Chapter 11

MIGRATION AND URBANIZATION

JEFFREY G. WILLIAMSON*

Harvard University

Contents

1. The problem 426
2. The urban transition 427
   2.1. Quantifying the urban transition 427
   2.2. Migrant selectivity bias 430
   2.3. Selectivity bias, the brain drain, and remittances 431
   2.4. City growth, migration, and labor absorption 433
3. Disequilibrating labor market shocks and equilibrating migrant responses 433
   3.1. Disequilibrium and wage gaps 433
   3.2. Are migrants rational? 435
   3.3. Are there too many city immigrants? 437
4. What does "overurbanization" mean? 439
   4.1. Push, pull, and the engines of city growth 439
   4.2. The urban bias 441
5. How do urban labor markets work? 442
   5.1. The evolution of conventional wisdom 442
   5.2. The Todaro model 443
   5.3. Critique: How do urban labor markets really work? 445
   5.4. Some evidence 446
6. Migration and city growth in general equilibrium: What are the driving forces? 449
   6.1. What drives Third World migration and city growth? 449
   6.2. Modeling migration and city growth 453
   6.3. Understanding the past and projecting the future 455
7. Where do we go from here? 459

References 461

*The research underlying this paper has been financed by NSF Grant No. SES-84-08210. The helpful comments of Charles Becker, Jere Behrman, Barry Eichengreen, Peter Lindert, Mark Montgomery, Richard Sabot, Oded Stark, Rick Steckel, Moshe Syrquin, Michael Todaro, and the participants in the Handbook of Development Economics project are greatly appreciated.

Handbook of Development Economics, Volume I, Edited by H. Chenery and T.N. Srinivasan
© Elsevier Science Publishers B.V., 1988
1. The problem

What explains the timing and the extent of the transition from a traditional rural to a modern urban society? Why does city growth speed up in early development and slow down in later stages? What role does migration play in the process, and do migrants make rational location decisions? Do urban labor markets serve to absorb urban immigrants quickly? Are rural emigrants driven by “push” conditions in the countryside or by “pull” conditions in the cities? Is the Third World “overurbanized”?

Speculation on these issues has never been in short supply. Friedrich Engels (1845, 1974) thought that Manchester’s booming growth in the early nineteenth century—and the urban decay associated with overcrowding—could easily be explained by the rapid development of Britain’s manufacturing. Ravenstein (1885, 1889) and Redford (1926, 1968) thought that rural–urban migration, and thus city growth, was conditioned by Malthusian forces, agricultural land scarcity, and enclosure. In short, Engels favored “pull”, whereas Ravenstein and Redford favored “push”. Meanwhile, Mayhew (1861) documented low-wage (“informal sector”) labor in London, viewing these urban poor as a reserve army driven there as an employment of last resort.

Despite a century and a half of debate, social scientists are still uncertain about the quantitative sources of the urban transition, how it can be influenced by policy, and if so whether it should be influenced by policy. While successful industrialization clearly fosters urbanization, what accounts for the “explosive” city growth in the Third World since the 1950s? The two principal hypotheses advanced in the literature are that rapid city growth and urbanization can be explained primarily by (1) unusually rapid rates of population growth pressing on limited farm acreage, pushing landless labor into the cities; and (2) economic forces pulling migrants into the cities. In the contemporary developing world these latter forces include: domestic policies that distort prices to favor cities (e.g. the domestic terms of trade have been twisted to “squeeze” agriculture); cheap energy prior to the first OPEC shock favoring the growth of energy-intensive urban sectors, thus creating urban jobs; the diffusion of technology from the developed world, which favors modern, large-scale urban industries; foreign capital flows into urban infrastructure, housing, power, transportation, and large-scale manufacturing—further augmenting the growth of cities in the Third World; and the liberalization of world trade since the late 1950s, which has stimulated demand for manufacturing exports produced in Third World cities.

Most demographers favor the first hypothesis. Exploding numbers of people must be employed, and a marginal agriculture with quasi-fixed arable land stocks cannot offer sufficient employment for the Malthusian glut created by the demographic transition. Marginal survival by hawking urban services may be the
only way a social system can absorb the population glut, and squalid urban living conditions have been an attribute of early stages of industrialization since Engels wrote of Manchester in the 1840s. The demographer, writing in the shadow of Malthus, is likely, therefore, to favor a causal sequence running from a population boom, to labor pushed off the land, to city immigration, and thus to rapid urban growth under squalid living conditions. This view has also had a profound influence on economists' thinking about development. It is central to Lewis's (1954) labor surplus model—a model that also worked well for the classical economists developing their paradigms of growth during the British industrial revolution. It is also central to the Todaro (1969) thesis that rising immigration to the city is associated with high and even rising rates of urban unemployment. On the other hand, most economists now tend to favor the second hypothesis; that is, an emphasis on those economic forces which contribute to urban pull.

The literature on Third World migration and urbanization is enormous and growing. Development analysts have learned much since we began to worry about such issues in the 1950s. A comprehensive review of that literature is beyond the scope of this chapter; nevertheless an assessment of the most important components is possible. Section 2 begins with the demographics of the urban transition. Economists should know these "stylized facts", but, as it turns out, their perception of the facts is often in error. Section 3 examines the labor markets which link city with countryside. The size and persistence of rural–urban wage gaps and their implications for national income allocative losses are explored. This is followed by an assessment of the extent to which migrants actually respond to the gaps. From private rationality we then turn to social optimality including the questions of whether there are "too many" city immigrants in the Third World and the desirability of suppressing migration through government policy. Section 4 dwells on the sources of city growth and the so-called Hoselitz thesis: Is manufacturing an engine of growth in the Third World and, if so, to what extent is that growth fueled by an "urban bias" in government policy? Section 5 raises the question of how urban labor markets work. Here, the main debate since 1969 has been over the Todaro thesis. This section explores the evidence and contrasting policies which have emerged from the debate. Section 6 returns to the central issue raised at the beginning of this chapter: What drives Third World city growth? It turns out that computable general equilibrium models of city growth offer an excellent device for organizing an answer to that question.

2. The urban transition

2.1. Quantifying the urban transition

The sixteen large Third World cities surveyed by Sinclair (1978, p. 15) had growth rates over the quarter century 1950–1975 ranging between 2.4 (Calcutta)
and 8.3 (Seoul) percent per annum. The average growth rate was 5.4 percent—a very high rate of growth indeed. Furthermore, by the end of the 1960s, slums and uncontrolled settlements housed a large share of most urban populations, almost 43 percent on average, ranging from a low of 14 percent in Amman to 90 percent in Addis Ababa. Such rapid growth implies rising density, congestion and urban land scarcity. Not surprisingly, rents have risen, living conditions have often deteriorated, and the delivery of public services has been minimal at best. Immigration, of course, played a key role in the process throughout the 1950s and 1960s, accounting for about 58 percent of population growth in eleven major cities reported by Sinclair. The range was from a low of 33 percent in Bogota to a high of 76 percent in Abidjan.

While such detailed demographic accounts are useful, there are 837 cities in the Third World with populations in excess of 100,000, so aggregation is necessary if analysis is to move from anecdote to "stylized fact". Furthermore, the Third World performance ought to be compared with earlier industrial revolutions if we are to isolate conventional from unusual city growth experience. Given the common alarmist view that Third World societies have "overurbanized", a brief look at the past might be a useful antidote. The most careful quantitative assessment can be found in a United Nations' publication, *Patterns of Urban and Rural Population Growth* (1980), written by Samuel Preston, as well as his own summary [Preston (1979)] of the volume.

In the first place, Third World city growth is not a new problem, but rather seems to have followed a relatively smooth "urban transition" since the late nineteenth century. While the urban share of the Third World's population rose sharply between 1925 and 1950 (from 9.3 to 16.7 percent), it had been on the rise since 1850. And while the share rose even more sharply between 1950 and 1975 (from 16.7 to 28 percent), Third World urbanization over the past century seems to obey the usual rules of logistic curves tracing the diffusion of new technologies. The point of inflexion in this case appears to have been in the quarter century terminating just after the first OPEC shock. City growth rates trace out a similar pattern, although, as reported in Table 11.1, there is considerable variety by level of development: Latin America reached a peak rate of city growth in the 1950s, East Asia (excluding China and Japan) in the 1960s, Africa in the late 1970s, and South Asia (according to projections) in the 1980s.

In contrast with earlier assertions by Hoselitz (1955, 1957) and Bairoch (1975), we now know that Third World urbanization experience has been fairly conventional by historical standards. Between 1875 and 1900, currently developed countries' urban share rose from 17.2 to 26.1 percent, about the same increase which took place in the Third World between 1950 and 1975, 16.7 to 28 percent [Preston (1979, p. 196)]. While the rate of urbanization in the Third World has not been exceptional, the rate of city growth has. Between 1875 and 1900, city populations in the currently developed countries rose by about 100 percent;
between 1950 and 1975, city populations in the Third World increased by 188 percent. City growth was faster in the Third World in part simply because overall rates of population growth were faster, in both city and countryside. Given the far higher rates of overall population growth in the Third World compared with nineteenth-century developing countries, it should come as no surprise that immigration accounts for a smaller share of Third World city growth. On average, immigration accounted for 39.3 percent of city growth in the 1960s [Preston (1979, p. 198)], while, as we shall see below, it was far higher during the First Industrial Revolution in England.

Table 11.2 summarizes English experience with the urban transition. Consistent with Preston's figures in Table 11.1, Todaro (1984, p. 13) estimates the rate of city growth in the Third World in the 1960s and 1970s to have been 4.32 percent per annum. While these city growth rates are double the English rates in Table 11.2, we must remember that everything was growing more rapidly in the Third World. Indeed, the overall population growth in the Third World [2.33 percent per annum: Todaro (1984, pp. 10-11)] was almost double that of England over the century 1776–1871 [1.23 percent per annum: Wrigley and Schofield (1981, p. 529)]. In any case, the English city immigration rates were not so much lower in the three decades 1776–1806, ranging between 1.10 and 1.91 percent per annum, than they were in the Third World, averaging 1.79 percent per annum. Thus, while England's growth was more gradual, recorded city immigration rates were on par with the so-called "exploding" Third World.

Now consider rural emigration. With the exception of the war-induced good times for English agriculture between 1801 and 1806, rural emigration took place at every point over the century. Furthermore, the rate of emigration about

---

**Table 11.1**

City growth rates in the Third World, 1950–2000 (percent per annum)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>City growth rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>4.42</td>
<td>4.85</td>
<td>4.97</td>
<td>5.10</td>
<td>5.00</td>
<td>4.56</td>
</tr>
<tr>
<td>Latin America</td>
<td>4.57</td>
<td>4.21</td>
<td>4.01</td>
<td>3.86</td>
<td>3.56</td>
<td>3.06</td>
</tr>
<tr>
<td>East Asia (exc. China and Japan)</td>
<td>4.16</td>
<td>5.20</td>
<td>4.52</td>
<td>4.00</td>
<td>3.33</td>
<td>2.36</td>
</tr>
<tr>
<td>South Asia</td>
<td>3.37</td>
<td>3.91</td>
<td>4.01</td>
<td>4.33</td>
<td>4.47</td>
<td>4.27</td>
</tr>
<tr>
<td>City minus rural growth rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>2.65</td>
<td>2.90</td>
<td>3.08</td>
<td>3.20</td>
<td>3.14</td>
<td>2.86</td>
</tr>
<tr>
<td>Latin America</td>
<td>3.34</td>
<td>3.19</td>
<td>3.19</td>
<td>3.03</td>
<td>2.74</td>
<td>2.29</td>
</tr>
<tr>
<td>East Asia (exc. China and Japan)</td>
<td>3.53</td>
<td>4.60</td>
<td>4.78</td>
<td>4.41</td>
<td>3.73</td>
<td>2.65</td>
</tr>
<tr>
<td>South Asia</td>
<td>1.55</td>
<td>1.72</td>
<td>1.88</td>
<td>2.18</td>
<td>2.65</td>
<td>3.20</td>
</tr>
</tbody>
</table>

doubled over the century. From 1816 to 1871, the rural emigration rate from the English countryside ranged from 0.87 to 2.10 percent per annum. The rates estimated for the Third World in the 1960s and the 1970s range between 0.97 and 1.37 [Preston (1979, p. 197), Kelley and Williamson (1984a, p. 93)]. In summary, judged by the standards of the First Industrial Revolution, the urban transition associated with ongoing industrial revolution in the Third World hardly seems exceptional.

2.2. Migrant selectivity bias

Cities are dynamic, but how much of that dynamism is attributable specifically to urban behavior, and how much to the fact that cities are simply full of young adults? Are cities different largely because of selectivity bias, young adults favoring the cities and shunning the countryside? Table 11.3 summarizes the evidence for both the Third World in 1970 and England in 1861. The rural–urban differences are significant although not spectacular; in both cases the cities tend to be full of young adults. The source of the rural–urban age differences is also

<table>
<thead>
<tr>
<th>Years</th>
<th>Urban share (%)</th>
<th>Annual percentage rate</th>
<th>Percent of city growth due to immigration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1776–1781</td>
<td>25.9</td>
<td>2.08</td>
<td>1.26</td>
</tr>
<tr>
<td>1781–1786</td>
<td>27.5</td>
<td>1.81</td>
<td>1.62</td>
</tr>
<tr>
<td>1786–1791</td>
<td>29.1</td>
<td>2.20</td>
<td>1.37</td>
</tr>
<tr>
<td>1791–1796</td>
<td>30.6</td>
<td>2.17</td>
<td>1.20</td>
</tr>
<tr>
<td>1796–1801</td>
<td>32.2</td>
<td>2.08</td>
<td>1.10</td>
</tr>
<tr>
<td>1801–1806</td>
<td>33.8</td>
<td>2.15</td>
<td>1.91</td>
</tr>
<tr>
<td>1806–1811</td>
<td>35.2</td>
<td>2.07</td>
<td>0.59</td>
</tr>
<tr>
<td>1811–1816</td>
<td>36.6</td>
<td>2.40</td>
<td>1.37</td>
</tr>
<tr>
<td>1816–1821</td>
<td>38.3</td>
<td>2.39</td>
<td>1.06</td>
</tr>
<tr>
<td>1821–1826</td>
<td>40.0</td>
<td>2.61</td>
<td>1.12</td>
</tr>
<tr>
<td>1826–1831</td>
<td>42.2</td>
<td>2.33</td>
<td>1.06</td>
</tr>
<tr>
<td>1831–1836</td>
<td>44.3</td>
<td>2.08</td>
<td>1.04</td>
</tr>
<tr>
<td>1836–1841</td>
<td>46.3</td>
<td>2.04</td>
<td>0.83</td>
</tr>
<tr>
<td>1841–1846</td>
<td>48.3</td>
<td>2.41</td>
<td>1.23</td>
</tr>
<tr>
<td>1846–1851</td>
<td>51.2</td>
<td>2.05</td>
<td>0.97</td>
</tr>
<tr>
<td>1851–1856</td>
<td>54.0</td>
<td>2.06</td>
<td>0.77</td>
</tr>
<tr>
<td>1856–1861</td>
<td>56.4</td>
<td>2.08</td>
<td>0.60</td>
</tr>
<tr>
<td>1861–1866</td>
<td>58.7</td>
<td>2.35</td>
<td>1.06</td>
</tr>
<tr>
<td>1866–1871</td>
<td>62.0</td>
<td>2.29</td>
<td>1.15</td>
</tr>
</tbody>
</table>

*Source: Williamson (1985a, Tables 4 and 5).*
Table 11.3
Age distributions in the contemporary Third World (1970) and in England (1861): Rural versus urban (%)

<table>
<thead>
<tr>
<th>Age class</th>
<th>Third World</th>
<th>England</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Urban</td>
<td>Rural</td>
</tr>
<tr>
<td>0-4</td>
<td>14.4</td>
<td>16.5</td>
</tr>
<tr>
<td>5-14</td>
<td>24.3</td>
<td>26.4</td>
</tr>
<tr>
<td>15-19</td>
<td>11.1</td>
<td>10.1</td>
</tr>
<tr>
<td>20-29</td>
<td>17.0</td>
<td>14.5</td>
</tr>
<tr>
<td>30-39</td>
<td>12.4</td>
<td>11.5</td>
</tr>
<tr>
<td>40-49</td>
<td>9.0</td>
<td>8.7</td>
</tr>
<tr>
<td>50-59</td>
<td>6.0</td>
<td>6.2</td>
</tr>
<tr>
<td>60+</td>
<td>5.8</td>
<td>6.0</td>
</tr>
<tr>
<td>Dependency rate</td>
<td>42.4</td>
<td>46.7</td>
</tr>
</tbody>
</table>


clear: migration tends to select young adults. Indeed, Rogers (1984) has constructed a “basic standard schedule” – reproduced in Figure 11.1 – which clearly reveals the clustering of migrants ranging from their late teens to their early thirties. Figure 11.1 also shows that immigrants to England’s cities in the 1850s had an even more pronounced age-selectivity bias than is the case now.

The economic implications of the young adult selectivity bias appears to be straightforward, although there has been little research devoted to an assessment of their quantitative importance. The higher urban activity rates would imply higher per capita incomes. The lower dependency rates would imply lower relief burdens in the city. Lower dependency rates also imply higher saving rates in the cities [Leff (1969), Coale and Hoover (1958)]. In addition, the young adult bias implies that the cities, ceteris paribus, should have higher crude birth rates, lower crude death rates, and thus eventually higher rates of growth in the natural labor supply as the urban transition unfolds. The moral of this demographic tale is that cities should increasingly find it easier to satisfy their growing labor requirements without the need for additional immigrants. And we do indeed find that over time the share of immigration in total urban population increase declines [Rogers (1984), Keyfitz (1980)].

2.3. Selectivity bias, the brain drain, and remittances

Concern with the rural “brain drain” is at least as old as the First Industrial Revolution in England [Graham (1892)]. Even taking into account only the rural rearing costs of children, the rural–urban transfer embodied in England’s rural
emigration in 1850 cost the rural sector the equivalent of 56 percent of its investment in that year [Williamson (1985b)]. While I am unaware of any similar data for Third World countries, if they are even close to the estimated value for the English case, then the brain drain certainly does take on an important dimension. The above estimate is, if anything, a lower bound, due to the presence of the migrant selectivity bias: rural emigrants are young adults whose present value is positive and large; and rural emigrants may also be drawn from the best and the brightest.

Unfortunately, we have little evidence on the economic importance of the brain drain. Equally important, we have very little evidence about rural returns on that investment flow, although the recent literature on the rural consequences of outmigration contains some evidence on urban–rural remittance flows [Johnson and Whitelaw (1974), Knowles and Anker (1981), Lucas and Stark (1985)]. Rempel and Lobdell (1978) undertook a comprehensive survey of this literature almost a decade ago and, based on some fifty studies, concluded the
following: (i) urban–rural remittances are very large in Africa, ranging from 10 to 13 percent of urban incomes; (ii) while there is some limited evidence supporting similar conclusions for Asia [Papola (1981, pp. 84 and 93)], there is almost none for Latin America; and (iii), remittances associated with international migration tend to be far higher. As Rempel and Lobdell point out, however, none of these studies controls for the funds emigrants bring with them, especially for those youths going to urban areas for additional education. When these emigrant funds are subtracted from immigrant remittances, the net remittances turn out to be small even for Africa, although this conclusion is disputed by Stark (1980a) and Lucas and Stark (1985).

The stakes in the debate on urban–rural remittances seem to be large. If remittances are indeed small, then rural emigration has important welfare and distributional implications. Furthermore, low remittances would imply high child “default rates” on rural parental investment, thus suggesting a force which might serve to lower fertility in the countryside. [See Caldwell (1982), Cain (1981), Williamson (1985c), Sundstrom and David (1986), Ransom and Sutch (1986). In contrast, see Stark (1981), Katz and Stark (1985).]

2.4. City growth, migration, and labor absorption

The remainder of this chapter will deal with the labor market forces which condition the stylized demographic facts presented above. Throughout, we need to view the migration and urbanization process in terms of a system of five equations: (1) rural and (2) urban labor supply conditions, (3) rural and (4) urban labor demand conditions, and (5) a migrant-clearing equation. While the literature tends to dwell most heavily on the last two of these – urban labor absorption and migration – the problem should be posed in general equilibrium. We shall try to do so wherever possible.

3. Disequilibrating labor market shocks and equilibrating migrant responses

3.1. Disequilibrium and wage gaps

The simplest multi-sectoral models assume long-run wage equalization. But even in such simple models, unbalanced output growth and derived labor demand can imply labor market disequilibrium and wage gaps in the short run. Migration, after all, responds only with a lag to wage differentials. Furthermore, when such models are consistently shocked by disequilibrating economic and demographic events, wage gaps can persist and even increase. Since such shocks are precisely the ingredients of industrial revolutions, development economists have come to expect the appearance of wage gaps in Third World countries, especially between rural agriculture and urban manufacturing.
Figure 11.2 describes the familiar two-sector example of labor market disequilibrium with wage gaps between agriculture (A) and manufacturing (M). The derived labor demand functions, \( AA' \) and \( MM' \), intersect at \( Z \) where the distribution of labor is optimal, and the prevailing wage is \( w^* \). In the presence of a wage gap like \( XY \), however, the distribution of labor is sub-optimal: here there are too many farm laborers and not enough manufacturing operatives. I have also shaded in the Harberger Triangle \( XYZ \), representing the allocative losses associated with this labor market failure. The economic significance of the market failure is conditioned by two forces, the size of the wage gap and the share of the labor force misallocated, \( \ell - \ell^* \).

How big are rural–urban wage gaps? Lewis (1954) thought the gap was about 30 percent when he was writing in the early 1950s, but a recent review suggests they are about 41 percent [Squire (1981, Table 30, p. 102)]. Real wage gaps for relatively homogeneous unskilled male labor were somewhat smaller in England at the end of the First Industrial Revolution, about 33 percent [Williamson (1986)].

There has always been considerable disagreement about the magnitude of these gaps. As Kannappan (1985, pp. 708–711) points out, the measurement problems are often overwhelming [see also Knight (1972), Taira (1973)]. Indeed, Behrman and Birdsall (1983) have argued that the addition of differences in schooling quality to the more conventional list of explanatory variables causes urban–rural earnings differentials to evaporate for Brazilian males.

There has also been considerable debate over trends in the real wage gap, in part motivated by the fact that rising wage gaps are a central premise of the Todaro model. As it turns out, there is little evidence confirming a rise in the
wage gap in the Third World [Kannappan (1985, pp. 712–713), Gregory (1975), Glytsos (1977)], but there is abundant evidence coming from the First Industrial Revolution. It appears that the nominal wage gap in England rose from about 16 percent in 1815 to a peak of 64 percent in 1851 [Williamson (1985d, p. 49)]. All of this evidence on wage gap trends, however, should be treated with caution since nominal, not real, wages are almost always employed in the analysis.

Based on evidence from Brazil [Fishlow (1972), Thomas (1982)], Peru [Thomas (1980)], the American South [Bellante (1979)], and nineteenth-century England [Williamson (1986)], I suspect nominal wage gaps are poor proxies for real wage differentials. Not only do rural areas have a lower cost-of-living, but it appears to be the non-food component that explains most of the differentials [Thomas (1980, p. 89), Williamson (1986)]. On the other hand, urban amenities associated with government services may well serve as a partial offset to the higher cost-of-living in the cities, since these amenities may be capitalized into rents.

Given an initial wage gap, what would the national income gains be if migration to the cities drove the wage gap to zero? This calculation has in fact been made for England in the midst of the First Industrial Revolution [Williamson (1986)], where the deadweight loss embodied in that Harberger Triangle turns out to have been only 0.5 percent of GNP. As is so often the case in such calculations [Leibenstein (1957), Dougherty and Selowsky (1973), Shoven and Whalley (1984)], the figure is trivial. One would have thought it would have been bigger given that an outmigration of about 20 percent of the labor force would have been necessary to achieve wage equalization. The 0.5 percent estimate may seem small for another reason. Figure 11.2 and the calculation for England assume that the derived labor demands in the two sectors remain stable in the face of massive outmigration to the cities. It seems far more likely that $AA'$ would shift upwards while $MM'$ would shift downwards in response to the migration since the ensuing agricultural output contraction and the urban output expansion would serve to improve agriculture’s terms of trade (depending, of course, on the price elasticities underlying the demands for $A$ and $M$). On the other hand, this “partial” general equilibrium calculation does not allow capital to chase after labor. When a full general equilibrium assessment is made of the British case, the deadweight loss is considerably larger, some 3.3 percent of GNP.

3.2. Are migrants rational?

Wage gaps imply some labor market disequilibrium, and the migration necessary to eliminate that disequilibrium is quite large in typical developing countries, like England in the first half of the nineteenth century. Do migrants in fact respond to wage gaps?

Since Sjaastad’s (1962) seminal work, an enormous amount of empirical evidence has accumulated on the determinants of individual migration. Most
studies confirm that migrants are motivated by earnings differentials, although it is hard to imagine that we would find otherwise. By far the best survey of this research can be found in Yap (1977) whose summary of the available data shows quite clearly that migrants are rational and do indeed respond to earnings differentials. For example, a recent study of Ahmedabad [Papola (1981, pp. 82–83)] reports informal sector workers at least doubled their earnings upon immigration to the city. Yet, and as Todaro insists, the chance of obtaining employment matters, and the expected wage often (but not always) yields better results in regression analysis than the wage alone. But the critical issue is not so much whether migrants respond to earnings differentials, but whether the rate of return to migration is “high” (implying insufficient migration), and whether the elasticity of migrants to wage differentials is “large” (implying low costs of migration). On this score, the literature is much too quiet.

As Yap (1977, pp. 253–254) points out, even though migrants improve their incomes by city immigration, they may have been limited to marginal employment. The evidence, such as it is, does not appear to support this view. While we shall have more to say about the issue below in Section 5 where the Todaro thesis is confronted, “income comparisons between migrants and natives suggest that the incomes of urban dwellers are more a function of education and skill levels than of migration status” [Yap (1977, p. 254)]. The same was true of urban immigrants in England’s cities in 1851 [Williamson (1985e)].

As noted above, Thomas (1980) has shown for Peru that the non-food component of the cost-of-living index explains most of the cost-of-living differential between city and countryside. To the extent that rents are the central ingredient of the cost-of-living differential, and they seem to have been in England in the 1830s [Williamson (1986)], that fact implies, ceteris paribus, that the nominal wage gap should increase over time as land scarcity in the city rises with urbanization. Furthermore, it suggests that government policy towards urban housing should be critical to the migration decision and to the cost of labor facing firms in the city. In addition, the availability and cost of public services should also matter to the migration decision, suggesting that government policy regarding urban social overhead (including schooling and health) should be an important determinant of city immigration rates [Linn (1983)]. To quote Yap (1977, p. 245):

When the destinations are cities, better living conditions, better educational opportunities…may…be important motivating factors. Since anti-poverty programs are apt to include improvements in urban services, it is useful to know whether programs to upgrade housing, sanitation, health, and public education facilities will also stimulate in-migration….

Unfortunately, econometric studies have only been able to confirm the general attraction of cities. They have not been able thus far to isolate the impact of
non-wage components of the migration decision, primarily because effective measures of urban services and quality-adjusted rents are hard to get. This surely should be a key goal of future research on Third World migration.

3.3. Are there too many city immigrants?

Given wage gaps favoring Third World cities, it would appear that the rate of outmigration from the countryside is too low. In the face of such evidence, how could anyone argue that there are too many city immigrants, that cities are too large and too many? The only way that both statements can be correct is, of course, if the social cost of the migrants exceeds their private cost, in which case Third World cities would represent another example of the “tragedy of the commons”.

There are only two ways that the too-many-city-immigrants argument can be accommodated. First, it may be argued that the marginal cost of urban public goods is far above the user price charged, and that the subsidy fails to get capitalized into rents. While this argument certainly sounds plausible, as far as I know there has been no attempt to assess its quantitative relevance. Furthermore, if the problem is underpricing urban public services, then the solution is not to close the cities down, but rather to get the prices right. Second, it may be argued that immigration generates important negative externalities. One such negative externality might be that the relative poverty of the urban migrants tends to create social and political tensions. Of course, it could be argued just as plausibly that the presence of external agglomeration economies implies that cities are too small and that the observed level of urban immigration is too little [Shukla and Stark (1985)]. In any case, it is never quite clear whose objective function is being considered in the too-many-city-immigrants argument. One can appreciate the planners' dilemma of dealing with what appears to be an overwhelming rush to the cities, or their complaints that the quality of urban services newly created rapidly deteriorates as they get stretched across an ever increasing set of users, but what about the urban and, indeed, the national residents who the planning bureaucracy serve?

Oddly enough, there is almost no attention in the development literature to the issue of who gains and who loses from an influx of unskilled labor to the city. The relative silence in this literature on the distributional impact of urban immigration is especially odd given the attention which economists have lavished on the distributional impact of international migrations.

In the simplest general equilibrium model, who gains and who loses from city immigration of unskilled labor seems clear enough. In the countryside, unskilled labor that stays behind should gain, to the extent that increased labor scarcity tends to raise their earnings [although the empirical evidence is mixed: Johnson
(1960), Gardner (1974), Greenwood (1978), and Schuh (1982)], and land rents should fall (in the absence of offsetting terms of trade effects). In the cities the immigrants gain, as we have seen. Unskilled city-born laborers and previous immigrants who are crowded-out by the competition from the new immigrants surely lose as the urban labor market becomes glutted with more unskilled labor. In contrast, skilled labor will gain as the increase in unskilled labor supplies augments output and drives up the demand for skills. Certainly capitalists gain for identical cheap labor reasons. And finally, to the extent that middle class and rich tend to consume heavily the personal services offered by the urban unskilled, they get an additional cost-of-living gain. Thus, while planners may push for restrictions on immigration—making their job of coping with city growth easier—one can hardly imagine any other class of city residents pushing for restrictions, especially given that the previous unskilled immigrants have very little political power. It is hardly surprising, therefore, that immigration restrictions have hardly been tried, and have often failed when tried.

A reduction of city immigration was one of the first policy options suggested for reducing urban poverty and unemployment in the Third World [Frank (1968), Yap (1976, p. 227)]. The policy makes some sense if the focus is solely on the welfare burden of the urban poor. However, the policy makes no sense at all if the focus is on national poverty and welfare burdens. After all, outmigration to the city is one of the best ways to escape rural poverty. Yet, it is quite apparent that some development analysts and many governments think there are too many city immigrants, and that policies should be devised to reduce their number. Preston's summary [Preston (1979, p. 195), U.N. (1980, p. 3)] indicates the popularity of this view:

Among the 116 developing countries that responded to the United Nations' "Fourth Population Inquiry Among Governments" conducted in 1978, only six declared the spatial distribution of their population to be "acceptable". Forty-two replied that it was "unacceptable to some extent", and 68 declared it to be "highly unacceptable". To another question addressed specifically to the desirability of current rates of rural-urban migration, only three countries expressed a desire to accelerate such migration. Twenty-three wished to maintain it at present levels, 76 to slow it down, and 14 to reverse it.

As Shukla and Stark (1985, p. 297) point out, things had not changed when a similar U.N. survey was taken in 1983.

Simmons (1979) documents Asian experience with slowing down metropolitan city growth, and Shaw (1978) focuses on housing as one key policy instrument to effect the same goal [see also Laquian (1981)]. In spite of such policies, however, governments have generally been ineffective in substantially reducing city immigration in the Third World. Their failure seems to stem from two sources: the evidence in support of the too-many-city-immigrants view is much too weak to
motivate policy-makers to do more; and the urban political forces that matter come from classes who clearly gain from city immigration.

4. What does “overurbanization” mean?

Much of the previous section dealt with what city immigration ought to be, but a parallel debate exists over the forces which have actually driven migration and city growth in the Third World. Part of this debate deals with “push” versus “pull” forces, and part of it deals with what has come to be called the “urban bias”. This debate was important in motivating the Todaro model and what might be called the new urban pessimism.

4.1. Push, pull, and the engines of city growth

One of the central arguments used to support the view of overurbanization is that Third World populations are supported by an unusually small industrial labor force [Preston (1979, p. 207)]. The evidence on this point was first marshalled by Hoselitz (1953, 1955, 1957) who found that Third World ratios of industrial employment to urban population were small in contemporary developing nations (Asia in particular) compared with currently developed countries in the late nineteenth and early twentieth centuries. Not too long ago, Bairoch (1975, p. 150) replicated Hoselitz’s finding:

The degree of urbanization in under-developed countries was the same in 1960 as it had been in Europe in 1880–85 when the latter’s percentage of working population engaged in manufacturing was twice that of the less-developed countries.

Using World Bank data from the 1960s, Berry and Sabot (1984, p. 106) have made the same point:

Whereas at the turn of the century the industrializing European nations absorbed almost half of their incremental labor force into industry each year, today the developing countries absorb less than 30% of their additional workers into industry.

However, given the more sophisticated level of social services now offered in major cities everywhere, and given the far more rapid rate of city population growth augmenting employment in the formal and informal construction sector, it is not at all clear that contemporary Third World cities should replicate the employment patterns which prevailed almost a century ago.
Nonetheless, Hoselitz's thesis has encouraged the view that Third World city growth is being driven by somewhat different forces than has been true historically. According to this view, sometime after 1930, "urbanization without industrialization" began to appear in the Third World [Bairoch (1975, p. 144)], and the problem took on alarmist dimensions in the 1960s when urban unemployment and underemployment captured our attention.

Since so much seems to ride on the issue, fresh new tests of the Hoselitz thesis have gained high research priority with the appearance of more employment data during the last two decades of Third World city growth. Based on an exhaustive assessment of industrial employment trends in the Third World between 1950 and 1970, Preston has been unable to find any evidence of deterioration in industry/urban ratios for the Third World as a whole [Preston (1979, pp. 207–209), U.N. (1980, pp. 17–19)]. Indeed, in the largest of the developing regions—Middle South Asia (including India)—the industry/urban ratio has risen. Latin America, on the other hand, seems to be an exception. Furthermore, where the urban service sector has been rising, "it is typically a result of rising fractions of professional, technical, and administrative personnel" [Preston (1979, p. 209)]. A more recent study by Gregory (1980) has offered an even sounder rejection of the Hoselitz thesis: employment in the tertiary sector has not grown at an explosive rate in the Third World, but rather has matched that of the secondary sector (p. 682); employment in sales and services has in fact grown at a slower rate than employment in production (p. 686); furthermore, there is no evidence to support the view that informal sector urban employment has grown at an exceptionally rapid rate (p. 696). Generally, urbanization does not seem to have outpaced industrialization in the Third World since 1950, although it is still true that cities in the developing world have lower industry/urban ratios than was the case at the turn of the century for currently developed countries.

While urbanization has not outpaced industrialization in the Third World, the higher service/urban ratios have still suggested to some that there are over-urbanization forces at work. Hoselitz's original findings encouraged the view that the urban service sector must be too large. Rather than viewing the growth of urban services as a response to buoyant demand, Hoselitz viewed that growth mainly as a response to immigrant-swollen labor supplies. He believed that migrants were pushed to the city by unfavorable employment conditions in the countryside, that glutted labor markets in the city could not be cleared by employment in high-wage modern sectors, and thus that city service sectors were simply offering employment of the last resort. Low-wage employment in the "unproductive" service sector boomed, thus assuring the finding of relatively low industry/urban ratios in the Third World by historical standards.

If Hoselitz is right, then urban service sectors—the informal sector in particular—should exhibit relatively low wages and be dominated by migrants. If he is wrong, then we should find no such evidence. And if he is wrong, then we should
be able to offer explanations for buoyant labor demand in urban services. How much of that demand might stem from government final demand? How much from employment multipliers emanating from the manufacturing export base? How much from the construction of social overhead and housing, the latter including squatter housing? The literature has yet to supply an effective accounting of these potential sources of labor demand for urban services, and it matters to our understanding of the city growth process.

The literature does offer answers to the following questions: Do urban service sectors in the Third World exhibit relatively low-wage employment? Are they dominated by migrants? Apparently migrants do not dominate urban service employment [Yap (1977, p. 255)]. Nor can it be argued that wages are relatively low in urban service sectors, including the informal sector [Yap (1977), Papola (1981), Kannappan (1985)]. Indeed, Mazumdar (1976) and others have questioned the view that urban informal service sectors are a point of entry for the immigrant any more than they are for the city-born. Furthermore, Udall (1976) has shown that when rural disturbances led to a dramatic labor migration into Bogota, the influx failed to increase the relative size of the service sector, a result in complete contradiction with the Hoselitz thesis. While we shall have more to say about the issue below when we examine urban labor markets in greater detail, it appears that the evidence is not consistent with Third World "overurbanization".

4.2. The urban bias

Of course, the Third World may be overurbanized for another reason. Lipton (1976) and Keyfitz (1982, pp. 662–667) have both reminded us that an "urban bias" favors city growth [see also Mamalakis (1969), Mitra (1977), Bates (1981), Moore (1984)]. There is no shortage of policies which have that effect: the domestic terms of trade is twisted against agriculture, thus encouraging more rapid rural emigration to the city than would have been true in the absence of such policies; tariff and exchange-rate management also serve to protect urban industry, fostering its expansion at agriculture's expense (although the higher capital intensities thus encouraged may offset urban employment effects); financial markets are manipulated by government policy to create cheap capital for favored urban industries (although, once again, the higher capital intensities thus encouraged may offset urban employment effects); and social overhead is allocated disproportionately to the cities (schools, health facilities, roads, water supplies, and electricity), all offered at less than user cost, and all financed from general tax revenues rather than from urban land taxes. While it is easy enough to list the sources of urban bias [and its magnitude: e.g. Little, Scitovsky and Scott (1970), Bale and Lutz (1981), Agarwala (1983)], it is quite another matter to
establish the impact of this bias on Third World city growth. I know of no study which has performed this exercise, and it is central to policy debate. While Montgomery (1985) has explored the correlations between Agarwala's (1983) indices of price distortions and urbanization among 31 developing countries, he finds it difficult to control for other forces driving city growth. It seems to me that this issue deserves high research priority, and I suspect it will require a full computable general equilibrium model to attack it properly.

5. How do urban labor markets work?

5.1. The evolution of conventional wisdom

Development economists in the 1950s viewed rural–urban migration and city growth with optimism. City immigration was seen as a favorable process whereby "surplus rural labor was withdrawn from traditional agriculture to provide cheap manpower to fuel a growing modern industrial complex" [Todaro (1980, p. 361)]. Demand pull was central to such development models [Lewis (1954), Fei and Ranis (1961)], and in the absence of labor-saving technological change in urban industry, city job creation insured a steady decline in low-productivity employment nation-wide as per capita incomes rose. Accumulation was the key constraint on economic transformation in such models, but with an elastic immigration response to favorable employment conditions, real wages tended to remain stable, profit shares in GNP rose, the domestic saving rate increased, and the accumulation constraint was released.

Things began to change in the 1960s. The appearance of large-scale household surveys documented for the first time what appeared to be high rates of unemployment throughout the Third World. By the end of the decade, "the dominant impression was that rates of urban unemployment were increasing and that urban unemployment in LDCs was a chronic problem" [Berry and Sabot (1984, p. 105)]. Furthermore, many economists were persuaded that underemployment touched even greater numbers than did open unemployment – the ILO suggesting that there were three or four underemployed for every person unemployed [Berry and Sabot (1984, p. 114, fn. 20)], and an urban informal service sector was viewed as the main holding area of low-wage underemployed. Indeed, informal service sector employment in the cities began to be viewed as a "mere manifestation of an impoverished rural 'surplus'" [Kannappan (1985, p. 702)].

The optimism of the 1950s was slowly being replaced by a neo-Malthusian pessimism, and urbanization debates shifted to gloomy accounts of unemployment, poverty and inequality. Demographers and labor economists began to construct an "excess supply-limited demand" paradigm [Fry (1979), Kannappan (1985, p. 699)]. Not only did this gloomy neo-Malthusian perspective discount
the role of demand-pull in the cities, but it even viewed immigration to the city with alarm.

5.2. The Todaro model

This neo-Malthusian pessimism was given theoretical rationalization in 1969 with the appearance of Michael Todaro’s model of labor migration and urban unemployment. Not surprisingly, the Todaro model took the profession by storm. After all, it had very attractive ingredients:

The idea that migrants compare expected gains with the current costs of being unemployed represented rather standard theory. The model’s appeal lay rather in the fact that it fitted well with three prevalent stereotypes: high wages in the modern sector; presumptions of mass unemployment; permissive or overly generous policies and/or articulate, militant labor movements [Kannappan (1985, p. 703)].

The Todaro model and its extensions [Harris and Todaro (1970), Zarembka (1972), Stiglitz (1974), Corden and Findlay (1975), Cole and Sanders (1985)] has enjoyed considerable popularity since 1969. The hypothesis is simple and elegant. Perhaps the most effective illustration can be found in Corden and Findlay (1975), reproduced in Figure 11.3. There are only two sectors analyzed in the figure, and labor is the only mobile factor there, but it is sufficient to illustrate the point. Under the extreme assumption of wage equalization through migration, and in the absence of wage rigidities, equilibrium is achieved at E (the point of intersection of the two labor demand curves, $AA'$ and $MM'$). Here wages are equalized at $w_\star = w_{\star\star}$, the urbanization level is $O_M L_M / L$ (the share of the total labor force, $L$, employed in urban jobs, $O_M L_M$), where $M$ denotes urban manufacturing and $A$ denotes agriculture. Wages are never equalized in the real world, of course, and so the model incorporates the widely-held belief that the wage in Third World manufacturing is pegged at artificially high levels, say at $\bar{w}_M$. If for the moment we assume unemployment away, then all of those who fail to secure the favored jobs in manufacturing would accept low-wage jobs in agriculture at $w_{\star\star}$. The model now allows for a wage gap between the two sectors, a feature of developing economies explored at length in Section 3.

Figure 11.3 makes it clear that the level of city employment would be choked off by the high wage in manufacturing, but would city immigration also fall off? Not necessarily. Indeed, the model was originally motivated by concerns with urban unemployment, as well as by the coexistence of dramatic city growth,

---

1 The Todaro model is also strikingly similar to Mincer’s (1976) two-sector framework used to assess the impact of minimum wage legislation in “covered” and “uncovered” sectors.
unemployment, and the expansion of the informal urban service sector where, it was alleged, low-wage underemployment prevailed. Todaro explains this apparent conflict (e.g. immigration in the face of urban unemployment and underemployment at very low wages) by developing an expectations hypothesis which, in its simplest form, states that the favored jobs are allocated by lottery, that the potential risk-neutral migrant calculates the expected value of that lottery ticket, and then compares it with the certain employment in the rural sector. Migration then takes place until the expected urban wage is equated to the rural wage. Given \( \bar{w}_M \), and a wage in informal urban services so low that it can be taken as zero, at what rural wage would the migrant be indifferent between city and countryside? If the probability of getting the favored job is simply the ratio of employment in manufacturing, \( L_M \), to the total urban labor pool, \( L_U \), then the expression

\[
\bar{w}_A = \left( \frac{L_M}{L_U} \right) \bar{w}_M
\]

indicates the agricultural wage at which the potential migrant is indifferent about employment location. This is in fact the \( qq' \) curve in Figure 11.3. The equilibrium agricultural wage is given by \( w_A \), and those unemployed or underemployed in the city (e.g. the size of the informal service sector plus those without any work at all) is thus given by \( L_{US} \).²

²The \( qq' \) curve is a rectangular hyperbola with unitary elasticity. The elasticity of the labor demand curve \( MM' \) is assumed to be less than unity in Figure 11.3, an assumption motivated by empirical evidence available for the Third World.
The new equilibrium at Z in Figure 11.3 seems to offer an attractive explanation for some of the stylized facts of Third World labor markets. It yields a wage gap, \( \bar{w}_M - w_A \), and urban low-wage employment or unemployment, \( L_{US} \). Moreover, when the dynamic implications of the model are explored, it turns out that an increase in the rate of manufacturing job creation need not cause any diminution in the size of the low-wage informal service sector. Indeed,

as long as the urban–rural [wage gap] continues to rise sufficiently fast to offset any sustained increase in the rate of job creation, then...the lure of relatively high permanent incomes will continue to attract a steady stream of rural migrants into the ever more congested urban slums [Todaro (1969, p. 147)].

Nor has Todaro changed his view since 1969. A decade later he stated that city immigration in the Third World is

the major contributing factor to the ubiquitous phenomenon of urban surplus labor and as a force that continues to exacerbate already serious urban unemployment problems caused by growing economic and structural imbalances between urban and rural areas [Todaro, (1980, p. 362)].

Furthermore, the model makes some firm assertions about how urban labor markets work and how immigrants are absorbed into that labor market. First, it asserts that immigrants earn lower incomes than non-immigrants, the latter having first claim to the favored jobs. Second, it asserts that immigrants have a higher incidence of unemployment. Third, it implies that wages are lower in informal service sector employment than in industrial employment. Fourth, it implies that immigrants earn less in the cities when they first arrive, than they earned in the rural areas they left.

5.3. Critique: How do urban labor markets really work?

There are five critical assumptions of the Todaro model which lead to its dramatic results. Each of these seems restrictive, or at the very least deserve far more research.

First, as Willis (1980, p. 396) points out, job allocation rules are not likely to obey the simple lottery mechanism embedded in the Todaro model. The literature on job search has grown considerably since 1969, and it all emphasizes the role of investment in the search. In contrast, there is no such explicit investment in the Todaro model, except for the actual decision to migrate. In fact, recent evidence from the Third World suggests that unskilled immigrants do not engage in long job searches, and that overt urban unemployment is an attribute of the skilled
rather than the unskilled [Yap (1976, 1977), Papola (1981)]. Indeed, this fact has encouraged the development of two-stream migration models and explicit attention to labor heterogeneity [Fields (1975), Fallon (1983), Cole and Sanders (1985)].

Second, there is no attention to informal sector labor market behavior in the Todaro model. In particular, we need to know far more about the sources of labor demand. After all, wages do clear that labor market, they are responsive to demand and supply, and they certainly do not settle to zero.

Third, there has been little evidence marshalled in support of the modern sector rigid-wage assumption [Montgomery (1985)]. This statement holds true for trade union pressure and minimum wage legislation. After all, wage differentials between urban formal and informal sectors could be explained just as well by appealing to firm-specific training costs [Mazumdar (1973, 1976), Stiglitz (1974)].

Fourth, there is the issue of discount rates and rational migrants. What matters to Todaro’s migrants is the present value of expected urban earnings compared with the present value of expected rural earnings. Given modern sector wages double those of rural wages, given some unemployment duration before a migrant secures the modern sector job, and given some discount rate, how long a time horizon would a potential migrant have to have before present values were equated? Cole and Sanders (1985, p. 485) have made that calculation where discount rates are allowed to vary between 5 and 15 percent. They conclude: “If one must assume very long time horizons, in some cases greater than 50 years, an alternative explanation of migration may be in order” [Cole and Sanders (1985, p. 485)].

Fifth, and perhaps most important, the model abstracts from many additional influences on the potential migrant’s decision. This is the thrust of much of Stark’s recent work on risk aversion [Stark and Levhari (1982), Katz and Stark (1986)], relative deprivation [Stark (1984)], and cooperative family games.

This debate is not merely academic nit-picking, since conflicting policy morals may emerge from the Todaro model and an alternative model which relaxes these critical assumptions. One of the first morals likely to be reversed is that “underemployment” in the informal service sector is socially unproductive. A second moral likely to be reversed is that rapid job creation in the modern sector fosters increasing urban unemployment. However, an important third moral is likely to remain unchanged: namely, development strategies should continue their recent emphasis on rural growth [Fields (1980, p. 390)].

5.4. Some evidence

To assess the empirical relevance of the Todaro model, we need answers to five key questions. First, are earnings in the urban informal sector less than those in
the modern sector? Indeed, are they less than those in rural areas? Second, do city migrants suffer relatively protracted unemployment during job searches? Third, do migrants earn less than non-migrants? Fourth, do “new” immigrants to the city earn less than their rural counterparts? Fifth, are migrants slow to assimilate into city labor markets?

Are earnings in the urban informal sector less than those in the modern sector? To begin with, labor force data do not support the proposition that urban economies can be cleanly separated into “capitalistic” and “traditional” sectors along dualistic lines [Kannappan (1985, pp. 705–708)]. Nor does the evidence support the view that wages in the modern sector are well above those in the informal sector, especially for manual workers [Kannappan (1985, pp. 708–712)]. Nor is there much evidence to support the view that manufacturing wages have outpaced agricultural wages in the Third World [Gregory (1975), Glytsos (1977), Bose (1978)]. The available data do not, therefore, support Todaro’s premise of rising wage gaps, a necessary premise to get his result that unemployment increases over time even in the face of modern sector job creation.

Is there high and rising urban unemployment in the Third World? Apparently not, at least based on 14 countries with adequate data across the 1960s and 1970s [Gregory (1980, p. 697), Berry and Sabot (1984, p. 109 and p. 115, fn. 32)]. Most of the early assertions which reached the opposite conclusion were based on poor data, a warning raised long ago by Turnham (1971) and more recently by Kannappan (1985, pp. 718–719). And who’s unemployed? Is it the urban immigrant? Not in Ahmedabad, where unemployment rates among the migrants were low and labor participation rates high [Papola (1981, p. 99)]. Nor in Colombia, Tanzania, and Malaysia where:

the pool of unemployed comprises predominantly non-heads of household, the young, and married women who typically do not bear main economic responsibility for others and who have access to transfer payments from parents or husbands. They can and frequently do remain unemployed for long periods… Most of the unemployed are relatively educated… [Berry and Sabot (1984, p. 111)].

In fact, the duration of job search seems to be low for city immigrants, especially among the unskilled. Based on a World Bank survey completed in 1975 [Sinclair (1978, pp. 50–51)], two-thirds of migrant job-seekers had found a job after one month, and a significant share of these had arranged for employment prior to the move. In Ahmedabad, the duration of job search was 1.5 months for the average migrant, and 53 percent of these were “sponsored” in some form or another [Papola (1981, pp. 83–84)]. Similar findings of short duration of job search are also reported in Yap’s survey [Yap (1977, p. 251)]. Todaro’s job-lottery and high-unemployment view of urban labor markets in the Third World simply fails to pass the test of evidence.
Do migrants earn less than non-migrants? Perhaps before we try to answer that question, we might first determine whether migrants dominate the informal service sector in Third World cities. The evidence appears to be mixed. While Papola (1981) can find no evidence to support the Todaro prediction in Ahmedabad, Yap (1976, p. 238) has found the opposite for Brazil, and others have found the same elsewhere in Latin America [Yap (1977, p. 255)]. The explanation, however, has less to do with their migrant status than with their age and skills [Yap (1976, p. 238)]. This emphasis on the heterogeneity of the immigrants to Third World cities has, in fact, motivated the recent development of models of two-stream migration. In any case, the evidence does not support the view that migrants earn less than non-migrants [Yap (1977), Nelson (1979), Mohan (1980)], leading Yap to her important conclusion that:

Income levels... seem to be more a function of an individual's human capital endowments than of his migration status. Accordingly, strategies to alleviate poverty should place more emphasis on raising the skill levels of the urban population than on restricting migration to the cities [Yap (1976, p. 241)].

Are the earnings of "new" city immigrants less than their rural counterparts? Are immigrants slow to assimilate into city labor markets? The answer to both questions is no. Yap (1977) and others have shown conclusively that city immigrants improve their income over their rural options immediately upon finding an urban job. Furthermore, the rate of assimilation and income improvement is very fast [Yap (1977, pp. 253–256)].

Most economic historians would have been surprised that the debate over Third World city immigrants' economic progress lasted so long, since the historical literature seems to be consistent with the view that migrants improved their earnings with the move, that they assimilated quickly into city labor markets, that duration of job-search was short and unemployment low, and that immigrants and native-born had similar earnings experience (after controlling for age, education and other components of human capital). This is certainly what the literature on nineteenth-century foreign migration into American cities suggests [Higgs (1971), Chiswick (1979), Hannon (1982), Williamson (1982b)]. It is also what the evidence from British cities in 1851 suggests [Williamson (1985e)]. It is hardly surprising, therefore, that the Third World evidence suggests the same.

In short, while there is some evidence which confirms the role of expectation in migration behavior [Todaro (1980, pp. 380–382)], other evidence refutes their importance [cited in Katz and Stark (1986, p. 134)]. In any case, the evidence appears to reject all the remaining premises in the Todaro model on how urban labor markets work.
6. Migration and city growth in general equilibrium: What are the driving forces?³

6.1. What drives Third World migration and city growth?

This question brings us full circle back to the issues raised at the beginning of this chapter. Regardless of the model of migration preferred and one’s view as to how urban labor markets work, the central question lurking behind all Third World urban debates is the quantitative importance of the underlying forces pushing and pulling migrants to the city. Oddly enough, while the debate has been intense, until recently the literature has provided little guidance as to which of the contending forces is doing most of the work.

As this chapter has already implied, there is certainly no shortage of hypotheses as to what drives Third World migration and city growth. Each of these hypotheses can be classified under one of three headings: first, the endogenous “limits” to city growth; second, exogenous external events; and third, exogenous internal events.

Based on the firmly-held belief that the current structure of an economy—including urbanization levels—can influence subsequent economic performance, macro models of Third World countries have stressed sectoral detail from the start. The classic examples are offered by the dual economy models pioneered by Lewis (1954), Fei and Ranis (1964), and Jorgenson (1961), the latter extended by Kelley, Williamson and Cheetham (1972). Central to these models and their more elaborate extensions are the output gains associated with resource transfers from traditional low-productivity sectors to modern high-productivity sectors. Such resource transfers—labor migration in particular—have obvious spatial implications, the most notable example being urbanization.

In the classic labor surplus version, urbanization augments aggregate output both through short-run efficiency gains and long-run growth effects. In the short run, labor is shifted from low to high marginal productivity employment. In the long run, accumulation rates are raised since saving rates are higher in the modern urban sectors—indeed, in the extreme version only capitalists save, and capital is an argument in the modern, urban-based production functions only. Hence, city immigration leads to higher savings, investment, and output growth [Stark (1980b, p. 97)]. Rising urban accumulation rates imply increased rates of modern-sector job vacancies, a rural–urban migration response, and further urbanization. Thus, output growth, trend acceleration, and increased urbaniza-

³This section draws heavily on previous collaborative work in Kelley and Williamson (1984a, especially pp. 7–10), Kelley and Williamson (1984b), Becker, Mills and Williamson (1986), and Becker, Williamson and Mills (in progress).
tion are the likely outcomes of the labor surplus model. The neoclassical dual economy model makes the same prediction, at least in the long run.

What forces tend to inhibit the rate of urbanization in these dual economy models? Obviously the share in urban areas cannot exceed 100 percent. Thus, in the very long run, urbanization rates must slow down as this limit is approached and city growth rates decline to the national population growth rate. In the medium term, however, increasing labor scarcity is typically the only source of retardation in the rate of urbanization, even when such models are expanded for demoeconomic simulation [see Sanderson (1980)]. The rise in the real wage serves to choke off the rise in the saving rate, to reduce the rate of urban capital accumulation, to retard the rate of increase in new urban job vacancies, and thus to limit urban growth. The ultimate source of the limits to urban growth in models of this sort is agriculture, either through the disappearance of a rural labor surplus and/or through the rise in the relative price of agricultural products—the key wage good in such models.

Nowhere in this account are competing, and potentially voracious, urban “unproductive” investment demands on the national savings pool considered. In addition, while inelastic agricultural land supply insures an eventual constraint on city growth through rising food costs and real wage increases (at least in relatively closed economies), nowhere is the impact of inelastic urban land supply on city rents—another key wage good—and urban cost-of-living considered. Nor, for that matter, is there any concern with inelastic urban land supply on density, crowding, and urban disamenities. Such models say nothing about the costs of urbanization, and are equally silent on the possible limits to urban growth generated within the growing urban sector itself.

It seems likely that more insight into the limits of urban growth might be gained by examining various urban costs that influence the migration decision, on the one hand, and rising urban investment requirements that compete with “productive” capital accumulation, on the other. As we pointed out above, first among these influences are inelastic urban land supplies. Urban land constraints serve to raise rents, augment urban relative to rural living costs, and inhibit city immigration. (At least in the long run immigration would be inhibited. In the short run, higher rents may induce residential construction, a booming demand for unskilled labor used there, and thus more immigration.) To the extent that rising rents and urban disamenities are both caused by density, crowding, and other manifestations of inelastic urban land supplies, then city rents reflect more than living costs alone, but the quality of urban life as well.

The housing-cum-social overhead investment requirements of city growth must also matter. “Unproductive” urban investments of this type, which do not create capacity for future urban employment, may well take priority over those forms of accumulation that do create capacity for future urban employment [Coale and Hoover (1958)]. In any case, unproductive urban investment requirements com-
pete directly with productive capital accumulation. Any model of urban growth must deal with these competing requirements since new urban housing-cum-social overhead investment requirements may serve to check urban growth. Of course, if the housing-cum-social overhead investment is forgone, then housing costs will rise and the quality of urban services fall, further discouraging city immigration. In short, the rise in city cost of living (and of doing business) may impose a limit to urban growth and/or the rise of urban unproductive investment requirements may diminish the rate of productive urban capital accumulation and new urban job vacancies, thus limiting urban growth.

Another set of forces driving migration and city growth are external events over which the economy has little control. Three of these have attracted special attention in the literature. First, in view of the relative capital intensity of city economies, Lewis (1977) has suggested that urban growth breeds foreign capital dependency and that a relative abundance of foreign capital must therefore be a significant determinant of urban growth. This hypothesis encourages the view that foreign capital inflows to the Third World must have played a critical role in accounting for the rapid urban growth up to 1973–74. After all, the Third World was a heavy recipient of foreign capital during the 1960s and early 1970s, reaching an average of about 3 percent of gross domestic product. It follows that the greater austerity in world capital markets since must have played a role in contributing to a city growth slow-down. Second, the relative price of traded goods in world markets should matter. While the relative price of urban manufactures has drifted downwards throughout the past quarter century, the decline was far less dramatic in the 1960s and early 1970s when Third World city growth was especially rapid. Surely we would like to get a clearer understanding of the influence of these world market conditions on migration and city growth. Third, since cities are energy intensive, the rise in the relative price of energy across the 1970s should have contributed to a city growth slow-down. Indeed, can part of the rapid city growth in the Third World in the 1960s and early 1970s be explained by cheap oil?

Another set of forces driving migration and city growth are exogenous internal events. Five of these are most important, and they lie at the heart of the debate. First, has agricultural land scarcity played an important role in pushing labor to the cities? Hoselitz (1955, 1957) thought that arable land shortages generated a powerful rural push, accounting for heavy rural emigration to the cities as well as for "overurbanization" compared with the nineteenth-century developing economies. If the price of agricultural products remains unchanged, then the economics is obvious: land scarcity breeds redundant labor, and the rural emigrants flee to the city. Even though the qualitative result is obvious, it is not at all clear how important agricultural land scarcity has been in accounting for Third World city growth. Furthermore, suppose the price of agricultural products rises in response to the land-scarcity-induced supply contraction? If the demand for agricultural
products is price inelastic, land scarcity may induce a decline in farm labor's marginal physical product, but the greater rise in price will serve to increase the value of farm labor's marginal product. Thus, land scarcity can, at least in theory, create rural pull not push. Second, there are Engel effects to consider. The conventional wisdom has it that "the income elasticity of demand for goods provides clues as to why cities and economic growth invariably seem to accompany each other" [Mohan (1979, pp. 6-7)]. As an economy grows, the proportion spent on food declines, increasing the relative demand for urban-based non-food products. Presumably, the faster the growth, the more rapid the demand shifts toward urban-based activities. While this conventional argument certainly makes sense, one can still doubt its empirical relevance for open economies where world demands and domestic supplies are far more critical in determining structural transformation and urbanization. Third, what about the urban bias? How much of the measured city growth in the Third World is due to price distortions (including commercial and exchange rate policy) which favor urban activities? Fourth, what about government manipulation of capital markets and, fifth, the distribution of public investment between rural and urban areas? While the literature is full of assertions about the magnitude of such policy forces, I am not aware of a single systematic effort to measure the quantitative impact of the urban bias.

However, among these internal events driving migration and city growth, two are especially important: unbalanced total factor productivity advance and Malthusian pressures. By unbalanced total factor productivity advance I mean simply that technological change is usually much more rapid in the modern, urban-based manufacturing sectors than in the traditional, rural-based primary product sectors. Traditional service sectors, of course, also tend to lag behind [Baumol (1967), Baumol, Blackman and Wolff (1985)]. The size of the bias and the magnitude of the unbalancedness vary across countries and over time, but they have been a technological fact of life since Britain's First Industrial Revolution, and in spite of past agricultural revolutions and the contemporary Green Revolution [for Britain, see Floud and McCloskey (1981, vol. 1, ch. 6) and Williamson (1985d, ch. 6); for America, see Williamson and Lindert (1980, ch. 7)].

While such supply-side forces are likely to be at the heart of city growth experience, demand is hardly irrelevant, although it is price elasticities that matter, not income elasticities and Engel effects. After all, if output demand is relatively price elastic, then sectoral total factor productivity growth tends to generate an elastic supply response rather than a relative price decline. This distinction is important since cost-reducing innovations will be passed on to users by falling prices in the inelastic demand case. Thus, the rise in the marginal physical product of factors used in a technologically dynamic sector will be partially offset by price declines, so that marginal value products rise by less, and
resource shifts to the technologically dynamic sector may therefore be minimal, and this includes labor. If, on average, urban sectors tend to have relatively high rates of total factor productivity growth, and if the demand for urban output is relatively price elastic, then final demand shifts towards the dynamic sectors, the derived demand for urban employment is augmented, urban job vacancies are created, migration responds, and city growth takes place. The higher are price elasticities of demand for urban output, the greater is the city growth impact of unbalanced productivity advance favoring the modern sectors. The more open is the economy to foreign trade, the more likely will those conditions be satisfied.

Finally, we have Malthusian forces to consider. Popular accounts of Third World urbanization and city growth often suggest that high rates of population growth lie at the core of the problem. Indeed, not too long ago a World Bank team asserted that "the increase in population growth of the 20th century is the single most important factor distinguishing present and past urbanization" [Beier et al. (1976, p. 365)]. While the assertion sounds plausible, this conventional wisdom has never been adequately tested. Certainly rapid population growth fosters rapid city growth, but it is not clear by how much. In addition, rapid population growth does not offer an explanation for Third World urbanization experience. On the contrary, classical trade theory has shown that a decline in the price of a factor due to an expansion of its supply (in this case the labor force, swollen by population pressure) will result in the relative expansion of the sector in which it is used most intensively. Since we believe that cities tend on average to contain the more capital-intensive activities, rapid population growth should lead to slower urbanization rates, not faster.

How are we to assess the quantitative importance of these forces driving Third World migration and city growth? If economic forces play a critical role in determining rural-urban migration, then urbanization and city growth are surely determined by those same forces. It follows that urbanization and city growth cannot be analyzed without explicit attention given to the interaction between rural and urban labor markets. Furthermore, those labor markets cannot be fully understood without explicit modeling of labor supply and demand forces in both the sending and receiving regions. In short, urbanization and city growth can only be understood by embedding the process in a general equilibrium model.

6.2. Modeling migration and city growth

Two computable general equilibrium (CGE) models have been used recently to address these topics, one constructed to confront the "typical" city growth experience of a group of 40 developing countries [Kelley and Williamson (1984a, 1984b)], and the other constructed to confront city growth experience in India [Becker, Mills and Williamson (1986)]. These two CGEs join a rapidly
A growing family of such models applied to developing countries [e.g. Adelman and Robinson (1978), Dervis, deMelo and Robinson (1982)], and to public finance and trade problems [Shoven and Whalley (1984)]. These two multisectoral models are predominantly neoclassical and contain considerable price endogeneity: households maximize utility and producers maximize profits. Optimizing behavior, however, faces constraints: households have limited endowments, migrants incur costs when they move, firms face capital scarcity, mortgage markets are absent, and so on. Furthermore, factor markets are imperfect and governments obey rules that often conflict with optimal resource allocation. Much of the novelty of the two models, however, lies with the inclusion of spatial variables likely to influence city growth, variables that have been omitted from all previous CGEs. These additional variables include squatter housing as well as more formal urban housing, as well as other non-tradables. In addition, the Indian model develops a two-stream view of migration, a view which is consistent with the critical response to the Todaro model discussed above in Section 5.

This chapter is not the place to elaborate and defend these models, but it does seem appropriate to summarize how they are used as well as the findings that have emerged from their application. I will focus on the Kelley and Williamson (KW) model in what follows.

The KW model was given empirical content by relying on the fictional construct of a “representative developing country”. This is simply the average experience of a large group of developing countries that satisfy the conditions that underlie the model’s theoretical structure. These conditions are: low per capita income in 1960; some per capita income growth over the past 25 years; primary reliance on domestic saving for accumulation; and that the country be a price taken in world markets. The last condition required exclusion of a number of countries (most notably, the OPEC countries) that have an impact on world prices of key resource-intensive products. Fairly extensive historical documentation on economic and demographic variables back to 1960 was necessary for a country to be included in the sample. The 40 countries that met these requirements accounted for about 80 percent of the Third World’s population (excluding China).

Using this data base, the model was used to simulate the urbanization process over the period 1960–80. Five groups of variables are taken as exogenous, each one central to debates regarding the sources of city growth listed above. The first group contains the prices of three main types of commodities that enter world trade—in imported fuels and raw materials; manufactured goods; and primary products. By comparing simulations based on pre-OPEC fuel price trends with those based on post-OPEC fuel price trends, for example, we can assess the influence of fuel scarcity on migration and city growth. The second exogenous variable is agricultural land stock growth, making it possible to assess the role of arable land scarcity as a source of push to the cities. Third, the level of foreign
capital inflow is exogenous, permitting us to explore the hypothesis that foreign capital has been an essential ingredient to migration and city growth. Fourth, rates of total factor productivity growth by sector are determined exogenously and are held to be "unbalanced" in favor of the modern sectors. Finally, the aggregate rates of population and labor force growth are exogenous, making it possible to vary the Malthusian burden so as to assess its importance in contributing to Third World migration and city growth.

Given historical trends in these five sets of exogenous variables, the KW model determines the rate of capital accumulation, investment in dwellings and training and skill development; the patterns of resource allocation and income distribution; the rate of industrialization; and, of course, trends in rural to urban migration and city growth. Although historical documentation is not available for many of these variables, the historical data base is adequate to test the model's ability to track Third World experience across the 1960s and 1970s. It performs quite well overall, and especially well for the migration and city growth experience which motivated the model in the first place. Growth rates of the city population average 4.6 percent per annum in the 40-country sample, and the model prediction is almost exactly the same, 4.7 percent per annum. The model also predicts an acceleration in city growth across the 1960s and early 1970s, conforming to the pre-inflexion point phase along logistic urbanization curves found so commonly in time series. Migration experience is also closely replicated by the model. The rural outmigration rate is predicted at 1.1 percent per annum, while Preston (1979) has estimated the outmigration rate to have been about 1 percent. Similarly, the per annum urban immigration rates are 2.0 in the model and 1.8 percent in Preston's estimates. Finally, the model predicts that 45 percent of the urban population increase can be attributed to immigration, a prediction that lies approximately midway between Preston's estimate of 39.3 percent and Keyfitz's (1980) estimate of 49 percent.

6.3. Understanding the past and projecting the future

There are many ways the KW model can be used, but I will illustrate with one example. First, the model is simulated under the actual exogenous conditions which prevailed up to 1980 as well as the most likely conditions that will prevail up to the year 2000 (the latter based on World Bank projections). This simulation is called BASELINE. Second, the model is simulated under the assumption that the exogenous conditions prior to 1973–74 and the OPEC oil price increase, continued up to 2000. This simulation is called STABLE. The results are reported in Table 11.4.

Whether we focus on BASELINE or STABLE, the model produces the conventional logistic curve with rising then falling city growth and immigration
Table 11.4
Urbanization, urban growth, and immigration in the Third World:
BASELINE and STABLE projections, 1960–2000

<table>
<thead>
<tr>
<th>Year</th>
<th>Population urban (%)</th>
<th>Ratio of percent labor force in mfg. to percent urban</th>
<th>Urban population growth rate (%)</th>
<th>Net rural outmigration rate (%)</th>
<th>Net urban immigration rate (%)</th>
<th>Net immigrant share of urban population increase (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASELINE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1960</td>
<td>32.6</td>
<td>0.374</td>
<td>5.56</td>
<td>1.41</td>
<td>2.91</td>
<td>45.1</td>
</tr>
<tr>
<td>1965</td>
<td>35.8</td>
<td>0.394</td>
<td>4.11</td>
<td>0.82</td>
<td>1.50</td>
<td>48.1</td>
</tr>
<tr>
<td>1970</td>
<td>39.9</td>
<td>0.406</td>
<td>5.15</td>
<td>1.60</td>
<td>2.51</td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>45.2</td>
<td>0.420</td>
<td>4.48</td>
<td>1.38</td>
<td>1.72</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>49.2</td>
<td>0.435</td>
<td>4.10</td>
<td>1.28</td>
<td>1.35</td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>54.3</td>
<td>0.450</td>
<td>4.84</td>
<td>2.23</td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>59.7</td>
<td>0.466</td>
<td>4.67</td>
<td>2.53</td>
<td>1.79</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>64.4</td>
<td>0.478</td>
<td>4.33</td>
<td>2.45</td>
<td>1.41</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>68.4</td>
<td>0.486</td>
<td>3.97</td>
<td>2.21</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>STABLE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1960</td>
<td>32.6</td>
<td>0.374</td>
<td>5.56</td>
<td>1.41</td>
<td>2.91</td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>35.8</td>
<td>0.394</td>
<td>4.11</td>
<td>0.82</td>
<td>1.50</td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td>39.9</td>
<td>0.406</td>
<td>5.15</td>
<td>1.60</td>
<td>2.51</td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>46.2</td>
<td>0.427</td>
<td>5.94</td>
<td>2.67</td>
<td>3.29</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>54.8</td>
<td>0.454</td>
<td>5.97</td>
<td>3.74</td>
<td>3.31</td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>62.9</td>
<td>0.480</td>
<td>4.98</td>
<td>3.75</td>
<td>2.35</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>69.3</td>
<td>0.493</td>
<td>4.32</td>
<td>3.64</td>
<td>1.70</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>74.2</td>
<td>0.499</td>
<td>3.79</td>
<td>3.26</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>77.9</td>
<td>0.501</td>
<td>3.43</td>
<td>2.85</td>
<td>0.84</td>
<td></td>
</tr>
</tbody>
</table>

Source: Kelley and Williamson (1984b, Table 3, p. 430).

rates. Based on these simulations, it certainly appears that urban growth problems will be far less severe by the end of the century, even though there will be no serious diminution in Malthusian population pressures in the Third World over the remainder of this century. Presumably, we will hear fewer complaints from the urban planners, the much-abused term "overurbanization" may disappear from our lexicon, and pessimists' stress on urban environmental decay might lose some of its urgency. Far slower urban growth rates by the year 2000 will make it easier to cope with the accumulated problems associated with decades of rapid urban expansion.

Table 11.4 offers more information of interest to the urban analyst. Consider, for example, the debate on overurbanization that was initiated by Hoselitz in the 1950s. To repeat, his thesis was that urbanization was outpacing industrialization in developing countries in the sense that urban population shares were large in relation to industrial employment shares, at least when compared with the historical performance of currently developed countries. This statistic is also produced by the model in the second column of Table 11.4, although "industry"
is limited to manufacturing. While Hoselitz found support for his thesis in data from the early 1950s, a United Nations study (1980, p. 13) which compared the 1970 with the 1950 statistics, concluded that "urban growth is no longer outpacing industrial growth: if anything, a slight reversal of the over-urbanization tendency has appeared". The model's predictions are consistent with the U.N. finding since the statistic rises, although at a declining rate, throughout the four decades. According to the United Nations, manufacturing has served as the "engine of urbanization" over the past two decades in the Third World, and according to both the BASELINE and the STABLE projections, it will continue to do so in the future.

What happens when the "representative developing country" experiences abrupt changes in the macroeconomic/demographic environment? It turns out that the urban transition can be sharply retarded, halted, or accelerated by changes in the economic/demographic environment over which most Third World countries have little control. This can be seen most clearly in Figure 11.4 where the STABLE and BASELINE predictions are plotted. Note the period from OPEC to the early 1980s. The BASELINE plot tells us what actually happened to Third World city growth in response to changing economic/demographic conditions after 1973–74, while STABLE tells us what would have happened had the pre-OPEC conditions persisted. It appears that the city growth slowdown in the late 1970s was attributable entirely to changes in the economic/demographic environment. Thus, conditions prior to 1973–74 were unusually favorable to rapid city growth. While urban immigration and city growth would have been high and rising in the absence of these favorable conditions, they were even higher due to their presence. Furthermore, the modest urban growth slowdown in the Third World during the remainder of the 1970s was initiated entirely by the appearance of unfavorable exogenous conditions. Had the favorable pre-OPEC conditions persisted after 1973, urban growth rates would have risen still further, making Third World urban problems even more "severe" than in fact was the case.

Which exogenous shocks mattered most? Population pressure? Unbalanced productivity advance? Adverse terms of trade between manufactures and primary products? The relative scarcity of imported fuels and raw materials? Increasing scarcity of arable land? Relative austerity in international capital markets and limits on the availability of foreign capital? Our technique for identifying the impact of various exogenous events on Third World city growth has been to employ historical counterfactuals—what if conditions had been different from those which actually prevailed after 1973–74?

Has arable land scarcity been an important quantitative ingredient of Third World city growth? The answer coming from counterfactual analysis is an unambiguous no. Is a relative abundance of foreign capital a significant determinant of urban growth? Apparently not. Is the increase in population growth in
the twentieth century the "single most important factor" distinguishing present from past urbanization? Once again, the answer is no. Had the Third World experienced the much lower population growth rates that prevailed in the industrial countries in the 1960s, the rate of immigration and urban growth would have still been very high, much higher, for example, than in Britain during the First Industrial Revolution. Furthermore, the forces of "demographic transition" over the next two decades are unlikely to play a major role in city growth in
the future. Finally, what about the pace and character of technological progress? In spite of conventional appeal to Engel effects and overall rates of productivity and income growth to get those effects, it turns out that it is the unbalanced character of that productivity growth which has done most of the work in the past and is likely to do so in the future. The unbalanced rate of technological progress in the Third World was the key condition accounting for the unusually rapid rates of urban growth in the 1960s and 1970s. It follows that if the productivity slow-down currently characterizing the industrial nations spills over into the industrializing Third World during the next two decades, Third World city growth rates will slow down as well.

7. Where do we go from here?

In spite of the attention which Third World migration and city growth has received since the late 1960s, a good share of the questions which really matter have yet to be adequately confronted.

Who gains and who loses from city immigration? While there has been a long-standing interest in the growth–distribution trade-off, and while urbanization has always played a role in such trade-off debates, nowhere in the literature can we find an adequate quantitative assessment of the distributional impact of city immigration itself. This is a rather puzzling state of affairs given the large and growing literature on the distributional implications of international migrations, especially so given that rural–urban migrations in the Third World far exceed in numbers contemporary international migrations, including Latin illegals over American borders today and guestworkers over European borders in the 1960s and 1970s. Who gains and who loses from these rural–urban migrations? The development literature has focused almost entirely on the migrants themselves, who, of course, gain from the move. But what about urban residents with whom they compete as substitutes in production and employment? What about skilled urban labor and capital with whom they are complements? And what about those left behind in rural areas? While it is likely that landlords lose, what about the landless and small owner-occupiers left behind? Here, the net impact may be more complex and subtle. And when we have finally assessed the impact of emigration on factor productivity of those rural resources left behind, what about urban–rural remittances from cityward movers to rural stayers? The size of those (net) remittances and their function has still to be properly documented, and distributional assessment of the impact of migration is likely to hinge on their magnitude.

Economic–demographic interactions need more serious empirical attention. Rural emigration is selective, especially of the young adult. What does this selectivity bias do to the stock of females at fertility risk in the countryside, and
thus to rural fertility rates and the demographic transition? And if high rural fertility is driven in part by parental "pension motives", how do parents respond to the high and rising emigration of their children? Does the emigration of children imply a default on parental investment in old age? Do parents respond by having fewer children? And what does the migrant selectivity bias do to the spatial distribution of the welfare burden? If the very old, the very young, and the disabled are left behind in rural area, what social systems can and should be devised to maintain their support to replace the income of the emigrants? And who should be targeted in any transfer to the rural poor left behind? The young (via more education, better health facilities and improved nutrition), who are themselves encouraged to migrate after such transfers have been made? Or the old and disabled?

What role does the "informal service vector" serve in the cities? The literature has always contained a persistent bias against services, viewing them as a parasitic activity with little social value. It is certainly hard to understand how such a view could persist given that services have accounted for 40 or 50 percent of national income since Britain began the First Industrial Revolution in 1780. Yet, Marx gave it little attention when writing in the midst of the First Industrial Revolution, and contemporary development economists tend to do the same. The best example of this is the so-called "informal service sector", labeled the "residuum" in Victorian England. The Todaro model views that sector as a holding area for the reserve army, where the employed are at the subsistence margin, and the unemployed are supported by transfers. Research over the past decade has shown this view to be totally inappropriate, but we still need to learn far more about how this sector works, the demands that it satisfies, and its link via the employment multiplier to the city's export base, namely manufacturing.

While the development literature has devoted an enormous amount of attention to empirical work on migration in the Third World, almost all of that work has had a micro focus. We have, as a result, learned a great deal about who migrates and about which components of the total quality of life induce the move - cheap housing, the quality of public services, current job vacancies, good wages, and good chances for higher paying jobs. We have also learned much about the elasticity of migration to such incentives by type of migrant. Yet, there has been almost no work done on the macro forces determining the magnitude of the flows, the rural-urban wage gap, and the intra-urban wage structure Computable general equilibrium models offer one way to attack these problems, but we need alternative approaches to improve our understanding of the macroeconomic and macroedemographic forces driving city migration and the structure of wages.

The literature is full of anecdote and assertion regarding the impact of policy on Third World migration and city growth. Yet, there has been almost no effort to quantify the impact of various policies, including any overall "urban bias".
Which policies are most effective in achieving which targets? What are the trade-offs between city migration and city size targets, on the one hand, and distribution and growth goals, on the other? Once again, computable general equilibrium models offer one useful way to attack these questions.

Finally, development economists would be well advised to take another careful look at history. There is very little that is unique about Third World migration and city growth. The overall quantitative patterns appear to replicate those of Britain during the First Industrial Revolution and, I suspect, most of the other nineteenth-century industrial revolutions which followed. Many of the debates over how urban labor markets operate, whether there is overurbanization, whether migration is driven by rural push or urban pull, who gains from migration, the determinants of wage gaps, and optimal investment in city social overhead were raised as early as the 1830s in Britain. It is folly for economists interested in Third World migration and city growth to remain ignorant of those debates and to treat the contemporary Third World experience as unique.

References

Becker, C.M., Williamson, J.G. and Mills, E.S. (in progress) *Indian urbanization and economic growth: Modeling the past, the present and the future.*


Ch. 11: Migration and Urbanization


