Extent and Sources of Pollution*

Pollution in São Paulo has been studied by, among others, Esteves and Gianneschi (1980); Gianneschi, Junior, and Salvador (1979); Licco, Oda, and Galvão Filho (1979); Pazzagliani and Greco (1981); and the World Bank.

Air pollutant discharges in the GSP currently amount to 8,000 tons daily. Carbon monoxide constitutes 65 percent of the emissions, sulfur dioxides 13 percent, hydrocarbons 10 percent, particulates 7 percent, and nitrogen oxides 5 percent. In 1978 daily air quality standards were exceeded 299 times for carbon monoxide, 121 times for particulate matter, and 17 times for sulfur dioxide. The highest concentrations of these substances in the same year were, respectively, 282 percent, 190 percent, and 219 percent more than their standards. During 1976–78 the average annual concentration of carbon monoxide in the more polluted areas was 12–13.5 parts per million (ppm), while that of particulates was 115–126 micrograms per cubic meter. These levels are comparable with those recorded in some of the largest cities in the United States during the early 1970s.<sup>2</sup>

The damage is compounded by São Paulo's climate and topography. Light winds, air stagnation, and frequent temperature inversions hold polluted air close to the ground. In the winter months of June, July, and August the concentration at ground level rises as much as 50–75 percent over the annual mean in some places, which aggravates respiratory illnesses.

Spatial variations in damages are sharp within the thirty-seven municipalities in the GSP. Nine municipalities contribute about 95 percent of total air emissions.

An analysis of the lead content in the blood of humans compared three neighboring areas: Embu-Guaçu, which is mildly polluted; São Paulo municipality, which is more polluted; and São Bernardo, which is heavily polluted (Fernicola and Azevedo 1979). The average lead concentrations in blood among representative groups of people are shown below.

	<i>Average lead in blood (micrograms per cubic meter)</i>	<i>Standard deviation</i>
Embu-Guaçu	11.2	5.6
São Paulo municipality	12.4	4.8
São Bernardo	20.5	5.7

In the GSP, about 94 percent of carbon monoxide, 73 percent of nitrogen oxides, and 72 percent of hydrocarbons are discharged by some 2 million vehicles. Industrial processes and stationary fuel combustion by approximately 30,000 establishments are responsible for about 80 percent of particulate matter, over 90 percent of sulfur dioxide emissions, 19 percent of hydrocarbons, and 25 percent of nitrogen oxides. Automotive traffic is clearly the largest source of total pollution, while industrial processes and stationary fuel combustion are the principal sources of particulates and sulfur dioxide. Roughly 65 percent of the industrial emissions are released by a relatively few industries in the nonmetallic, chemical, and metallurgic categories. According to Gianneschi, Junior, and Salvador (1979), about 90 percent of particulates and 74 percent of sulfur dioxide are discharged, respectively, by only 5 percent and 10 percent of the sources; at the other extreme, 3 percent of the particulates and 10 percent of the sulfur dioxide come from 75 percent and 73 percent, respectively, of the least polluting industries.

### *The Policy Focus*

Federal legislation has established uniform national air quality standards and broad guidelines for their achievement (Companhia Estadual de Tecnologia de Saneamento Ambiental [CETESB] 1979a, b; Empresa Metropolitana de Planejamento de São Paulo [EMPLASA] 1979; Pazzaglini and Greco 1981). By and large, uniform abatement by polluting industries has also been sought but thus far has been enforced unevenly in different areas. Enforcement of abatement has been effected through the CETESB, the state's pollution control agency, with a system of fines and penalties which has been more effective in the GSP than elsewhere. Thus, at the moment there is considerable spatial differentiation in pollution abatement.

One approach adopted to contain pollution has been to indirectly influence environmental quality by ma-

nipulating the growth of industrial output. This policy has resulted partly from the current strategy of trying to prevent a worsening of the problem rather than reduce existing emissions. Toward this goal, the locational decisions of new producers have been influenced through licensing and zoning procedures. The CETESB now licenses all new industrial investment in the state of São Paulo.

In the future, options other than those that directly affect output may be expected to increase. Already there is a \$187 million project for air and water pollution abatement which regulates a small number of industrial plants that account for a very high percentage of total pollution. The air pollution control is aimed primarily at particulates, and the water pollution control focuses on toxic wastes. Loans are offered to enable plants to meet state and federal quality standards. It is chiefly plants in the GSP that qualify for the loans. In operation the project is expected to stress whatever turns out to be the least-cost control strategy for reducing smoke and meeting air quality standards. Since air quality is measured in the area near a factory (rather than in the smokestack), greater efforts and higher control costs are to be expected in the more polluted areas.

### **The Costs of Pollution Control**

Estimates of the capital and operating costs of pollution control equipment and spare parts for 285 industries that contribute 97 percent of industrial particulate pollution in the GSP have been compiled. (These estimates are CETESB projections that used information for industries without control equipment.) For the fifty-three industries that account for 90 percent of particulate effluents, a 94 percent abatement—that is, 85 percent of industrial particulates and 55 percent of all particulates—might be achieved in 1977 at a total cost of about 418 million cruzeiros (about \$30 million) or about 3,200 cruzeiros (\$225) a ton of emissions.<sup>3</sup> If a ten-year life of the equipment and a 10 percent interest rate are assumed, the annualized investment cost is \$4.75 million. Addition of a (liberal) 25 percent for annual labor and maintenance costs yields an overall annual cost of about \$6 million. If 25–50 percent of the GSP's 11 million people are assumed to be directly affected by pollution,<sup>4</sup> a per capita annual cost of about \$1.1–\$2.2 is implied. Elimination of an additional 7 percent of smoke from 232 more industries (or, the collection of about 91 percent of industrial particulates and about 70 percent of all particulate effluents) more than doubles the total cost, to about 964 million cruzeiros (\$68 million in 1977), and the cost a ton to 6,800 cruzeiros (\$480).

The Medio Paraiba area, which has a population of

over 1 million, is a small but heavily industrialized and heavily polluted area in the state of São Paulo.<sup>5</sup> Approximate total particulate emissions from point sources are over 42,466 tons a year. According to Kowalczyk (1980) the cost of the best control technology, which achieves about 98 percent abatement, is \$16.6 million (in 1979 dollars), or about \$550 a ton if the technology is applied to the four sources which account for 72 percent (30,759 tons) of all industrial pollution in the area. If 25 percent annual labor costs are assumed, an annualized investment plus labor cost figure is \$3.25 million to abate 30,759 tons annually. If only 75 percent of the people are directly injured by the airborne discharges, this means an annual cost of \$4.33 (1979 dollars) for each person affected.

Licco, Oda, and Galvão Filho (1979) estimate that in many Brazilian industries effluent per unit of output (the emission factor  $e$ ) is constant. In other industries it

declines somewhat with the volume of production,  $X$ , as shown in the following examples.

Industry	$e$ , in kilograms per ton of output $X$
Steel: basic oxygen process	$6X$
Iron	$9X$
Cement: calcination furnace	$6X$
Ceramics	$9.5X^{0.8}$
Ammonium phosphate	$10X^{0.8}$

If the emission factor declines significantly with output, it would be advantageous to require less than proportionate abatement from larger producers and thereby allow those who pollute less per unit of output to produce more.

Economies of scale in pollution control are implied in Kowalczyk (1980). As table 13-1 shows, the control costs per ton of emissions for the largest polluters, who contribute 90 percent of the particulates in the GSP, are

**Table 13-1. Control Costs for Industries That Contribute 97 Percent of Particulate Emissions in the GSP**  
(thousands of cruzeiros, except where otherwise specified)

Industry	Number of industries <sup>a</sup>	Emissions (tons per year)	Control efficiency (percent)	Control costs (1977 estimates)						Total cost (1980-81 estimate)
				Capital costs		Installation costs		Total cost		
				Total	Per ton	Total	Per ton	Total	Per ton	
Nonmetallic	110	107,015	n.a.	n.a.	n.a.	n.a.	n.a.	443,870	4.15	1,194,600
Average		973	n.a.	n.a.	n.a.	n.a.	n.a.	4,035	4.15	10,860
Top 90 percent	36	103,770	95	110,810	1.07	101,620	0.98	212,430	2.05	660,340
Average		2,882	95	3,078	1.07	2,822	0.98	5,900	2.05	18,340
High <sup>b</sup>		9,600	n.a.	10,000	1.04	10,000	1.04	20,000	2.08	54,000
Low <sup>b</sup>		400	n.a.	2,600	6.50	1,600	4.00	4,200	10.50	11,300
Chemicals	20	16,030	n.a.	n.a.	n.a.	n.a.	n.a.	122,190	7.62	328,900
Average		802	n.a.	n.a.	n.a.	n.a.	n.a.	6,110	7.62	16,455
Top 90 percent	6	15,360	80	n.a.	2.81	31,900	2.08	75,100	4.90	202,000
Average		2,560	80	7,200	2.81	5,317	2.08	12,517	4.90	33,670
High <sup>b</sup>		3,300	n.a.	6,550	1.98	5,200	1.59	11,800	3.57	31,760
Low <sup>b</sup>		980	n.a.	6,550	6.68	3,950	4.03	10,500	10.71	28,300
Metallurgic	115	17,110	n.a.	n.a.	n.a.	n.a.	n.a.	397,270	23.22	1,069,170
Average		110	n.a.	n.a.	n.a.	n.a.	n.a.	2,563	23.22	6,898
Top 90 percent	11	12,560	98	68,000	5.41	62,000	4.94	130,000	10.35	350,110
Average		1,142	98	6,182	5.41	5,636	4.94	11,818	10.35	31,800
High <sup>b</sup>		2,200	n.a.	10,700	4.86	10,700	4.86	21,400	9.72	57,590
Low <sup>b</sup>		430	n.a.	3,000	6.98	2,000	4.65	5,000	11.63	13,460
All industries	285	140,155	n.a.	n.a.	n.a.	n.a.	n.a.	964,330	6.88	2,592,670
Average		492	n.a.	n.a.	n.a.	n.a.	n.a.	3,384	6.88	9,097
Top 90 percent	53	101,690	98	220,401	1.68	195,520	1.48	417,530	3.16	1,212,450
Average		3,485	94	4,190	1.68	3,690	1.48	7,878	3.16	22,800

n.a. Not available.

a. Particulates involve a total of about 1,150 enterprises.

b. Excluding extremes.

Source: CETESB estimates; Kowalczyk (1980).

significantly lower than for the other polluters. Thus, it may be advantageous to focus control policies on the largest polluters. If an 85 percent overall abatement were achieved, the effort to control the remaining 10 percent—the smaller polluters—would raise overall costs, compared with a strategy of relying only on the top 90 percent of the polluters (assuming a control efficiency of 94 percent).<sup>6</sup>

The efficiency of control equipment and the abatement cost for each unit also vary among industries (see table 13-1). Nonmetallic industries, which contribute the bulk of particulate emissions (76 percent excluding fuel combustion), face significantly lower control costs than do chemical and metallurgic industries. (Industries excluded from the table experience still higher control costs.)

The range in cost differences is apparent when a low control cost industry, cement, is compared with iron, a high control cost industry (see table 13-2). As a greater variety of equipment becomes available in the market and experience accumulates, however, these cost differences are likely to diminish. The existing interindustry cost variation nevertheless implies potential savings from differentiation between producers on cost grounds.

### The Benefits of Pollution Control: Health Effects

The health effects of pollution are discussed in Lave and Seskin (1977) and in Smith (1976).

#### *Air Pollution and Human Health*

Fernicola and Lima (1979) have evaluated the degree of exposure of the population of São Paulo to carbon monoxide. To determine the carboxihemoglobin (COHb) content by the spectrophotometric method, 327 blood samples were collected from adults. Thirty samples for each control group were taken from residents of Embu-Gauçu, an area considered to have a low carbon monoxide concentration. The following significant results were obtained.

Group	Concentration of COHb in blood (percent)
Traffic policemen	
Smokers	6.3 ± 2.07
Nonsmokers	2.1 ± 0.68
Bus drivers	
Smokers	4.6 ± 1.94
Nonsmokers	1.6 ± 0.48
Control	
Smokers	3.8 ± 1.74
Nonsmokers	0.8 ± 0.21

Thus, if the national air quality standard of 9 ppm for carbon monoxide were met, COHb would be kept to about 2 percent for nonsmokers. The relation between increasing cardiovascular disease and COHb content of over 2 percent has been cited in Fernicola and Lima (1979).

Mendes (1976) has examined the effect of air pollution on mortality in an epidemiological study that correlated deaths for each day in the csp in 1973 with meteorological conditions and air pollution levels. Five peaks in deaths were observed; at least one, on August 1, was clearly related to a dramatic worsening of air pollution. Particulate and sulfur dioxide concentrations rose after July 25 and reached very high levels by July 30. On August 1 sulfur dioxide reached a peak in Capuava of 452 micrograms per cubic meter, in Aclimação of 371, in Tatauape of 292, and in Cerqueira Cesar of 288. On the same day deaths reached a high for the year of 299, compared with an annual average of 228. Deaths of people over 65 or less than 1 year old and deaths attributed to respiratory diseases also peaked on the same day. In all, a close correlation between concentration of sulfur dioxide, total deaths, and deaths resulting from respiratory diseases (particularly for people over 65 or less than 1 year old) was observed statistically between July 25 and August 8. If the same study were carried out for 1974 and 1975, when air pollution was worse, more conclusive evidence on the relation of pollution to the deterioration of human health might have been found.

#### *Air Pollution and Mortality*

A regression technique, after Lave and Seskin (1977), has been used to try to isolate a pollution-mortality relation (see Thomas 1980 for details). It should be emphasized at the outset that the purpose is to explore the existence of any significant association between the two phenomena in São Paulo rather than to derive the exact effect of pollution on mortality. As will be clear, the absence of an adequate data base limits the explanatory power of the equations. More generally, one should not expect to adequately capture relatively short-term pollution-mortality relations, which must be based on long-term exposure to the effluents. Furthermore, these regressions measure the marginal effects of pollution on mortality with other variables held constant, but in reality a simultaneous explanation of the changes in other variables as well as in pollution and mortality might be required.

Annual cross-sectional data for 1977 for the thirty-seven municipalities in the csp were analyzed with the use of linear multivariate regression analysis. The mortality rate (deaths per 10,000 population) is the dependent variable. Death is recorded at the place of domicile

**Table 13-2. Pollution Control Costs for Some High-Pollution Industries:  
São Paulo and Rio de Janeiro, 1980**  
(1979 U.S. dollars)

Industry	Average collection efficiency	Metropolitan São Paulo			Metropolitan Rio de Janeiro			Cost- benefit ratio
		Particulate emissions		Capital cost per ton	Particulate emissions		Capital cost per ton	
		Tons per year	Percent		Tons per year	Percent		
Nonmetallic								
Cement	98	38,917	28.39	67	25,500	20.69	69	1.0
Quarrying	95	31,467	22.95	222	64,300	52.17	124	13.6
Ceramic								
Clay	90	19,900	14.52	292	9,930	8.06	94	n.a.
Gypsum lime	90	2,419	1.76	952	...	...	...	n.a.
Asphalt	98	10,745	7.84	358	21,588	17.52	169	17.9
Concrete	66	3,370	2.46	490	316	0.26	563	62.1
Glass	80*	785	0.57	8,577	...	...	n.a.	n.a.
Chemicals	80*	12,252	8.94	625	...	...	n.a.	n.a.
Metallurgic								
Steel	98	6,863	5.01	1,058	1,268	1.03	1,879	13.6
Iron foundries	98	10,364	7.56	1,860	343	0.28	6,467	45.6
Total or average	94	137,082	100.00	468	123,245	100.00	151	n.a.

... Zero or negligible.

n.a. Not available.

\* Approximation.

Note: Ratio of particulate emission, São Paulo/Rio, 1.11. Ratio of control cost, São Paulo/Rio, 3.10.

Source: Based on Kowalczyk (1980).

rather than of death, although it is not known how long the person had lived there. Data on total mortality are used first. Then, on the assumption that the pollution effect is longer term, children less than 1 year old and less than 10 years old are excluded alternatively.

The independent variables that represent environmental, physical, and socioeconomic characteristics are pollution level, population density, average income level, hospital beds per person, and the percentage of the population 65 years and older.<sup>7</sup> Since air quality data are not available at the municipal level, annual emission data were used instead. The emission data for particulates, sulfur dioxide, and carbon monoxide are available for the six municipalities which account for about 80 percent of industrial pollution. For the other municipalities levels of particulates and sulfur dioxide had to be approximated with the use of estimated emission factors (emissions for a quantity of industrial output), and the level of carbon monoxide had to be estimated from an emission factor based on the number of vehicles. All the emission variables are expressed as a fraction of the area in square kilometers, to better reflect their concentration. Since estimates of particulates are far more reliable than estimates of sulfur dioxide and carbon monoxide, the latter two pollutants are excluded in one set of calculations.

Population density is an important determinant of the incremental damages from a given amount of pollution. The age of the population clearly increases the mortality rate. Average personal income is expected to affect mortality inversely; estimated per capita value added by municipalities, which is admittedly a crude proxy for the income of people living in those locations, had to be relied on. For the municipalities the availability of medical service was approximated with data on hospital beds per person, as a socioeconomic variable.

Equation 13-1 regresses the total mortality rate (deaths per 10,000 population) for the thirty-seven municipalities on total pollution (tons a square kilometer) and the socioeconomic variables for 1977.

$$\begin{aligned}
 (13-1) \quad TM = & 76.214 + 0.120 PM - 0.068 SO_2 \\
 & \quad (1.51) \quad (-1.03) \\
 & - 0.020 CO + 4.357 P \geq 65 \\
 & (-0.67) \quad (1.25) \\
 & - 67.893 VA - 0.248 HB \\
 & (-0.43) \quad (-1.46) \\
 & - 0.002 P/km^2 + e \\
 & (-1.43)
 \end{aligned}$$

$$R^2 = 0.341$$

where  $TM$  is the total mortality rate for the municipality,  $PM$  is mean particulate matter,  $SO_2$  is mean sulfur dioxide,  $CO$  is mean carbon monoxide,  $P_{\geq 65}$  is the percentage of the population aged 65 and older,  $VA$  is per capita value added,  $HB$  is hospital beds per 1,000 people, and  $P/km^2$  is population density. The units of the variables are given in table 13-3. The numbers in parentheses are  $t$ -statistics.

Only 34 percent of the variation in the total mortality rate is explained by the seven independent variables ( $R^2 = 0.341$ ).<sup>8</sup> The exclusion of important socioeconomic variables in the equation may partly account for its low explanatory power. The noteworthy result is the positive and significant coefficient for particulates. Because particulate data are the most reliable, this result is encouraging. The other pollution variables, mean sulfur dioxide and mean carbon dioxide, have no significant effect in the regressions, but the poor quality of the data may account for this result. As expected, the percentage

of the population over 65 is consistently and significantly related positively to the death rate. The variable for per capita hospital beds also has a significant and negative effect on the mortality rate. Population density has a negative effect, which is contrary to expectation.<sup>9</sup> That income is not a significant factor in any of the equations is probably explained by the weakness of the data.

According to equation 13-1, an annual increase of 1 ton of particulates per square kilometer in the GSP over 1977 particulate levels is associated with an increase in the mortality rate of 12 deaths per million people. On the basis of a mean particulate concentration of about 17 tons per square kilometer<sup>10</sup> and a mean mortality rate of 8,830 per million, a 50 percent reduction in industrial particulates alone<sup>11</sup> would be associated with a 1.2 percent reduction in the mortality rate. Inclusion of better estimates of mean sulfur dioxide and carbon monoxide levels may be expected to raise this effect significantly.<sup>12</sup>

Table 13-3. Total Mortality per 10,000 Population, Thirty-seven Municipalities, GSP, 1977

Value	Particulates			Particulates, $SO_2$ , $CO$		
	$TMT$	$TM_1$	$TM_9$	$TMT$	$TM_1$	$TM_9$
$R^2$	0.175	0.343	0.366	0.341	0.451	0.470
Constant	69.079	22.290	25.618	76.214	27.913	30.759
Air pollution variables (tons a square kilometer a year)						
$PM$	0.006* (1.423)	0.002 (0.629)	0.002 (0.584)	0.120* (1.509)	0.075* (1.630)	0.075* (1.460)
$SO_2$	—	—	—	-0.068 (-1.031)	-0.032 (-0.843)	-0.030 (-0.694)
$CO$	—	—	—	-0.020 (-0.667)	-0.127 (0.715)	-0.017 (-0.848)
Socioeconomic variables						
$P_{\geq 65}$ (percent of total)	5.331* (1.606)	7.449** (3.984)	8.723** (4.140)	4.357 (1.245)	6.633** (3.290)	8.015** (3.525)
$VA$ (cruzeiros per capita)	-189.581 (-1.251)	14.330 (0.168)	52.335 (0.543)	-67.893 (-0.430)	61.350 (0.674)	112.000 (1.091)
$HB$ (per 1,000 population)	-0.251* (-1.393)	-0.166* (-1.640)	-0.228* (-1.996)	-0.248* (-1.458)	-0.162* (-1.547)	-0.225** (-2.037)
$P/km^2$	—	—	—	-0.002* (-1.431)	-0.119* (-1.599)	-0.001* (-1.432)

— Not applicable.

Notes: The numbers in parentheses are  $t$ -statistics.

\* Significant at 10 percent level.

\*\* Significant at 5 percent level.

Variables:

$TMT$  = total mortality rate

$TM_1$  = mortality rate excluding children 1 year old and younger

$TM_9$  = mortality rate excluding children 9 years old and younger

$PM$  = particulate matter

$SO_2$  = sulfur dioxide

$CO$  = carbon monoxide

$P_{\geq 65}$  = population 65 years and older

$VA$  = value added

$HB$  = hospital beds

$P/km^2$  = population per square kilometer

With child mortality excluded, an increase of 1 ton of particulates per square kilometer may be associated with an increase of about 8 deaths per million. Given a mean mortality rate of about 5,690 per million for the population over a year old, the implication is that a 50 percent reduction in particulates alone is associated with an approximate 1.2 percent decrease in the mortality rate of noninfants.

As a further check on the above results, data for seven highly polluted subdistricts within the São Paulo municipality over the six-year period 1973–78 were analyzed.<sup>13</sup> Data on total mortality for all age groups were available from Secretaria de Economia e Planejamento (SEPLAN). Mean monthly pollution readings by the CETESB, in micrograms per cubic meter for particulates, sulfur dioxide, and carbon monoxide, were averaged for 1973–78.<sup>14</sup> The weakest link in the data is income; per capita income was estimated rather roughly by EEMPLASA for various zones which were aggregated to correspond to the subdistricts.

Equation 13-2 regresses the total mortality rate (deaths per 10,000 population) for the seven subdistricts during 1973–78 on mean and maximum particulate and sulfur dioxide concentrations in micrograms per cubic meter. It is assumed that family income grew by 3 percent in real terms between 1973 and 1978. (Other results are in table 13-4.)

$$\begin{aligned}
 (13-2) \quad TM' &= -134.99 + 0.888 PM \text{ mean} \\
 &\quad (5.72) \\
 &+ 0.153 SO_2 + 10.655 P \geq 60 \\
 &\quad (1.27) \quad (4.28) \\
 &+ 0.009 I \\
 &\quad (2.83)5 \\
 &- 0.003 P/km^2 + e \\
 &\quad (4.76) \quad R^2 = 0.579
 \end{aligned}$$

where  $TM'$  is the subdistrict mortality rate,  $PM \text{ mean}$  and  $SO_2 \text{ mean}$  are the average particulate and sulfur dioxide concentrations, respectively,  $P \geq 60$  is the percentage of the population aged 60 or older,  $I$  is per capita income, and  $P/km^2$  is population density. In table 13-4  $PM \text{ max}$  and  $SO_2 \text{ max}$  are annual averages of monthly maximum values of particulate and sulfur dioxide concentrations.

About 60 percent of the variation in mortality among the subdistricts is explained in equation 13-2. Inclusion of  $PM \text{ max}$  and  $SO_2 \text{ max}$  raises the explanatory power to 66 percent.<sup>15</sup> A change in the assumption of constant income to a 3 percent growth in income has little effect. Similarly, the deletion of the variables  $PM \text{ max}$ ,  $SO_2 \text{ max}$ , and  $P/km^2$  does not affect principal results appreciably.

Equation 13-2 implies a high and significant associa-

tion—an improvement by 1 microgram per cubic meter is associated with a decrease in mortality of 8 deaths per 100,000—which may be explained by several factors. The seven subdistricts face about the worst pollution problem, and marginal benefits from control at existing pollution levels are high. The pollution variable may also be a proxy for excluded socioeconomic variables. The relatively low level of public services in some of the more polluted areas may contribute to higher mortality rates. Thus, the strong marginal mortality effect found should not be fully attributed to pollution, but the consistently significant coefficient of particulate concentration may be additional evidence of a significant incremental pollution-mortality effect in areas that are already heavily polluted.

### Policy Implications

It was indicated that if 25–50 percent of the GSP's population is assumed to be directly affected by pollution, the per capita costs of a 94 percent abatement by the industries (or of a 55 percent reduction in total particulate emissions in the area) would be \$1.1–\$2.2. A per capita cost figure of about \$4 is implied by Kowalczyk (1980) for the abatement of 71 percent of industrial particulates in the heavily industrialized Medio Paraíba area, on the assumption that only 75 percent of the people are directly hurt. It was also suggested (with caveats) that air pollution is statistically significant in explaining variations in the total death rate in the GSP.

Although no dollar value was placed on the health benefits associated with pollution control, the orders of magnitude of costs of control suggest that net benefits could possibly be obtained from a substantial reduction—say 85–90 percent—in the stock of pollution. In the GSP smoke abatement can be efficiently achieved by focusing on a relatively few large polluters, and this is the current policy. Concentration on the large nonmetallic industrial polluters may also make sense on economic grounds, but such a discriminatory approach may not be practical.

Of the methods of reducing pollution, the strategy of affecting industrial output is relatively inefficient because it does not lead to substitutions in the production processes which could lower the cost of controls. The preferred approach would be to directly control pollution either by an emission tax or by pollution abatement standards, which induce producers to cut back their use of pollutant fuels or find substitutes for them and to adopt control equipment. Direct fuel restrictions provide no incentives for the use of control equipment. Similarly, requirements to use abatement equipment give no inducements to restrict or modify fuel use.

The higher intensity of industrial activity and the

Table 13-4. Total Mortality per 10,000 Population for Seven Subdistricts, São Paulo Municipalities, 1973-78

Value	Constant real income			3 percent annual growth in real income		
	A	B	C	A	B	C
$R^2$	0.317	0.587	0.667	0.314	0.579	0.663
Constant	-38.330	-136.380	-124.025	-37.548	-134.990	-122.236
Air pollution variables (micrograms per cubic meter)						
<i>PM</i> mean	0.472** (2.94)	0.894** (5.81)	1.097 (5.44)	0.469** (2.91)	0.888** (5.72)	1.119** (5.48)
$SO_2$ mean	0.040 (0.27)	0.160* (1.34)	0.062 (0.493)	0.033 (0.225)	0.153 (1.27)	0.065 (0.51)
<i>PM</i> max	—	—	-0.220** (-2.25)	—	—	-0.235 (-2.38)
$SO_2$ max	—	—	0.147* (1.33)	—	—	0.134 (1.21)
Socioeconomic variables						
$P \geq 60$ (percent of total)	3.312* (1.33)	10.55* (4.28)	9.112** (3.87)	3.499* (1.40)	10.655 (4.28)	9.209* (3.89)
Per capita income (cruzeiros)	0.003 (0.78)	0.009 (2.97)	0.010** (3.53)	0.002 (0.67)	0.009** (2.83)	0.010** (3.44)
$P/km^2$	—	-0.003 (-4.85)	-0.003** (-5.42)	—	-0.003** (-4.76)	-0.003** (-5.37)

— Not applicable.

Notes: Numbers in parentheses are *t*-statistics. Variables are defined in table 13-3. Columns A: variables *PM* max,  $SO_2$  max, and  $P/km^2$  excluded; columns B: variables *PM* max and  $SO_2$  max excluded; columns C: all variables included.

\* Significant at 10 percent level.

\*\* Significant at 5 percent level.

greater population density in São Paulo, combined with the city's unfavorable geographic and climatic characteristics, imply higher incremental damages there than in Rio. According to the 1970 Industrial Census, the value of production in São Paulo was 8.6 times that in Rio. The difference in particulate effluents indicated in Kowalczyk (1980) for 1979 is only 11 percent, which appears to be the result of a serious underestimation of fugitive dust in São Paulo. When fugitive dust sources from quarries and asphalt plants are excluded, São Paulo's effluents exceed Rio's by 145 percent. Given São Paulo's larger area, however, pollution concentration in the two places would not be of the same order of magnitude. Population density in São Paulo exceeds that in Rio (by about 17 percent in 1970 and more at present), which would lead to a significantly higher marginal benefit of pollution control in São Paulo.

Rio's industries seem to face lower control costs than those in São Paulo (table 13-2). With quarries and asphalt plants included, the difference in cost per ton of pollution control between the two places is 210 percent. Because of the inaccuracy of data on fugitive dust, however, a better comparison is provided by excluding quarries and asphalt plants; then the cost difference is 60 percent. The large difference in the control cost for

asphalt plants is primarily attributable to the inclusion of fugitive dust emissions in the Rio inventory. Ceramic plants in Rio, which are small in size, can use water spray systems instead of more expensive bag filters. Rio's higher cost per ton for foundries and steel mills appears to be a result of smaller-size facilities, for which the cost per ton of initial abatement hardware is higher.

Thus, while higher apparent pollution damages in São Paulo call for more stringent abatement than in Rio, the generally higher range of control costs in São Paulo limits the desirability of controls. In the absence of better estimates of benefits and costs, one cannot conclude whether it would be socially beneficial to treat São Paulo and Rio differently under an antipollution strategy. More information would make possible a cost-benefit ranking of the various control options for each area, with implications for differential treatment, to aid in identifying the proper priorities.

Much greater differences in damages from pollution in the São Paulo-Rio comparison may be expected if smaller towns are also considered. The desirable approach would be to set widely different abatement goals for different locations. Two other strategies may be considered: one that requires uniform percentage abatement in all areas and another that sets equal maximum

allowable emission standards for all places. In principle Brazil has uniform emission standards; in practice they are enforced only selectively and through more or less uniform percentage abatement requirements. The resulting spatial differentiation, although not necessarily ideal, would be better than that implied by the implementation of identical emission standards.

The least amount of spatial shifts in production would be induced by the uniform abatement (percentage collection) policy. If, in addition, new sources are singled out for regulation, or if they face significantly stricter restrictions, the existing locational patterns would tend to be frozen. In contrast, the desirable policy would normally require more abatement in the more damaged places and thereby encourage some displacement of existing output and induce future entrants away. An extreme but not unrealistic possibility is that an optimal policy would concern itself only with highly built-up urban centers and would set no controls on highly rural situations. As a result industrial plants would be induced to move away from population centers, which would help to cut down on the total cost of antipollution policies.

## Notes

1. On the basis of projections in Hauser and Gardner (1982) and United Nations (1980). Problems arising from conflicting definitions of boundaries make comparisons difficult. Today Tokyo and New York would also be larger than São Paulo if Yokohama and northeastern New Jersey were included in the respective metropolitan areas.

2. In New York City and Chicago, for instance, sulfur dioxide concentration averaged 150–160 micrograms per cubic meter and suspended particulates 180–190 micrograms per cubic meter in 1972. The average carbon monoxide concentration recorded in a representative area in New York City in 1975 was 4.0 parts per million.

3. The welfare cost of pollution control, as in Thomas (1980, 1982), would be somewhat lower than these measures.

4. The São Paulo municipality, parts of which are heavily polluted, contains over 65 percent of the GSP's population. The heavily industrialized areas of Santo André, São Caetano, São Bernardo, Osasco, Mogi das Cruzes, Guarulhos, and Diadema account for another 20 percent.

5. A 1980 population estimate for the industrialized municipalities that constitute this area is 1.16 million, or a population density of 224 persons per square kilometer.

6. It does not follow, however, that to meet overall abatement standards which are significantly higher than 85 percent it would be desirable to raise the standards for the larger producers above 94 percent. Instead, as the overall standards are raised, it would most likely be desirable to include the smaller polluters in the operation.

7. Data sources are Secretaria de Economia e Planejamento (SEPLAN); Fundação Sistema Estadual de Análises de Dados Estatísticos (SEADE); Secretaria de Estado de Saúde, Centro de Informações de Saúde (CIS); and Secretaria de Fazenda.

8. Elimination of infant mortality and child mortality raises the explanatory power to nearly 50 percent. The coefficient on particulates falls, however, presumably because of the vulnerability of babies to pollution.

9. This may suggest that density is endogenous within the model in adjusting for pollution levels (that is, the impact of particulate concentration on density is negative), and thus density has a negative relation to mortality.

10. About 137,000 tons of annual particulate emissions by industries is divided by the GSP's area of 7,951 square kilometers.

11. This means about a 33 percent reduction in total particulate emissions, since industries contribute 65 percent.

12. Lave and Seskin (1977) found that for 117 Standard Metropolitan Statistical Areas (SMSAs) in the United States for 1960 a 50 percent reduction in particulates and sulfates was associated with a 4.7 percent decrease in the mortality rate.

13. The subdistricts are Aclimação, Cerqueira César, Consolação, Indianapolis, Lapa, Santa Cecília, and Tatuapé.

14. Other data sources were SEPLAN, SEADE, CIS, and EMLASA.

15. The negative coefficient of *PM max* is unexpected and is contrary to findings reported earlier by Mendes (1976).

## Bibliography

- Companhia Estadual de Tecnologia de Saneamento Ambiental (CETESB). 1979a. *Legislação básica poluição ambiental estadual e federal*. São Paulo.
- . 1979b. *Legislação básica regiões metropolitanas proteção dos mananciais e zoneamento industrial*. São Paulo.
- Empresa Metropolitana de Planejamento de São Paulo (EMLASA). 1979. *Zoneamento industrial metropolitano, a experiência da Grande São Paulo*. São Paulo.
- Esteves, Silvio S., and Adenir Gianneschi. 1980. "O sistema do inventário de fontes de poluição ambiental como suporte no controle de poluição do ar na região metropolitana de São Paulo." International Clean Air Congress, Buenos Aires. Processed.
- Fernicola, Nilda G. G., and Fausto Azevedo. 1979. "Níveis de chumbo e atividades do ácido daminlevulene-codedesratasa no sangue da população de São Paulo (Brasil)." Presented at the 10th Congresso Brasileiro de Engenharia Sanitária e Ambiental, Manaus, January 21–26. Processed.
- Fernicola, Nilda G. G., and Ernesto R. Lima. 1979. "Avaliação do grau de exposição de amostras populacionais de São Paulo ao monóxido de carbono." *Rev. Saúde Pública*, no. 13. São Paulo.
- Friedlaender, Ann F., ed. 1979. *Approaches to Controlling Air Pollution*. Cambridge, Mass.: MIT Press.

- Gianneschi, A., A. P. Júnior, and N. N. B. Salvador. 1979. "Inventario das fontes de poluição no estado de São Paulo." *Trabalhos da CETESB*, vol. 1. São Paulo.
- Hauser, Philip M., and Robert E. Gardner. 1982. "Urban Future: Trends and Prospects." In Philip M. Hauser and others, eds., *Population and the Urban Future*. Albany, N.Y.: State University of New York.
- Kowalczyk, J. F. 1980. "Rio de Janeiro Industrial Pollution Control Project." Latin America and the Caribbean, Country Programs Department, World Bank, Washington, D.C. Processed.
- Lave, Lester B., and Eugene P. Seskin. 1977. *Air Pollution and Human Health*. Resources for the Future Series. Baltimore, Md.: Johns Hopkins University Press.
- Licco, E. A., H. I. Oda, and J. B. Galvão Filho. 1979. "Padroes de emissão para controle de material particulado na região da Grande São Paulo." *Trabalhos da CETESB*. São Paulo.
- Mendes, Rene. 1976. "Avaliação de efeitos da poluição do ar sobre a saúde, através do estudo da mortalidade diária na Grande São Paulo—1973." CETESB Project 36176, part 1. São Paulo.
- Mills, Edwin S. 1978. *The Economics of Environmental Quality*. New York: W. W. Norton.
- Pazzagli, Filho M., and M. A. Greco. 1981. *Zoneamento industrial em áreas críticas de poluição*. São Paulo, CNDU, Ministry of the Interior, State of São Paulo.
- Smith, V. K. 1976. *The Consequences of Air Pollution*. Cambridge, Mass.: Ballinger.
- Thomas, Vinod. 1980. "Welfare Cost of Pollution Control." *Journal of Environmental Economics and Management*, vol. 7, pp. 90–102.
- . 1981. *Pollution Control in São Paulo, Brazil: Costs, Benefits, and Effects on Industrial Location*. World Bank Staff Working Paper 501. Washington, D.C.
- . 1982. "Welfare Analysis of Pollution Control with Spatial Alternatives." In George S. Tolley and Philip Graves, eds., *Environmental Policy*. vol. 2: *Air Quality*. Cambridge, Mass.: Ballinger.
- . 1985. "Evaluating Pollution Control: The Case of São Paulo, Brazil." *Journal of Development Economics*, vol. 19, pp. 133–46.
- Tolley, George S., Philip E. Graves, and John L. Gardner, eds. 1979. *Urban Growth Policy in a Market Economy*. New York: Academic Press.
- United Nations. 1980. *Patterns of Urban and Rural Population Growth*. Population Studies 68, table 4-7. New York.

## *Improving the Effectiveness of Urban Projects*

Douglas H. Keare

Earlier chapters have reviewed the urban policy options that are available to developing countries. A variety of policies intended to alleviate urban problems and improve the functioning of cities have been used or proposed for use by developing countries. The results to date have been mixed and controversial, and there is an emerging consensus that better evaluation of the actual and possible outcomes of urban programs would be helpful to policymakers who must choose between policies. This chapter examines whether rigorous evaluation of projects can assist in improving the efficiency and effectiveness of future urban policymaking and in formulating and implementing projects. Although much of the chapter deals with specific shelter projects financed by the World Bank, the lessons drawn from such cases are likely to apply to policymaking as a whole.

In 1969 the World Bank began a lending program that was explicitly designed to respond to the compounding problems of cities in developing countries. The emphasis was on ameliorating the housing problems that faced the burgeoning low-income populations in the cities by accelerating the rate of increase in the supply of basic shelter and by facilitating access to the expanded supply by the lowest income groups. The strategy developed was to provide secure tenure and basic services in new sites and services areas and in slum and squatter areas designated for upgrading and to rely on the participants to complete or improve their houses through progressive development, with a modicum of additional assistance.<sup>1</sup> The first such project was approved in June 1972. Since then a sizable lending program has developed and has accounted for over \$2 billion in loans for sixty-two projects through 1981.

An early feature of this unprecedented endeavor was a

pilot program to rigorously evaluate selected early projects. Preparation for the evaluation began with approval of the first project in 1972, but it was mid-1975 before enough projects were ready for implementation. The housing program was launched with the assistance of the International Development Research Centre (IDRC) of Canada. Programs in Senegal (the first project), El Salvador, the Philippines, and Zambia were selected for study, and annual conferences were held to discuss the findings. The concluding conference for this phase of the work, held in November 1980 in Washington, D.C., was attended by project managers and researchers from the countries involved and from other interested countries. Subsequently the evaluators' attention has turned to the publication and dissemination of the programs' more important results and to evaluations of projects in other countries.

This chapter argues that evaluation programs can do much to guide public sector spending (or lending) programs. This was not the primary aim of this evaluation program at its outset; then the emphasis was on measuring a broad range of expected impacts on the participant populations.<sup>2</sup> As the program progressed, however, and initial experiences and evaluation findings were digested, the focus narrowed to issues that were seen to be both significant and amenable to rigorous evaluation or research. As a consequence, although the study of impacts has remained a central concern, it has been joined by the assessment of the effectiveness of programs and projects and of certain of their components.

Ex post assessment of experience makes clear that there can be a synergism of activities in a jointly programmed sequence of projects. These activities are an evaluation component, policy reviews at various critical

junctures, and associated research. As has happened in this program, project experience, evaluation findings, and research results all inform each stage of project design and each policy review. At the same time, project and policy requirements, coupled with evaluation research experience, can help to adjust the designs of evaluation and research programs to maintain responsiveness to these operational needs. The following matrix provides a helpful way of viewing this interaction.

	<i>Activity</i>	
	<i>Continuous</i>	<i>Periodic</i>
Action	Project	Policy
Study	Evaluation	Research

The general background against which effectiveness is reviewed in this chapter is one of success. The progressive development model, exemplified by the sites and services and area upgrading projects, has demonstrated its validity in a broad range of circumstances, and the recorded impacts on the participating populations and their housing conditions have been in the expected directions and of significant magnitude. Self-help construction methods have proved relatively efficient, and the impacts of projects on the housing stock have been generally greater than anticipated. The projects have been affordable and generally accessible for the target populations. Although it is still too early for some of the developmental results to have been recorded, those measurements which have been completed indicate that the projects' impacts on the socioeconomic conditions of participants have been in the directions expected. And, notably, the projects have not had negative impacts on expenditures for food and other basic necessities (see Bamberger, González-Polio, and Sae-Hau 1982; Bamberger, Sanyal, and Valverde 1982; Keare and Jiménez 1983; Keare and Parris 1982).

Notwithstanding this general record of success, the projects have encountered some problems and produced some unexpected results, some of which are detailed in the following section. An analysis of the projects' successes and shortcomings supports recommendations, advanced below, that future projects endeavor to push standards and costs still lower, include explicit provisions and opportunities for rental arrangements, and incorporate credit provisions that are more nearly tailored to the needs of the targeted families. Furthermore, it is observed that these objectives will be more easily pursued and also more effective if a rigorous and relentless approach is taken to improving cost recovery performance in the programs. The final section, "The Contributions of Evaluation," provides some examples of how operationally useful knowledge and understanding have been developed more rapidly, completely, and convincingly through the use of evaluation and associated

research than could have occurred through direct assessment of project experience alone. It concludes with some observations on how the efficiency of interaction between evaluation and research on the one hand and project design and policy formulation on the other can be improved.

### Project Efficiency and Effectiveness

Despite differences in the goals and operational procedures of the four shelter programs studied, the programs are comparable enough to warrant preliminary investigation of their efficiency and effectiveness according to a single set of criteria. Eight project aspects appear to be of particular financial, economic, and social importance in all four cases: project planning and design, selection of project beneficiaries, construction methods, materials loan programs, housing completion and occupancy, maintenance of housing and infrastructure, cost recovery, and community participation. Criteria that appear to be useful in gauging the efficiency of these components include speed of implementation, cost, quality of housing or services, accessibility by target populations, replicability, and flexibility of implementation.

There are tradeoffs among these evaluative criteria. Higher-quality housing may be more costly to construct and less affordable for low-income target groups. Greater attention to selection procedures may increase target group accessibility but raise the costs of identifying appropriate families. The sometimes conflicting objectives have to be reconciled according to a crude relative weighting of each, as determined by project management and other policymakers, and no precise model for the assessment is yet available. The discussion below briefly summarizes the constraints and options involved in applying the criteria to each project aspect and reports on selected initial results from the four projects.

### Project Planning and Design

Speed of implementation can affect the cost of both house construction and provision of services. If progressive development is too slow administrative costs per plot will rise and intended benefits will be delayed and possibly reduced, but pressure to construct houses too rapidly, without appropriate credit, can result in affordability problems for families. High, rigidly enforced design standards for quality housing are inexorably linked to higher costs, may contribute to delays, and may limit participation by low-income families and restrict the flexibility of households' responses. Reductions in capi-

tal costs, however, may produce only illusory gains because they are offset (or more than offset) by increases in operating and maintenance costs. Although costs can sometimes be reduced by the use of domestic rather than imported materials, families may prefer to build with more expensive materials and recoup the additional outlays through higher returns from renting part of the space and from eventual resale.

There is no evidence that design standards in the projects have been intolerably high. Studies of turnover do not, for example, reveal a higher than average incidence of departure from project sites among the poorer participants.<sup>3</sup> Two facts, however, seem clear: entry costs can be reduced so as to facilitate accessibility for poorer households without adverse consequences for physical results, and, even if costs remain the same, changes can be made in the mix of components and services offered that would make them more desirable. In particular, the evaluation program and other studies have shown that the projects under review have placed somewhat too high a premium on service levels and may have overlooked opportunities to facilitate increases in housing space—an attribute that has turned out to be more highly valued than was anticipated in the initial project designs. Although families in El Salvador have on the whole been satisfied with the design features of project housing, demand studies have revealed that they also view favorably the larger lot sizes (and lower service levels) of *colonias ilegales* (unauthorized settlements) if this option is available in a particular city.<sup>4</sup> Similarly, in the Philippines, where project lot sizes in the Tondo area of Manila have been constricted by the existing high densities, families have responded, following the re-blocking process, by building additional stories to increase living space.<sup>5</sup> Improved services, especially toilet facilities, have been less highly valued by upgrading families than had been anticipated by project designers.

In Zambia the plot sizes offered have been large enough to satisfy families, but apparently even larger lots would have been preferred. A partial explanation is that families have discovered that they can rent out space despite project prohibitions. The outcome, however, is biased by the fact that Zambian participants are not charged for land—which has potentially serious consequences both for project costs and replicability and for the allocation of a scarce factor of production, land. The primitive status of the building materials industry and the excess demand for good housing in Zambia help explain Zambian families' preferences for constructing houses of imported rather than domestic materials: they incur added costs for higher quality and recoup those costs by renting out space. In these circumstances efforts to encourage participants to use local materials

(for example, by producing their own soil-cement blocks), thereby reducing costs and potentially increasing long-term efficiency, have so far lacked appeal. In Senegal a similar desire for additional space rather than highly serviced facilities has been demonstrated by the choices of the low-income control groups in Guediawaye and Grand Yoff. The relatively expensive standard designs were presented to the project families in such a way that the costs, in time and money, of obtaining approval for simpler designs were perceived to be substantial. This belief may well have led to families' accepting housing standards that were higher than desired.

These findings point to the advisability of providing a wider range of building options to participant families. Households appear to demand substantially different combinations of plot size, contractor construction, and service facilities; they also wish to have the option of adding to or further upgrading their homes and of renting out space. Despite the counterarguments of some national housing planners, it appears that scaled-down housing standards are a precondition for affordable (but still desirable) housing for low-income populations. Rental arrangements, too, appear to be both desirable and feasible for target populations, although strict enforcement of cost recovery is mandatory to avoid the creation of undue subsidy of rentier groups within the projects. These issues will be treated below.

### *Selection of Project Beneficiaries*

The selection procedure also affects the efficiency of sites and services programs.<sup>6</sup> Project managers are typically concerned about both the upper and the lower income bounds for selections from the applicant pool. Care must be taken to avoid including in the project too many high-income families, some of whom may attempt to falsify income statistics to participate, but it is also necessary to ensure that low-income participants will be able to afford project costs. The speed of the selection process itself involves tradeoffs. If the screening program takes too long, households may become dissatisfied and withdraw from consideration; if it is carried out too rapidly, errors may occur.

Restricted definitions of income may also lead to biased selection standards. If selection is based only on earned income rather than on total income from all sources, projects may exclude large numbers of households which, by virtue of the combined resources within their extended families, can afford the programs. Furthermore, the criterion will be biased against female-headed households, which rely more on transfer payments from kin than do male-headed households. It must be noted, however, that verification of the true

incomes of all applicants is a costly procedure. Statistical sampling procedures will probably have to be used and the inherent errors accepted.

The evaluation studies have not discovered seriously anomalous income distributions in any of the four projects. An acceptable range of income groups appears to have been incorporated into the programs.<sup>7</sup> Some problems with underreporting of incomes arose in the Dakar project, but a more general finding has been that some families, particularly female-headed households and those employed in the informal sector, initially suffered from a measure of unintended discrimination in the selection stage, owing to problems in verifying income from nonformal sources.<sup>8</sup> Several of the El Salvador subprojects have gone so far as to include higher proportions of female-headed families than exist in the population at large. All four country projects have included families with a greater range of incomes (and more diverse sources of those incomes) than had been foreseen. Consequently, future programming efforts may well have to find a way to assess both household and extrahousehold incomes and characteristics to determine an appropriate range of income bounds for project participation, rather than necessarily modifying the selection processes themselves.

The evaluations have demonstrated that transfers from the extended family and other kin and nonkin networks account for an important and stable share of participant families' incomes. In the sample of low-income households in Santa Ana, El Salvador, transfers constituted 66 percent of the incomes of 58 percent of the families in the lowest income decile; for the next lowest decile the transfers were 25 percent of the income of 48 percent of the families. In the Tondo sample transfers accounted for 54 percent of the incomes of 43 percent of the families in the lowest third of the income distribution. Furthermore, econometric analyses indicate that transfer incomes have a contractual character and are devoted more than proportionately to expenditures which can be described as meeting basic needs. Participation in projects (which confers investment opportunities) induces additional transfers (Kaufmann 1982).

### *Construction Methods*

The projects have placed substantial premiums on identifying construction methods that reduce costs and contribute to efficient implementation. Among the cost-efficient methods are project-provided, contractor-built housing; construction through mutual help; self-help in which families hire a contractor; self-help in which families themselves do the contracting and hire

and supervise individual workers; and self-help in which families build their homes themselves. Each method involves different construction rates, costs, and implications for housing quality, replicability, and accessibility, and the tradeoffs involved are considerable. The mutual help method may be slower (although not necessarily less expensive) than certain self-help methods, which may in turn be slower (but cheaper) than contractor construction. The scheduled timing for mutual help methods may either attract or discourage families with different preferences and varying formal and informal work schedules; scheduling is a serious issue. Use of skilled builders can raise project costs in the short run, but their participation in project design and execution may also contribute to replicable plans which scale down costs, including maintenance costs, over time.

Family-based construction methods using progressive development have proved viable in sites and services and upgrading projects. It has been estimated that certain families in El Salvador have saved up to 30 percent of costs by building their own houses rather than hiring contractors, and these houses have been judged to be of a quality comparable to those constructed by skilled builders. Combinations of self-help and mutual help can also produce housing of acceptable quality at costs similar to or lower than those charged by contractors. Yet such methods can absorb great amounts of supervision and elapsed time. In El Salvador the mutual help phase of the first projects, although it produced good results, lasted more than forty weeks, since work was carried out only on weekends. The original design also underestimated the opportunity costs of labor in general, and it was found that skilled workers and other small businessmen, in particular, could have been much more profitably employed elsewhere during these work periods.<sup>9</sup> Furthermore, households headed by single adults and especially by females typically found participation both undesirable and difficult, owing to several constraints on the flexibility of their time. Sweat equity does not appear to have been available for own construction in the relatively abundant quantities assumed, and as many as 51 percent of the households in the Santa Ana and Sonsonate projects used only hired labor to bring their core housing up to habitable levels. (See table 14-1 for a comparison of housing construction methods in El Salvador.)

These findings imply that market solutions for housing construction, such as seeking income transfers, hiring labor, and balancing accounts through the household's own greater participation in the labor force and rental of part of the structure, ought to be considered in the estimation of housing costs and other aspects of project design in future shelter programs. Pure self-

Table 14-1. *Housing Construction Methods in El Salvador: Santa Ana and Sonsonate*

Item	Hired labor only	Unpaid family labor only <sup>a</sup>	Hired and unpaid family labor	All options
Percentage of participants	51	27	21	100
Average number of person-days of hired labor	46.0	—	33.2	42.4
Average weekly wage income (colones)	124.0	120.0	155.0	130.0
Average number of household members with construction experience	0.03	0.41	0.12	0.16
Average number of person-days of unpaid labor	—	44.3	45.0	44.6
Average weekly nonwage income (colones)	14.0	9.0	6.0	10.8

— Not applicable.

Note: Figures may not add to totals because of rounding.

a. Regression analysis indicates that wage rate and proxy variables (such as technical experience, number of adults ages 17 to 60, and the male-female mix within a household) best measure a household's ability to build its own dwelling. There is also some evidence that households that use self-help have lower elasticities of housing demand than others.

Source: Data from Fundación Salvadoreña de Desarrollo y Vivienda Mínima.

help, it seems, should be envisioned as only one of a set of diverse methods that can be employed for efficient housing construction and consolidation. As the El Salvador data demonstrate, the amounts of self-construction undertaken will depend, for each household, on the productivity and the opportunity costs of the family members who engage in the associated activities and hence, on the level and sources of household income and the family members' construction skills, among other factors.

### Materials Loan Programs

In the projects under consideration, loans for housing construction were limited to credits for materials. Efficiency evaluations must thus gauge whether the materials purchased have been appropriate and affordable for housing consolidation and have been distributed speedily enough to participants, and whether other forms of housing loans might have been advisable. The credits have frequently been tied to purchase of materials from project stores. Although such approaches are designed to compensate for supply failures and to capture for participants the savings inherent in bulk purchases, newly created project stores may have high administrative costs, encourage overly high building standards, and fail to recommend local materials which, though acceptable for use, do not fall under the stores' control. Although sheer mechanical replicability may appear to be facilitated through this relatively simple tied credit option, the technique can be compromised if

the materials are too costly or too difficult to acquire and distribute.

In Zambia problems in stocking on-site stores led to considerable delays in distributing materials, and families queued for as many as twenty hours a week to receive items. Thefts of materials also occurred.

Restriction of purchases of materials to project stores may have compelled families to buy higher-quality supplies than necessary or desired. It appears that in the El Pepito and San José del Pino projects in El Salvador materials could have been purchased directly at wholesale prices, thus lessening or obliterating the expected cost-reducing advantages of the stores' bulk purchases.

In the Philippines the project stores did not stock the cement, hollow blocks, or steel bars most desired by households. Tondo dwellers, on the whole, appear to have found cheaper (but acceptable) materials elsewhere, with the Housing Materials Loans Program (HMLP) providing only 25 percent of beneficiaries with construction materials.

The credit programs themselves have been deficient in certain respects. In the Philippines, Senegal, and Zambia, for example, the maximum sizes of loans available were insufficient to cover the requirements of many households that sought credit to finance construction. Total funds have sometimes been exhausted before all requests have been met.<sup>10</sup> In the face of grossly inadequate information on the demand for credit, it is understandable that project designers were loath to encourage families' overextending themselves financially and

sought to avoid this outcome by restricting both the size of individual loans and the total funds available. It appears, however, that these restrictions had undesirable effects on some participants.<sup>11</sup>

Notwithstanding these difficulties, studies to date have concluded that credit programs have been useful and that families should simply be permitted greater flexibility in seeking housing materials in the market (except, perhaps, in countries such as Zambia where materials are in acutely short supply). The empirical fact that project families hire labor for a substantial proportion of housing construction argues further for consideration of more general lending, where appropriate, that would cover the costs of labor, particularly skilled building help, as well as that of materials. Some scholars contend, moreover, that the substitutability of labor for materials in housing construction may be limited, so that tying credit to materials generally may pose constraints which are only slightly less severe than tying to materials from specific stores.<sup>12</sup>

Little evidence exists as yet in the economic, financial, or even sociological literature which would offer guidance on the optimum forms and amounts of credit. It is evident that research into both demand for credit and the responses of credit markets is a prime necessity. Studies now under way in the World Bank stress investigation of informal credit markets, particularly their linkages with household expenditures for basic needs, and this emphasis should enable the research to contribute to improvements in the design of urban shelter projects.

### *Housing Completion and Occupancy*

Occupancy of plots in sites and services projects is another factor to be assessed in determining the efficiency of project implementation. Unless sites are occupied relatively rapidly, costs can escalate through added interest incurred during construction, through the real effects of inflation, and through the expense for families of maintaining one residence while another is being built. Arguments for replicability are thus likely to be vitiated by slow inhabitation of project areas, since construction delays may cause affordability problems for poorer families and induce project managers to include more high-income households than originally desired to achieve full occupancy within a reasonable period.

To date, land acquisition and the installation of basic services have proved to be the most serious hindrances to speedy project implementation in the early (infrastructure) stages. Whether because of difficulties in alienating public land, as in Zambia, or in finding affordable and accessible terrain for lots, as in El Salvador, project managers have had to struggle literally for years

to obtain land entitlement rights. These delays, and others outside project control, have exacerbated the already difficult problems of coordination with other agencies responsible for water, lighting, and other basic services. The National Housing Authority in the Philippines is obliged to negotiate with no fewer than eleven other agencies to deliver inputs in a mutually complementary fashion.

Delays in the installation and effective functioning of services may also delay house consolidation and plot occupancy.<sup>13</sup> Families in turn are confronted with deciding when, how, and in what sequence to construct the house and move to the plot. If they can move to the plot immediately and commence construction, there will be no problem for the individual family.<sup>14</sup> Several factors may, however, delay the move and force families into a situation in which they are simultaneously paying rent on their current residence and charges on their new plot. The two principal such factors are lack of services and the absence of a habitable structure on the site. Others are distance of the new site from the old (and from the workplace) and inadequate credit.<sup>15</sup> World Bank project designers responded quickly to this situation by adding a measure of core construction to most projects. Doubtless this method is one way of effecting earlier occupancy and avoiding certain costs, but it also adds, sometimes substantially, to direct costs and is generally a less flexible option. An urgent need of the lending program is to evaluate the tradeoffs between this and more flexible approaches which may be sustained by more innovative credit mechanisms.

### *Maintenance of Housing and Infrastructure*

Maintenance of infrastructure facilities and housing must also be designed and managed efficiently. Because these project features are often the responsibility of overburdened and underfunded local government agencies, there is a potential problem of inadequate maintenance. A related issue is the nature of maintenance costs: if these costs are too high, whether because of low capital investment or poor design, the benefits of low-cost housing can be rather quickly dissipated; if design standards are kept unreasonably low and do not facilitate maintenance, project facilities will decline in quality. Though it is not entirely clear how the assignment of responsibility for maintenance might affect access to projects for certain groups (such as higher-risk households with less steady incomes), it is apparent that inadequate planning for maintenance can negatively affect government decisions regarding replicability.

Because the evaluated projects have been installed for only two years or less, definitive statements on project maintenance cannot be offered. Yet problems with gar-

bage collection have already arisen in the Lusaka project. At first view the cause appeared to be difficulties in servicing the garbage trucks and the resulting deterioration of vehicles, which led to dissatisfaction with the trash collection services and negatively affected cost repayments by families. Further investigation showed that families were not totally satisfied with the services even when they were working—they wanted the services brought closer to their houses. These doubts led to other questions about the capacities of internal roads, and hence about overall costs and charges. Project experience and evaluation have verified that these various features—design, maintenance, cost recovery, and community participation—are inextricably related and, although it is difficult, better means of dealing with them simultaneously must be sought.<sup>16</sup> Because nearly all project services will require maintenance, planners will have to take additional steps in the future to ensure that both cities and their inhabitants are financially and institutionally prepared to continue maintaining facilities over time, especially in the light of the cost recovery problems already encountered.

### *Cost Recovery*

Given the desirable scope and scale of urban shelter programs, cost recovery has since the inception of these programs been deemed a crucial feature of project design because inability to collect payments would doubtless prejudice replicability. Delays or dissatisfaction with project services can make cost recovery more difficult, and failure to recoup payments may in turn raise the costs of the project for executing agencies and for those families in the program who do pay for services. Government reluctance to make defaulters pay may be a form of disguised subsidy to participants. Such a tactic may permit lower-income families to live in housing projects they might otherwise be unable to afford, but that has not been the inevitable result. Subsidized programs typically have severely limited supply capacities, and in situations in which output is artificially restricted, higher-income families have proved adept in capturing the rationed benefits. In any case, the costs will eventually have to be met, either directly or indirectly, by others. Because the causes of default are complex—involving affordability concerns, the adequacy of administrative machinery, and the roles of political pressure groups—it is difficult to isolate the contributions of individual factors to cost recovery problems.

Serious cost recovery problems have arisen in the Lusaka program: more than 50 percent of families in some upgrading projects are in arrears, and some households do not appear to have paid anything. In this instance it has been possible to clarify causes. The initial

assumption was that affordability problems must have contributed to the result, but evaluation studies have established that repayment performance has not varied with income. The principal factor in the slow rate of collections has been the absence of political will on the part of the responsible political party to seek repayments. The party did little to prompt communities to repay development costs to the Lusaka City Council (LCC), the project's administrative agency. The LCC has now taken steps to recover costs more systematically, but efforts are hindered by the absence of an efficient recordkeeping system to maintain up-to-date information on defaulters and amounts owed and by the lack of a flexible mechanism for collecting payments. The underlying problem is the long-standing tradition of subsidized housing in Zambia, the potentially harmful effect of which was inadequately evaluated during project design and appraisal. Families, moreover, have not been well informed about repayment schedules or about the purposes for which the money collected was to be used.<sup>17</sup>

Collections were also slow to develop in the early projects in the Philippines. The initial investigation found that the poor early performance did not stem from inability or unwillingness to pay or from poor collection procedures. Rather, the cost recovery objective was insufficiently integrated into project planning, implementation, and monitoring procedures. Thus, project activities have been articulated not in a critical path framework but rather in a serial framework. Typically, when physical implementation has been completed and collections could ostensibly begin, some neglected feature has delayed collection: titling has not been completed (or even started), specific cost recovery procedures (which involve negotiations with local governments) have not been worked out, the participants have not been properly informed of their obligations, and so on. Once the cost recovery objective is cast in an appropriate management framework, it will be necessary to use available management information to monitor progress toward this objective, as with all others.

The El Salvador program offers a strong counterexample. Because of its small size and its nonprofit status, which requires it to achieve cost recovery to remain in operation, the Fundación Salvadoreña de Desarrollo y Vivienda Mínima (FSDVM) has achieved an excellent repayment record for its shelter program. Through material help and collaboration on important community matters families are brought to understand that repayments are necessary to the survival of project facilities. The FSDVM uses effectively a phalanx of incentives and penalties to achieve cost recovery. These have included support to community organizations in return for their assistance in the collection of payments, careful screening procedures to select participants able to pay project

fees, and visits by lawyers to families who have fallen more than three months behind in their payments to press them to pay. Eviction, although seldom used to date, is an option which the Fundación will not hesitate to use if necessary.

These experiences reveal the usefulness of designing projects in a local public finance context, where the ability to operate and maintain infrastructure and services can be directly related to the resulting benefits as well as to collections or cost recovery. From the outset of a project, collection mechanisms ought to be designed with community participation in mind, and with regular, relatively convenient payment schedules that are thoroughly explained. Participants must be made aware of the direct relation between installment payments and loan repayments on the one hand and goods and services received on the other, as well as what the consequences of default will be, both for themselves and for the project as a whole. Methods for improving cost recovery performance which have been developed in other Bank-financed projects have included discounts for early repayment and the withholding of materials or house consolidation loans from families that default on lot installments. Collection methods based on property values have been used and could be extended. In sum, whereas affordable projects and workable cost recovery methods can be designed without undue difficulty, the real problems are in implementation. The requirements—in addition to political will—are better-integrated planning and considerably more attention during implementation.

### *Community Participation*

Finally, community participation in project implementation, and particularly in design, can have decided effects on efficient achievement of objectives, including maintenance and cost recovery objectives. Although popular support and consultation are indispensable for eliminating resistance to project interventions, lengthy discussion of procedures or training of project *promotores* in organizational skills, for example, can delay implementation schedules substantially. In some instances project acceptability can be improved and costs reduced if community groups accept direct responsibility for the operation and maintenance of facilities, but devising new administrative structures and procedures for such purposes can be costly. Effective community participation is itself a goal of certain projects—for example those of the FSDVM in El Salvador—and hence some project managers may be willing to invest more resources in developing community institutions than are others who are more committed merely to attaining physical goals.

The evaluated projects have witnessed both the positive and negative effects of active community involvement and noninvolvement in project implementation. Because of its commitment to stimulating and using community participation in achieving social goals, the FSDVM in El Salvador has compiled a generally noteworthy record of fostering feelings of responsibility for project facilities and services and for collections. Required mutual help construction has also engendered collective action toward social aims. Yet there have been several cases in which groups have organized against the FSDVM and refused to make payments until basic services such as water were provided.<sup>18</sup> Community organizations, it seems, cannot be expected to be simply passive or neutral variables in project implementation, for they can act as stimuli or as deterrents to reaching objectives, depending on the concordance of popular and program goals.

The Lusaka project, too, has used community involvement to its advantage in some cases and has failed to employ it sufficiently in others. Through consultation with groups in upgrading areas and those scheduled to be moved to overspill areas, major disturbances were avoided. This accomplishment appears to have been an important one, in that squatter areas had traditionally been centers of unrest. Community groups, for example, were organized at the planning stage into road planning committees, which made recommendations for modifications in road layouts (and the resulting dislocations of certain families) that were acceptable to all parties. The committees have continued to be instrumental in other decisionmaking functions that affect community well-being. Other components, including mutual help, have been less successful. Although community resources were successfully applied to the provision of clinics, other efforts, such as one to provide multipurpose community centers, have not engendered popular support.

The Philippines evaluation team has noted that a marked passivity characterizes group meetings where project plans are discussed. Tondo dwellers seem regularly to follow *barangay* (community) leadership in declaring preferences for particular services or procedures, even when these involve significant physical disruptions in reblocking their homes. Although this form of hierarchical organization, which involves consultation with *barangay* leadership and community groups, appears to have led, as in Zambia, to the circumvention of major social disruptions during the reblocking process, the evaluation team has recommended that more comprehensive information campaigns be conducted before and during relocation so that households understand as completely as possible what options are open to them. The options may include actions which do not necessar-

ily fit the preferences of *barangay* leaders, but inclusion of such choices may lead to a broader-based community participation than the estimable grass-roots movement already achieved.

### Some Implications for Project Design and Policy Formulation

The analysis of the eight project effectiveness components presented in the preceding section does not begin to exhaust the ways in which shelter programs might be evaluated. Strictly comparable data have not been generated in every instance, given the wide variance in procedures and in supply and demand features in the four country projects. This framework for gauging project efficiency nevertheless constitutes a useful tool for beginning to assess crossnationally the options and constraints in housing programs which attempt to serve the needs of low-income and lower-middle-income urban populations. In this section some of the lessons for designing projects and for addressing wider policy issues are discussed.

#### *Market Forces and Flexibility in Design*

A general feature of the findings is that market forces among the poor who seek shelter in rapidly urbanizing developing countries are stronger and more easily released from constraints than had been anticipated. Within the limits determined by household economies, families exhibit varying behavior in housing markets (as is generally corroborated by evidence from other projects). The planning and execution of the progressive development process, and the size, quality, and use of project houses, will vary enormously as well.

In the realm of project design, it has been shown that substantial opportunity costs render it uneconomic for large proportions of target populations to utilize family labor in constructing or improving their houses. It follows that families should not be forced to use their own labor. Nevertheless, the initial decision to encourage use of own labor in self-help processes was a wise one. This conclusion is confirmed by the quality of houses constructed using family labor in whole or in part and by the efficiency of the process for families with relatively low opportunity costs and relatively high levels of construction skills. Similarly, the findings with respect to materials stores, types of materials used, and the amounts of rental accommodations constructed indicate that families make a wide variety of choices, again usually for good reasons. The proper interpretation of these results is that a broad range of solutions should be encouraged, or at least allowed.

In the realm of policy, project experience and evaluation findings confirm that housing markets are quite easily unfettered, and that once they are unfettered, they respond vigorously. This argues for a strategy that makes it possible for more families to participate and that does not seek to constrain the development they choose to undertake. The first consideration implies making tenure, basic services, and housebuilding resources more easily available to lower-income families; affordability and cost recovery findings indicate that this is a risk that can be run. The second consideration implies the removal of restrictions (such as those on rentals) that act, directly or indirectly, to suppress construction.

The range of preferences implied by the above examples argues unambiguously for housing solutions that permit substantial flexibility to households. The optimal degree of flexibility is not known and may be undeterminable, but evaluation findings suggest that project planners should leave more to the market and resist temptations to build too many components into projects. Rather, they should try to be selective and should think of projects as series of interventions that may eventually prompt the system to produce all the desirable sets of complementary components and services.<sup>19</sup>

In this view, then, projects should be kept simple—designed for realistic results and speedy implementation.<sup>20</sup> Key elements such as appropriate locations, reasonably secure tenure, and efficient credit mechanisms must be in place, but beyond these the set of needed components may vary greatly. Far from denying the complexity of urban problems, this approach requires careful and complex analysis to relate objectives to resources and constraints, which leads to more accurate choices about what to do first in a given set of circumstances. Along with continually improving provisions for maintenance, cost recovery, and community participation, this approach may make possible more rapid and more easily replicable improvements in housing for more people.

#### *Affordability and Project Design*

Although affordability was not an explicit objective of early projects, evaluation experience has shown that housing can be produced that is affordable for households in the lowest quintile of most urban income distributions. If social equity objectives as well as efficiency goals are to be met as directly and as rapidly as is desirable, efforts over 1985–95 will have to be more concerted than those already successfully carried out. Though each particular set of circumstances dictates ways in which performance may be improved, evaluation findings suggest four principal ways of improving deliv-

ery of project benefits to the poorest urban dwellers: revise standards, and thus entry costs, to lower levels; increase the opportunities for rental arrangements in projects; tailor credit arrangements to the revealed needs of a population which may have substantial access to informal transfers; and insist on excellent cost recovery performance in projects so designed. Urban project designers in the World Bank have been making considerable progress along these lines in recent years.

The first two approaches to broadening the benefits of shelter projects are not necessarily complementary and need to be handled with sensitivity. Concerning standards, there are two crucial issues: the land density of settlement, and service levels.<sup>21</sup> The first issue is closely tied to plot size, although the two concepts are not synonymous. Given the scarcity of urban land, particularly in the very large metropolitan areas, urban projects should have densities that are as high as possible, consistent with the economic demands of the target populations and with reasonable health standards. Other things being equal, this strategy suggests small plot sizes—smaller, in general, than those considered during 1971–75, when the projects evaluated in this program were being devised. This does not imply less living space per individual or per household, however. Plot size and living area will vary greatly with the amount of vertical construction, which in turn will be determined by a number of considerations, including land price, rental market conditions, the availability of credit, and the incremental costs of constructing additional stories. The relevant factors, and thus the feasibility, profitability, and amount of vertical building, vary from one locale to another.

The early projects evaluated under the present program did nothing to encourage multistory construction, primarily because two of the four projects were in African cities of modest size where populations and land prices were not yet at levels that would force consideration of vertical building. In addition, multistory designs were explicitly avoided to minimize risks in the early pilot projects. There is now considerable evidence, however, that builders in the informal sector are capable of raising structures to substantial heights and will do so if appropriate incentives are present. The task of future projects will be to see that the incentives are there and particularly that potential biases against multistory construction are avoided.<sup>22</sup>

Analyses of available data on the first upgraded areas to be reblocked within the Tondo area in Manila show that adjustments in vertical structure took place in short order. The average lot size was reduced by 12 percent (from 65.3 to 57.6 square meters), but households increased the average constructed area by 34 percent (from 40.2 to 54 square meters) within the first three

months after reblocking. To a large degree this was accomplished by increasing the amount of vertical building: a quarter of the households (and half of those that previously occupied single-story dwellings) added a second story for the first time, and the average number of floors increased from 1.49 to 1.66. Projections on the basis of early trends suggest that this figure may by now have reached about 1.85.

The most up-to-date information on a large urban housing market is for Cairo,<sup>23</sup> where the character of the housing market has been changing rapidly as a result of increasing demand that impinges on an inelastic supply of land in a large metropolitan area. A substantial share of the increase in housing has consistently been accounted for by the informal sector.<sup>24</sup> The expansion has occurred both by extension of the urbanized area and by the addition of stories to existing buildings. As a result, informal housing that is initially constructed with one or two stories eventually reaches two to five stories, and the average number of floors per building has increased from 2.09 in 1976 to 2.45 in 1981. The economics of this process are reasonably clear. The cost of a medium-size informal lot in a medium-price location in Cairo is now approximately two and a half to three times the cost of constructing a 50-square-meter dwelling. Thus, conservatively speaking, unit costs per square meter can be reduced by as much as a third by adding one story and by nearly half by adding two. Furthermore, the direct costs of an additional story will be no more than 75 percent of the costs of building a new single-story house on the minimum-size lot at the periphery.<sup>25</sup>

Evaluation results indicate that a key objective of project design should be to bring (or keep) the unit price and quantity of land within affordable limits for the target population so that the poor are not overextended by participation or the affluent unduly encouraged to participate.<sup>26</sup> Fortunately, there are signs that significant improvements in practices of the recent past are possible. That there is a sound basis for such a favorable outlook is confirmed by project experience in El Salvador, where both the programs of the FSDVM and the use of evaluation are of long standing.

Advances have been essentially of two kinds: improvements in layout, which have increased the proportion of residential area to total area, and stimulation of two-story construction. Improvements in layouts were achieved principally by significantly reducing the proportion of land assigned to vehicular traffic and parking (through restriction of these functions to peripheral areas) and by grouping houses around miniparks or green areas which provide access to the interior plots and serve as semiprivate recreation areas. Families thus trade off some private area for shared use of areas. The

success of these design choices is illustrated by the fact that residential area as a percentage of total area has increased from an average of about 50 percent in the earliest projects in El Salvador to 70–80 percent in the most recent ones.<sup>27</sup>

At the same time, there has been a steady increase in El Salvador in the number of units per hectare of residential area. This figure has slowly increased as plot sizes have been reduced. The greatest increase in unit densities has come in an experimental project (Conacaste) in San Salvador, where a two-story design for some of the units has permitted the construction of 140 units per hectare, as against an average of 80–95 units in other recent projects. The cost-effectiveness of this result appears quite compelling for large urban centers where land prices are sufficiently high to encourage the substitution of capital (construction) for land. This example implies that a two-story unit (25 square meters of constructed area on a 32-square-meter lot) costs about 15 percent less than a 25-square-meter single-story house on a 60-square-meter lot. Furthermore, rising land prices are inexorably shifting the balance further in favor of multistory construction and making such construction economical over more of the urban landscape. Of course, Salvadorean cities are not yet large enough to provoke the strong incentives for vertical construction evident in the examples for very large cities such as Cairo and Manila.

Two other design refinements which can be used to reduce unit costs are the elimination of unnecessary construction costs and the omission of undesired services or service levels. These considerations too are best approached from the perspective of the diversity of demand and require the preservation of a range of options as well as flexibility in their use. The design of core units can be used to illustrate the choices involved. General judgments concerning the desirable extent of the core will depend on the relative benefits and costs of the possible construction methods.<sup>28</sup> Precise estimation of these costs and benefits is an aim of current research, and the findings should be factored into small unit models, such as the Bertaud model cited in note 27, to yield the requisite planning guidance. A further point to be emphasized is that what is shown to be desirable on average is not necessarily desirable for all concerned. For example, even if calculations showed that core housing has on balance an unfavorable benefit-cost ratio, a significant proportion of potential residents might benefit from the option of core housing. Conversely, although the overall benefit-cost ratio might be favorable, that result should not obscure the fact that a significant number of potential participants would benefit from options other than core housing and its corre-

sponding financial obligations. Each possibility must be treated squarely in project design.

There are several related considerations regarding service levels. Water supply provides a somewhat simplistic illustration. The questions in this case will be whether to have individual household connections or some kind of communal supply (such as standpipes) and when to introduce individual connections for a particular populace.<sup>29</sup> The problem is analogous to that of core housing: whenever possible, options should be offered. Because not all households can afford individual connections, access to some form of communal supply should be an option. What proportion of participants requires this option and what the range of options should be are essentially empirical demand questions that are not easily assessed.<sup>30</sup> The planning question is one of designing economical total reticulation systems (including fire hydrant and waste disposal systems) over time.

### *The Rental Option and Other Considerations*

Even with the provision of options and careful control of costs, a substantial portion of families, particularly among the poorest, will have access to project shelter only to the extent that rental accommodations are expanded along with ownership opportunities. This conclusion is based on the fact that a significant proportion of households in all of the countries covered by the evaluation rents rather than owns housing. This finding does not imply that all those renting do so as a matter of preference. The choices in a given situation depend on a particular confluence of supply and demand and will be biased toward rentals in the circumstances of restricted supply and overcrowding that prevail in many, if not most, developing-country cities. The point is not that rental accommodations and ownership opportunities should be expanded in step with each other, nor even that project planners should concern themselves with precise ratios, but only that in most instances both should expand.<sup>31</sup> The evaluation research has not indicated any particular need to stimulate the rental market, although appropriate improvements to credit systems (embracing loans to contractors as well as to purchasers) would increase the elasticity of supply of rental accommodations. The most important action is to ensure against avoidable constraints on rental arrangements, such as the restrictions and prohibitions designed into some of the early projects financed by the World Bank.

There is apparently some degree of resistance on the part of developing-country policymakers to rental options, and project planners must take that resistance into account. The underlying notion that renting does

not contribute to suitable solutions to urban shelter problems seems untenable in light of the evidence that large proportions of households in all countries choose to rent. The notion expresses a fear that the renting poor will be exploited by a rentier class—a phenomenon that exists to some extent everywhere. The best way to minimize this tendency is to expand the total housing supply and ensure, by monitoring for specific abuses, that full cost recovery is enforced so that projects do not subsidize a rentier class.<sup>32</sup> A well-executed strategy of this kind will avoid the feared outcome much more successfully than such misguided policies as rent controls.

Finally, it will be necessary to experiment with means (other than renting out housing space) of enabling the poorest households, particularly, to utilize houses and plots as earning assets, for commercial or small manufacturing enterprises or for growing food.

To reiterate, flexibility is crucial in seeking the mix of components that will serve the housing needs of target urban populations. The outlook for progress along these lines is encouraging. For example, at the same time that it adopted the sites and services concept, the World Bank, realizing the limitations of that concept, began to develop the complementary area upgrading approach. Recent years have also brought a significant broadening of options and modification of concepts under both approaches, leading, for example, to a change from materials loans to construction loans, to acceptance of the need for rental components, and to numerous other adaptations. The best illustration of such adaptability thus far has been in El Salvador, where the FSDVM has modified its mutual help program significantly, pushed its already low standards still lower in some projects, and developed the layout and design innovations mentioned above. Analogous changes have been made in the developing programs of the National Housing Authority (NHA) in the Philippines.<sup>33</sup>

Analyses aimed at assessing the relative merits of the area upgrading and sites and services approaches have suggested that, far from being distinct strategies, they belong to a continuum of possibilities with countless nuances. The processes involved in restructuring a wooden house on a Tondo lot, repositioning it within a new grid established by reblocking decisions, moving it to an adjacent overspill area, or even removing it to the Dagat-Dagatan sites and services area four kilometers away are at least conceptually similar. To be sure, time, distance, cost, and the extent of retention of the original structure all vary. The point to be stressed is that they vary along a continuous frame of reference or response and that the differences are frequently very small. Leaving aside the extremes, the principal effort should be to focus policymakers' attention on this continuum and on

the scope for improving the total housing stock progressively over time, rather than on the fine-tuning involved in selecting particular types of projects.

Along with these changes in design, which lead to simpler and lower standards and costs and an appropriate balance between ownership and rental options, should come concomitant changes in credit facilities. Not much is known yet about the real demand for credit, and experimentation has only begun, but evaluations of experience do permit the statement of a few principles. Not all participating families will require institutional credit, and the design of credit programs should be directed to those that do. The size of credit programs should be reckoned in terms of the numbers who need credit and their requirements rather than a rationing of supplies among the entire population. Credit should also be apportioned to builders who contemplate rental accommodations, not only to owner-occupiers.

### *Cost Recovery*

These remarks lead to a final, related observation: shelter program maintenance, collections, and cost recovery—and their interrelations—should be vital considerations for designers and managers. They should be dealt with from the beginning of project identification, and projects should not be executed that do not have a favorable relation between revenue extracted and benefits delivered as well as between benefits and costs. For replicability, programs must be designed so that they return, directly and indirectly, approximately as much to public revenues as they require in expenditures.<sup>34</sup> Beyond this, more must be done to make policymakers aware of the potential that projects offer for improving the local fiscal situation. In particular, projects should be used to prompt fuller cost recovery not only among the target groups themselves but, even more important, among higher-income strata if, as in Zambia and many other countries, their housing and services are being subsidized.

It is not claimed here that this is an easy task. On the contrary, precisely because it is difficult, further experimentation is necessary, and planning should include considerable experimentation with various forms of community participation and contribution. Fortunately, the record on this issue is encouraging. Just as they have a desire and capacity to improve their own homes, the urban poor have a desire and capacity for greater participation in their own governance. In the past these capacities have been left largely untapped by program designers, but it has been demonstrated that projects can make effective use of them. More can be done, however, to facilitate project implementation, im-

prove the record on collections, and articulate demands for specific components in future projects within an overall strategy for progressive social change and development. Judicious involvement of the community at all these levels must be more fully explored in planning, project design, evaluation, and associated research.

### The Contributions of Evaluation

All of the above findings have stemmed in large measure from the World Bank's Evaluation Program.<sup>35</sup> Evaluation and the associated research have contributed to the advancement of project design and policy formulation in four principal ways. First, evaluation has strengthened the modeling paradigms which frame the design of both projects and research. Second, it has contributed over time to better identification of the uses and users of all management information, including that provided by evaluation. Third, it has contributed to the refinement of indicators and measurements. Fourth, it has helped to break new ground with respect to certain relations between key variables. Each of these contributions will be discussed briefly.

By providing a rigorous framework, evaluation has contributed to keeping a framework in place notwithstanding alterations that were desired or required as project and program lessons were learned. By contributing both to the alterations and to the stability of the basic edifice, evaluation has helped to bring the framework into clearer focus. Examples of such contributions to better understanding include putting self-help and mutual help into appropriate perspectives, identifying the roles of rentals and transfers in the development process, and spelling out the needs for designing better housing finance milieux and credit schemes.

In the realm of identifying uses and users, evaluation has operated to narrow what was a sizable gap at the outset. The initial evaluation design, which was based on quasi-experimental design and statistical rigor and which sought to address a large proportion of the relevant questions, was extremely ambitious and excessively demanding. It was in everyone's interest to seize opportunities to improve the focus and become more selective as the program progressed and each cycle of findings was reviewed. Equally important, process evaluation, which did not play a major role in the beginning, was blended in progressively over time. Panel studies, which did form part of the original design, were used more as time went on. As evaluation results attracted managers' attention, the managers asked evaluators to study and report on increasing numbers of immediate problems. This required increased use of and experience with evaluability assessment, rapid feedback, and other

process evaluation techniques. More recently, the World Bank initiated a first experiment that employs participant observers in the field.

An illustration of improved and extended measurements is the demand studies carried out in El Salvador and the Philippines under the Evaluation Program. Along with similar studies in Korea and Colombia under the World Bank's Research Program, these have helped to lay the groundwork for a subsequent generation of urban projects which, by virtue of improved estimates, may be designed and priced more nearly in accordance with the demands of the intended beneficiaries. Also as a result of this work, additional research is under way in these and other countries to extend the range of circumstances (income, climate, city size, culture, and so on) for which there are reasonable demand estimates. In addition, the pioneering work in this program on hedonics analysis as applied to housing has demonstrated the overall cost-effectiveness of the progressive development approach at the same time that it has begun to develop information on the relative contributions of various components to overall project benefits.

The Evaluation Program can truly be said to have pioneered the investigation of transfers within the extended family and other kin and nonkin networks. That such transfers can be important was first discovered somewhat incidentally in a study of control groups in Dakar, Senegal, in 1977. The phenomenon was thereafter followed closely in all projects, particularly in the Tondo project in Manila, the Philippines, where the budgets of a small sample of families were followed closely over three years. Data from the evaluations in El Salvador, the Philippines, and Senegal, as well as from a related investigation in Cartagena, Colombia, have been used in several research papers. These have demonstrated, among other things, that transfers are very important in the budgets of the poor, making up close to half the total incomes of half the families in the lowest third of the income distribution; that they tend to flow from better-off to poorer households within the spectrum of the urban poor; and that they tend to be sensitive to household composition and employment status (households headed by females or unemployed workers are more likely to receive transfers than those headed by employed males) as well as to shifts in income. Furthermore, it appears that there is a basic underlying contract by which resources are provided to poorer households to meet basic needs or to improve housing when opportunities such as these projects afford are present. Currently the research is being extended outside the evaluation framework to exploit the best data bases and the most encompassing sample frames available. This work also addresses the important issue of urban-rural transfers.

It is worth noting in closing that these concluding

statements can be made and documented in large measure because not only the projects being evaluated but also the evaluation program itself have been under constant scrutiny. Thus, evaluation methods and processes were evaluated along with evaluation results at each of six annual conferences in which project managers, field evaluators, and responsible officials of the sponsoring agencies participated. A detailed record of these discussions and the consequent decisions has been kept. This made possible the meta-evaluation, or "evaluation of evaluation programs," on which these concluding remarks have been based.

## Notes

1. Progressive development is a method of housing construction or upgrading that is achieved through staged development over extended periods and involves considerable flexibility in housing design, materials used, and the family's (self-help) contribution to the construction process.
2. At the time, experience with such approaches to shelter development was sparse, and the bulk of information—most of it, at this stage, speculation—came from observation of spontaneous "invasions" of land by squatters in several countries. Empirical research had been extremely scanty, and rigorous evaluation of experience was nonexistent.
3. It has not been possible as yet to analyze whether present levels of project costs and standards systematically contribute to underrepresentation of lower-income households from the start of projects.
4. In some interior cities of El Salvador lower population densities and good access to well water have made the larger plots and fewer services of the *colonias* more attractive than the options designed by the project agency on the basis of its earlier experience in the larger cities.
5. Reblocking is the process of applying, insofar as possible, normal subdivision processes to the less orderly division of space that results from most illegal occupations of land by squatters. The physical subdivision process requires a counterpart process of mapping prior "claims" in the old order into accepted titles in the new.
6. Explicit selection does not occur within area upgrading projects, since they typically deal with pre-existing populations.
7. In El Salvador and Zambia, where the genuine concern of project administrations that the lower income groups be reached predated any involvement by international agencies such as the World Bank, the selection procedures appear to have had a particularly good record of selecting families from within prescribed income ranges.
8. This discrimination has been corrected insofar as possible.
9. What is more, contrary to what one might expect from experience in developed countries, where formal sector employment predominates, for the majority of participants in the

El Salvador projects evaluated, weekends were the least favorable times to participate in mutual help, as this was when their (informal sector) earnings were typically the highest.

10. Use of the term *demand* has been avoided up to this point, for very little is known about it. This was true when these projects were designed, and although the research program has contributed materially to great advances in the understanding of the demand for housing in recent years, the (derived) demand for credit has been little studied to date.

11. Furthermore, the evaluation's findings about income transfers indicate that the two types of restriction may be incompatible. That is, since families which are alike in other respects have very different access to transfers, they must be supposed to have commensurately different demands for credit. If this is so, a given loan fund, whether or not adequate in itself, will not be optimally allocated among families by rationing a small amount to each.

12. The evaluation findings are so far consistent with this finding (see Jiménez 1982). In addition, the administrative mechanism for ensuring materials purchase may carry its own restrictions and added costs.

13. This paragraph encapsulates the frustrations experienced in the Senegal project, the only one of the four to experience serious difficulties.

14. A social problem, however, may still exist. If significant numbers of families that have moved to new sites fail to build permanent structures within a reasonable period, it may be difficult to obtain and sustain policymakers' support for the sites and services option.

15. Another way of viewing this tradeoff is that if a family is forced into this situation and chooses nevertheless to construct a house with its own labor, the cost of this self-help method should include the cost of rent during the construction period—which, in El Salvador, has run up to six months. This added cost would not be present, or present to the same extent, in some of the "higher-cost" options, for example, lots with core housing.

16. In this connection it is noteworthy that in the same project consideration has been given to vesting the responsibility for maintenance of communal standpipes, where similar problems could be anticipated, in community groups. It has also been suggested that if a separate garbage collection fee existed, garbage collection services might be improved. The adequacy of service would then affect, and be affected by, payment of this fee alone and not the entire cost recovery effort.

17. For additional details on project collection efforts and on the constraints on the LCC's handling of arrearages, see Bamberger, Sanyal, and Valverde (1982), chs. 9 and 10.

18. In one case the FSDVM promoted, or at least actively encouraged, a protest directed against the recalcitrant state water utility. This demonstration inadvertently provided a model for organizing pressure on the Fundación itself later.

19. A model to guide practice in the adaptation of project components to actual demand is not yet available. A high priority should be given to developing one.

20. As pointed out above, there were substantial delays in the implementation of projects already evaluated. More recent experience indicates, however, that implementation delays for

urban projects are very near the mean for all Bank-financed projects.

21. The timing of the introduction of various services is a third, ancillary consideration.

22. Policymakers are sometimes reluctant to promote or even to permit two-story construction on the grounds that a significant proportion of families would be unable to build "suitable" additional stories. Evaluation findings that show the extent to which families hire others to do their building should help to redefine this risk.

23. Mayo (1982). Comparable information is becoming available under a World Bank research project on one or more cities in Colombia, El Salvador, India, Indonesia, Korea, and the Philippines.

24. Informal housing—defined for the purposes of this study as housing which has been constructed without official permission to convert land from agricultural use, without a building permit, or in violation of building codes—constituted about 90 percent of new housing starts in 1971–76 and 75 percent after 1976.

25. Infrastructure costs for building to these implied densities may, however, reduce the advantages of the central location.

26. Although the plot sizes of past projects have generally been acceptable to the target populations, little information is available as to optimal plot size distributions in projects. This important area should be given priority in evaluation and related research, particularly in the developing world's largest and most rapidly growing cities.

27. Such success has been one factor that has motivated the World Bank to develop small unit planning models to deal efficiently with tradeoffs in project layout design. See, for example, PADCO and the World Bank (1981).

28. Such calculations will include the benefits, if they can be demonstrated, of providing some amount of core housing to minimize reticulation costs, encourage use of sanitation facilities, or stimulate construction by making it easier to move in at once.

29. There are, of course, many other factors which affect unit costs, such as capacity, hours of service, and quality of water. Attention is limited here to factors which directly affect project area design.

30. Demand analyses from all four countries in the original evaluation program show that significant proportions of the target group are satisfied with communal water supply and pit latrines (Zambia project participants), choose low service level options if they are available (El Salvador and Senegal control groups), and place a relatively low hedonic evaluation on sanitation facilities (Philippines project participants).

31. There may be instances where crowding has been of such proportions that a significant expansion of owner-occupied housing will imply that the absolute number of renting families decreases.

32. To guide this aspect of the program, analyses of the

indirect effects on housing markets of changes in supply wrought by large projects need to be performed within evaluation research.

33. Few specified examples can at present be cited for the other two countries in the evaluation program. This is principally a result of their later phasing and slower development and the absence to date of second projects. Nevertheless, a similar pragmatism has been evinced in response to problems encountered in the implementation of the first projects.

34. This does not imply that there should be no subsidy from higher-income groups, either within projects or over a broader spectrum, but rather that the subsidy should be on a scale that can be sustained in a large program over time.

35. This is not to claim that operational personnel would have remained in ignorance without a rigorous evaluation program. Sometimes the first insights have come from the evaluation program, but just as often evaluation has merely confirmed suspicions that arose from project experience. The key point is that, by virtue of the rigor of its approaches, evaluation has assisted the learning process to proceed more confidently—and that is an important contribution.

## Bibliography

- Bamberger, Michael, Edgardo González-Polio, and Umnuay Sae-Hau. 1982. *Evaluation of the First El Salvador Sites and Services Project*. World Bank Staff Working Paper 549. Washington, D.C.
- Bamberger, Michael, Bishwapura Sanyal, and Nelson Valverde. 1982. *Evaluation of Sites and Services Projects: The Experience from Lusaka, Zambia*. World Bank Staff Working Paper 548. Washington, D.C.
- Jiménez, Emmanuel. 1982. "The Economics of Self-Help Housing: Theory and Some Evidence from a Developing Country." *Journal of Urban Economics*, vol. 11, no. 2 (March), pp. 205–28.
- Kaufmann, Daniel. 1982. "Social Interactions as a Strategy of Economic Survival among the Urban Poor." Ph.D. dissertation. Harvard University, Cambridge, Mass. Processed.
- Keare, Douglas H., and Emmanuel Jiménez. 1983. *Progressive Development and Affordability in the Design of Urban Shelter Projects*. World Bank Staff Working Paper 560. Washington, D.C.
- Keare, Douglas H., and Scott Parris. 1982. *Evaluation of Shelter Programs for the Urban Poor: Principal Findings*. World Bank Staff Working Paper 547. Washington, D.C.
- Mayo, Stephen. 1982. "Informal Housing in Egypt." Cambridge, Mass.: ABT Associates.
- PADCO and the World Bank. 1981. *The Bertaud Model: A Model for the Analysis of Alternatives for Low-Income Shelter in the Developing World*. World Bank, Urban Development Department Technical Paper 2. Washington, D.C.

## *Urban Economic Policy: Directions for the Future*

George S. Tolley and Vinod Thomas

The past quarter-century has witnessed rapid urbanization in the developing world. Today the developing countries contain fifteen of the twenty-five cities with populations over 5 million. Although urbanization is associated with economic growth in the minds of many, the scale of urbanization in developing countries is also viewed with alarm. The observed concentration of poverty, unemployment, housing, and transport problems in some highly built-up urban centers such as Mexico City, São Paulo, Bombay, Cairo, Jakarta, and Manila has given rise to fear of the consequences of present trends.

The rate and scale of the recent urbanization phenomenon have been so large that they are not yet adequately understood. Urban theory offers explanations of the workings of urbanization in developed economies: theories of housing markets, locational decisions, traffic and transport modes, and labor markets have been advanced. Are the same analytical approaches applicable in developing countries as well? Is developing-country urbanization a distinct phenomenon? Are there urban policy conclusions for developing countries that differ markedly from those for the industrial economies of today?

This book has emphasized similarities in urbanization in developing and developed countries but has also brought out important differences. Many of the analytical methods adopted in this work were developed in the context of experience in advanced countries. Basic behavioral and economic explanations behind rent and density gradients and city formation are relevant in developing as well as developed countries, and even the range of parametrical estimates of some key demand variables in the urban economy may be similar. The

forces of urbanization that originate from population growth, urban-rural productivity, and composition of demand increases are also widely applicable. The book has laid out some of these factors in the developing-country context to assist in an understanding of urbanization trends. The authors differ in their emphasis on the relative importance of these forces, perhaps because the experiences of the countries examined vary.

Some of the differences between developed and developing countries are perhaps more interesting than the similarities. Even if the tools of analysis are not necessarily different, the present-day urbanization experience in the developing world brings to focus more complex and difficult, if not different, policy issues. Urban problems in developing countries are more severe than those in the developed countries today because of the former's much faster rate of urbanization at present. In addition, the problems are also bigger than those that the advanced economies experienced when they were developing. The main reason is not that the pace of urbanization is different now than earlier, but that urban populations and agglomerations are larger today, partly because of larger overall population. Moreover, in developing countries today intraurban transport costs are relatively lower than those in previous time periods and allow flatter rent gradients within cities, larger city sizes and more concentration. (Reductions in intracity transport costs may have been greater than those in intercity costs.) Urban centers in developing countries also seem to have been able to achieve greater technology transfer and productivity gains than rural areas, which has caused correspondingly stronger pressures to urbanize in an increasingly interrelated world economy. Individual countries obviously face widely differing cir-

cumstances; factors range from low and stagnant agricultural productivity in many African countries to special city incentives and government urban subsidies that provide added impetus to urbanization in many parts of the world.

Although the urban economy is exceedingly complex, this book has adopted the view that an understanding of economic behavior in the urban context is more useful than is the construction of large-scale urban models. Much more comparative analytic work remains to be done, however, in promoting a better understanding of developing-country urbanization. The urban-rural linkages vary among countries, as do the importance of macroeconomic policy on urbanization and the impact of regional considerations that affect urbanization. Finally, international trade is an important determinant of the nature and composition of production and population within a country. Understanding these phenomena in greater detail constitutes an agenda for further work.

### Urbanization Problems in Mixed Economies

Urban problems are best understood in the context of the economic development of nations. Urbanization as such is neither the source nor the enemy of development. The mainsprings of productivity growth make agglomeration and urbanization attractive and bestow net economic benefits, at least for some time. Urbanization can proceed too rapidly, however, and make some developing-country cities too big from the economic and management points of view. Some important aspects of the overurbanization issue are negative externalities, such as crowding, pollution, and transport failure, which represent a drain on the benefits associated with urbanization. Overurbanization, if it occurs, is difficult to correct and may best be addressed through actions aimed directly at the externalities rather than at city size.

Often urban poverty is identified as the urban problem, and quick-fix solutions are sought to eradicate it. The basic reasons for urban poverty, however, relate to overall poverty in the country and to inadequate gainful employment. Nationwide policies for generating long-term income and employment and for promoting a smoother functioning of labor markets are perhaps the best strategy for reducing urban poverty. Sometimes informal sector activities, such as shoestring enterprises and street vending, are falsely viewed as unproductive, and their elimination is associated with poverty alleviation. Again, the level of economic development of the country determines the mix of employment opportunities, and informal activities could represent a valuable

source of employment and incomes which should not be discouraged.

Antipoverty efforts in urban areas are best served over the long term by employment and labor market policies that increase labor productivity and demand more rapidly than labor supply, with special emphasis on lower skill levels. Policies to assist labor mobility across occupations and locations are also vital. In this context the urban employment and poverty record can be improved in many developing-country settings by measures to reduce the natural growth of population and to improve labor force participation and education. Massive rural-urban migration is often held responsible for urban poverty and unemployment, but the migration issue is a complex one in which urban-urban outflows play as important a role as rural-urban migration. The economic incentives to migration derive from the economy-wide context, and measures at an individual city level—such as restrictions on informal activities and squatter settlements and withholding of urban infrastructure for new entrants—are likely to be ineffective and inefficient. Urban subsidies for production and consumption often serve to increase city inflows, and a reduction in such special incentives could help to keep migration flows at manageable levels.

### Lessons under Central Planning

China's experience illustrates the difficulty of influencing income gaps between the city and the countryside. China has not relied on migration and seems to have worked much harder on more direct measures than other societies. Through the encouragement of bootstrap operations villagers have been induced to help build an infrastructure of waterworks, roads, level fields, schools, public health care, and administration. The government has tried to help directly by raising prices for agricultural products, but a closer examination shows that the drive toward rapid industrial growth has helped shape government investments and industrial prices to the detriment of agricultural growth. And when it comes to state-subsidized services, including housing, health care, education, and cheap food, it is still the urban sector that has reaped the greatest share of the benefits. That these conditions are found in many other developing societies seems to illustrate the continuing problems of urban and industrial bias even in a society that has overtly renounced such a policy.

China has now begun to correct some of these difficulties. The government has helped induce a rapid rise in rural incomes over the past few years by again raising rural purchase prices, by increasing agricultural investment, and by removing several other bureaucratic con-

straints on farmer activities. The rise in rural incomes may still not be greater than the increase for urban dwellers, but absolute living standards have begun to improve significantly. Some of the gap between city and countryside may be erased less by helping agriculture itself than by allowing peasants to open more small industry in rural market centers. This phenomenon has not been discussed in this book because many of the market centers are excluded from China's urban statistics. Yet with the prohibition on migration to cities, these small market centers seem to be providing an outlet for peasants which exerts a positive influence on rural incomes and possibly leads to a diffuse pattern of urbanization throughout the countryside.

More peasants may also be slipping into cities. With the increasing role of peasant free markets in cities, the reduction in the number of rationed goods in some cities, new construction activities that can use unskilled labor, and a general relaxation of bureaucratic rule, the number of illegal migrants in cities may be growing. Increasing foreign trade may also create pressure for such major port cities as Shanghai to grow in a way that they have not in the past. All these forces may lead to a slightly different Chinese urban structure than we have seen in the past three decades, but much of the basic structure is likely to remain intact. Centralized control over investment, state jobs, and many rations continues to give the government a way of controlling the growth of cities. And much of the new growth of rural market centers would promote a more balanced pattern of urban growth, similar to that witnessed in earlier decades.

### Urban Concentration and Urban Policy

The distribution of city size is a subject of intense public debate in countries with divergent economic managements. The issue—a variant on the concern about urban concentration—relates mostly to the alleged polarization of economic activity in one or a few urban centers. Such concentration does often exist, as in many Latin American countries; it is less marked in India. Considerable work has been done on the classification of cities according to the structure of the economic base, but a classification of desirable city size distributions has yet to emerge.

Nationwide planning for a certain city size distribution does not seem promising, even if such an outcome might look attractive. Even limited efforts to create cities and growth poles have turned out to be excessively costly, as have policies to promote concentration. A macroeconomic policy posture that does not favor or penalize urbanization seems to be the most desirable

approach from an economic viewpoint. Experiences with the concentration incentives in Brazil and the decentralization inducements in Korea and Venezuela which are documented in this book lend support to this view.

The indirect urban effects of economic policies are often ignored and are not well understood. Many countries in Asia and Africa have historically adopted price policies that hurt agricultural productivity and thereby indirectly favored urban production. Consumer subsidies in urban areas have been widespread. Macroeconomic policies have also indirectly and unintentionally promoted a concentration of production of nontradable goods in urban areas. An overvalued exchange rate implicitly taxes tradables and shifts resources to nontradables. To the extent that the agricultural sector contains relatively more tradables than does the rest of the economy (including services), on average, such an exchange rate policy would promote nonagriculture in relation to agriculture.

A reversal of the process of urbanization is unlikely to be feasible, nor is it desirable as national policy if economic development is to be sustained. At the same time, it should be noted that a country might wish to pursue some type of decentralization of economic activities for reasons not readily measurable in economic terms. Often, large concentrations of people and activities present political problems. Another motive is that a larger number of smaller cities may be more desirable, from a defense and strategic viewpoint, than a few densely populated large cities. One problem is how to take account of these goals legitimately rather than letting them be used as mere rationalizations for undesirable or ineffective policies.

A relevant issue seems to be whether and how to marginally affect the degree and nature of urbanization, including the size distribution of cities. Clearly, political, social, and economic pressures influence policy options, and policy choices are seldom straightforward. Nevertheless, the analysis of urbanization in this book yields four types of suggestions for economic policy.

First, policies might attempt to take account of externalities that inadvertently affect urbanization. Externalities—whether negative or positive—are not usually accounted for in the activity of the private sector. Public policy might correct this failure by, for example, adopting appropriate pricing policies for public services and taxing public ills such as pollution and congestion. Such interventions are likely to influence the locational choices of people and economic activities.

Second, government policy could address vital infrastructural issues. Bottlenecks in transport and other infrastructure constrain urban development across regions and affect the size distribution of cities. The key

issue concerns the optimal level and composition of investment in infrastructure. In this connection, improved benefit-cost evaluation techniques and efficient application of policies would be essential.

Third, direct incentives to influence urbanization and city size might be introduced only selectively in cases where net benefits can be shown. Policy experience to date is not encouraging. The challenge, therefore, is to determine in which instances direct government intervention to promote or retard city size is warranted. A fundamental need is the development of rigorous and reliable means of making quantitative estimates of the effects of these policies.

Finally, several urban problems require direct solutions. Policies for greater efficiency in urban housing, transport, finance, and other areas are often needed. The experience with large-scale housing projects and capital-intensive transit systems has been unpromising. The evidence in this book supports smaller and more divisible projects which respond more directly to the demands and financial means of consumers. Expenditures on rehabilitation, operation, and maintenance of urban facilities, in addition to capital deepening, can have particularly high payoffs. Efficient financial and economic management of cities as urban centers reach extraordinary sizes is an increasing challenge. The relation between urban projects and policies on the one hand and management efficiency in cities on the other is

a crucial question. In better understanding this issue, knowledge from other disciplines needs to be combined with economic considerations, and intermediate goals must be contrasted with those connected with long-term economic development.

## Bibliography

- Findlay, Sally. 1977. *Planning for Internal Migration*. Washington, D.C.: U.S. Bureau of the Census.
- Ingram, Gregory K., and Alan Carroll. 1981. "The Spatial Structure of Latin American Cities." *Journal of Urban Economics*, vol. 9, no. 2 (March), pp. 257-73.
- Linn, Johannes F. 1983. *Cities in the Developing World: Policies for Their Equitable and Efficient Growth*. New York: Oxford University Press.
- Mazumdar, Dipak. 1976. "The Urban Informal Sector." *World Development*, vol. 4, no. 9, pp. 655-79.
- Mills, Edwin S., and Charles M. Becker. 1986. *Studies in Indian Urban Development*. New York: Oxford University Press.
- Mohan, Rakesh. *Labor Force Participation in a Developing Metropolis: Does Sex Matter?* World Bank Staff Working Paper 749. Washington, D.C.
- Renaud, Bertrand. 1981. *National Urbanization Policy in Developing Countries*. New York: Oxford University Press.
- Schaefer, Kalman, and Cheywa Spindel. 1976. *São Paulo: Urban Development and Employment*. Geneva: International Labour Office.

The most recent World Bank publications are described in the catalog *New Publications*, which is issued in the spring and fall of each year. The complete backlist of publications is shown in the annual *Index of Publications*, which contains an alphabetical title list and indexes of subjects, authors, and countries and regions; it is of value principally to libraries and institutional purchasers. The continuing research program is described in *The World Bank Research Program: Abstracts of Current Studies*, which is issued annually. The latest edition of each is available free of charge from Publications Sales Unit, The World Bank, 1818 H Street, N.W., Washington, D.C. 20433, U.S.A., or from Publications, The World Bank, 66, avenue d'Iéna, 75116 Paris, France.





**THE WORLD BANK**

Urban problems, already acute in the developing world, threaten to become more severe as population growth and rural-to-urban migration continue. To promote a better understanding of urbanization, this volume offers a framework for analyzing urbanization processes and evaluating policies to cope with urban problems. Among the topics considered are the causes of rural-urban migration, the connection between urbanization and economic development, urban poverty, differences in country experiences, the goals and effects of policies that attempt to influence urbanization, and urban management issues in such areas as housing, transport, finance, and pollution. Approaches in both market and centrally planned economies are examined.

Written by experts from the World Bank and the academic world, the studies build on progress to date and make new contributions that should be of value to policymakers, project analysts, scholars, and, in general, persons interested in urban problems. Introductory and concluding chapters by the editors draw together the findings and bring out policy implications.

George S. Tolley is a professor of economics at the University of Chicago. Vinod Thomas is chief of the Country Analysis and Projections Division, Economic Analysis and Projections Department, The World Bank.

**WORLD BANK PUBLICATIONS OF RELATED INTEREST**

**Metropolitan Management: The Asian Experience**

*K. C. Sivaramakrishnan and Leslie Green*

Published for the World Bank by Oxford University Press

**Studies in Indian Urban Development**

*Edwin S. Mills and Charles M. Becker*

Published for the World Bank by Oxford University Press

**Cities in Conflict: Studies in the Planning and Management of Asian Cities**

*John P. Lea and John M. Courtney, editors*

**Cities in the Developing World: Policies for Their Equitable and Efficient Growth**

*Johannes F. Linn*

Published for the World Bank by Oxford University Press

**Urban Land Policies: Issues and Opportunities**

*Harold B. Dunkerley, editor*

Published for the World Bank by Oxford University Press

**National Urbanization Policy in Developing Countries**

*Bertrand M. Renaud*

Published for the World Bank by Oxford University Press

**ISBN 0-8213-0786-X**